

Table des matières

Remerciements	3
Table des matières	4
Introduction	8
Chapitre I	12
Analyse bibliographique sur le goyavier-fraise	
I- Connaissance de la plante	12
1- Situation taxonomique	12
2- Origine, distribution, écologie	12
3- Botanique, génétique	13
4- Phénologie	14
5- Biologie florale et reproduction	14
6- Le goyavier-fraise, plante envahissante	15
II- Agronomie, valorisation	15
1- Multiplication	15
2- Techniques culturales	16
3- Ravageurs et maladies	17
4- Composition du fruit, utilisations	17
Chapitre II	19
Premiers éléments sur la phénologie du goyavier-fraise et son comportement en culture. Conséquences pour la construction d'un modèle d'élaboration du nombre de fruits et de leur calibre.	
I- La phénologie du goyavier-fraise en milieu naturel à la Réunion	20
II- L'azote déclenche l'émission de pousses florifères chez le goyavier-fraise	30
III- Un modèle général de l'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise.	43
1- Conséquences des résultats précédents	43
1.1- Conséquences agronomiques	44
1.2- Conséquences sur la structure du modèle	45
2- Structure du modèle général de l'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise	46
Chapitre III	50
Etude et modélisation de la floraison du goyavier-fraise	
I- Application d'un modèle de floraison au goyavier-fraise. Analyse des écarts entre les sorties du modèle et les données	51

	5
II- Un modèle stochastique pour décrire la floraison de plants à floraison asynchrone	59
III- Synthèse et conclusions pour le modèle de floraison du goyavier-fraise	79
Chapitre IV	82
Etude de la fécondation et de la fructification	
I- Introduction	82
II- Essai de pollinisation contrôlée	83
1- Matériels et méthodes	83
2- Résultats et discussion	85
3- Conséquences pour la modélisation de la pollinisation et de la fécondation-fructification	88
III- Effet de différents facteurs sur la distribution du nombre de graines par fruit à l'échelle du plant	89
1- Le modèle général de distribution du nombre de graines par fruit	89
1.1- La distribution du nombre d'ovules fécondés par fleur	90
1.2- La fructification	91
1.3- Le modèle général de distribution des nombres de graines par fruit	93
2- Effet de l'origine génétique du pollen sur les paramètres du modèle général, et simplification du modèle.	93
2.1- Matériels et méthodes	93
2.2- Résultats	94
2.2.1- Estimation des paramètres du modèle général complet	94
2.2.2- Simplification du modèle général	96
2.3- Conclusions	98
3- Explicitation des paramètres du modèle général: le modèle en λ	98
3.1- Construction du modèle en λ	98
3.2- Estimation des paramètres du modèle en λ	100
3.3- Simplification du modèle en λ	101
4- Discussion - Conclusions	103
Chapitre V	106
Etude et modélisation de la croissance du fruit du goyavier-fraise	
I- Couplage des facteurs de croissance et des erreurs statistiques à un modèle de croissance pour prévoir la distribution des poids de fruit	107
II- Synthèse de la partie précédente et conclusion pour le modèle de prédiction de la distribution des poids de fruit à maturité.	128
III- Modèle de prédiction de la distribution des dates de récolte	129
1- Introduction	129
2- Matériels et méthodes	130
3- Résultats et discussion	132

Chapitre VI	139
Discussion et conclusion	
I- La problématique	139
II- La démarche	140
1- La domestication basée sur une approche agronomique	140
2- La démarche expérimentale et la modélisation	142
III- Analyse critique des principaux résultats	144
1- Phénologie du goyavier-fraise en conditions naturelles	144
2- Déclenchement de pousses florifères par un apport d'azote	145
3- La floraison	145
4- La fécondation et la fructification	147
5- La croissance du fruit	147
IV- Les retombées pratiques de cette étude	150
V- La construction du modèle global d'élaboration du nombre de fruits et de leur calibre	152
VI- La démarche est-elle généralisable?	153
1- La domestication par une approche agronomique	153
2- Généralisation des modèles proposés	155
Littérature citée	157
Annexes	169
Annexe 1: Relations entre la température et la durée des intervalles fertilisation - demi floraison et demi floraison - demi récolte	170
1- Matériels et méthodes	170
1.1- Le modèle thermique	170
1.2- Les données sur les cycles de production	171
1.3- Les données thermiques	171
1.4- Analyses statistiques	172
1.4.1- Estimation de la température de base	172
1.4.2- Validation des modèles	173
2- Résultats	173
3- Discussion	177
Annexe 2: Détermination du nombre d'ovules dans les fleurs	181
1- Matériels et méthodes	181
2- Résultats et discussion	181
Annexe 3: Calcul de la vraisemblance d'un déroulement de floraison observé	185
1- Utilisation d'une distribution multinomiale.	185
2- Utilisation des probabilités conditionnelles	185
3- Conclusion	187

Annexe 4: Tableaux de présentation synthétique des modèles	188
Annexe 5: Planches photos	193
Liste des figures	195
Liste des tableaux	197
Résumé	199
Summary	200

Introduction

Les Hauts humides de l'île de la Réunion sont soumis à des conditions pédoclimatiques difficiles caractérisées par une pluviométrie annuelle importante (5000-8000 mm), des températures plus fraîches que sur le littoral, et des andosols désaturés peu mécanisables et de faible valeur agricole (Raunet, 1991). Les exploitations agricoles y sont majoritairement familiales, avec une SAU généralement comprise entre 5 et 10 ha. Le faire valoir direct est dominant. Les systèmes de production sont variés, de la monoculture de la canne à sucre à des systèmes entièrement diversifiés. La culture principale est la canne à sucre (82% de la SAU dans les Hauts de l'est de l'île), mais sur des terres de qualité très moyenne, difficilement mécanisables et souvent isolées par manque de route correcte, cette culture est à la limite de la rentabilité. La survie des exploitations est alors liée à un processus de diversification partielle ou totale des cultures (Galan, 1995). Les conditions pédoclimatiques limitent les possibilités de choix et les cultures fruitières apparaissent les plus opportunes. Mais les espèces traditionnelles telles que l'ananas, le litchi, les agrumes ou la banane se trouvent confrontées à des problèmes de production non rencontrés dans des zones plus favorables à leur culture.

Ce constat a conduit le Commissariat à l'Aménagement des Hauts, avec le soutien financier des collectivités locales, de l'Etat et de l'Europe, à demander au Département des productions fruitières et horticolas du Centre de coopération internationale en recherche agronomique pour le développement (Cirad-Flhor) en 1992 d'engager une opération d'expérimentation en milieu réel avec comme objectif l'élaboration de références techniques et économiques pour des cultures innovantes de diversification adaptées aux conditions pédoclimatiques et socio-économiques des exploitations de la zone.

Les espèces innovantes étudiées dans ce cadre ont été choisies sur la base de quatre critères relatifs à l'adaptation aux conditions pédoclimatiques et aux structures des exploitations de la zone, ainsi qu'au potentiel commercial:

- bonne adaptation aux conditions pédoclimatiques des Hauts humides,
- existence de débouchés commerciaux,
- connaissance du produit par la population afin d'éviter l'effet de nouveauté,
- si possible, complémentarité des périodes et des temps de travaux requis par la nouvelle culture avec ceux de la canne à sucre.

Les trois premiers critères nous ont amené à rechercher parmi des espèces déjà présentes à l'état spontané dans la zone d'étude, plutôt que d'en introduire de nouvelles. Le goyavier-fraise (*Psidium cattleianum* Sabine) a été retenu pour sa bonne adaptation à la zone et son potentiel économique. C'est une espèce arbustive qui a été introduite à la Réunion au début du 19^{ème} siècle. Il s'est rapidement développé dans les écologies humides entre 0 et 1300 m d'altitude et son aire de répartition est estimée à 12 000 ha (Raunet, 1991). Les deux variétés botaniques sont présentes sur l'île: *P. cattleianum* var. *cattleianum*, à fruits rouges, et *P. cattleianum* var. *lucidum* à fruits jaunes (Annexe 5). La première est la plus fréquente et intéresse les transformateurs grâce à son arôme et à sa couleur rouge. Elle a été retenue pour cette opération de recherche et, sauf précision particulière, le terme goyavier-fraise sera relatif à cette variété botanique dans la suite de ce travail. Il produit des baies très appréciées de la population réunionnaise qui les récolte en milieu naturel. Des récoltes plus conséquentes sont réalisées par des cueilleurs saisonniers et alimentent l'industrie agro-alimentaire (confiture,

sorbets, boissons alcoolisées), ce qui confère à ce fruit une importance économique locale (Lapierre, 1996).

Le goyavier-fraise n'est pas cultivé dans le monde et très peu de références agronomiques existent. Cette espèce peut donc être considérée comme une espèce sauvage, et notre démarche globale comme sa domestication, c'est à dire le passage de l'état spontané à l'état cultivé. Depuis quelques décennies, l'identification et la domestication d'espèces sauvages à potentiel agricole se sont développées comme moyens pour élargir la gamme d'espèces cultivées, avec des retombées économiques attendues au niveau local, national ou international (Kupzow, 1980; Nerd *et al.*, 1990). Les espèces sont identifiées pour leur potentiel dans différents domaines tels que la production de fruits (Jankiewicz *et al.*, 1986; Schultes, 1989; Reid et Gast, 1993; Clement et da Silva Filho, 1994) ou de légumes (Wagenvoort *et al.*, 1989), l'industrie (von Witzke et Röbbelen, 1988; Hof *et al.*, 1999), ou la médecine (Do Ngoc *et al.*, 1995). La valorisation des ressources génétiques locales à buts multiples (fruits, fourrage, bois) dans des systèmes agro-forestiers relève également de stratégies de domestication (Quevedo et Zuluaga, 1994; Leakey et Simons, 1998). De manière générale, la sélection génétique en fonction de critères d'intérêt nutritionnel, pharmaceutique, industriel ou agronomique bien définis est une des premières étapes du processus de domestication. Les études agronomiques sont ensuite conduites sur le matériel végétal sélectionné.

Plusieurs auteurs notent que le goyavier-fraise se reproduit fidèlement par semis (Fouqué, 1972; Popenoe, 1974). Cette caractéristique est aussi valide à la Réunion où il existe peu de variabilité phénotypique sur les caractères végétatifs (port de l'arbre, forme et taille des feuilles). Des caractères agronomiques importants tels que le nombre de fruits par plant et le calibre du fruit sont variables à différents niveaux (entre fruits d'un même plant pour le calibre (Fig. 1), entre plants pour une même récolte et entre récoltes pour un même plant), et cette variabilité semble peu dépendante du génotype. La couleur du fruit est peu variable dans des conditions de production données et paraît également peu liée au génotype. Ces éléments suggèrent que le gain génétique sur ces caractères apporté par un travail de sélection risque d'être faible, avec des résultats attendus au mieux à moyen terme. Par contre, le gain agronomique, lié à l'application de techniques culturales adaptées, peut être élevé et rapide à condition d'explicitier la variabilité observée, et que cette dernière soit au moins partiellement due à des facteurs sur lesquels l'agriculteur peut agir. Nous avons donc engagé la domestication du goyavier-fraise par une approche agronomique sur du matériel végétal de semis en cherchant à expliciter la variabilité observée sur le nombre de fruits produits et sur leur calibre. L'objectif est la production économiquement viable de fruits de qualité par des techniques culturales adaptées.

La notion de rendement commercial, qui est la variable d'intérêt pour l'agriculteur, est de plus en plus intégrée dans les études en arboriculture fruitière (Doyle *et al.*, 1989; Atkins, 1990; Agostini, 1995; Stover *et al.*, 2001). Il est défini comme un rendement quantitatif pondéré par une fonction indicée sur la qualité (Agostini, 1995). Il fait référence d'une part au nombre de fruits produits, et d'autre part à un ou plusieurs critères de qualité des fruits dont la rémunération est différenciée. Pour beaucoup de cultures fruitières, le calibre du fruit est un critère majeur de qualité. Chez le goyavier-fraise, le calibre du fruit a des retombées économiques importantes et à deux niveaux: une meilleure valorisation commerciale des fruits de gros calibre, notamment en frais, et une baisse des coûts de production. La récolte demande en effet beaucoup de main-d'œuvre et représente 65% des coûts de production. La cadence de récolte, exprimée en kilos de fruits récoltés par personne et par unité de temps,

augmente avec le calibre des fruits, ce qui diminue les temps de récolte pour une production donnée. Une valeur moyenne du calibre du fruit apparaît cependant insuffisante compte tenu de la variabilité observée chez le goyavier-fraise (Fig. 1), comme chez d'autres fruits (Habib *et al.*, 1991; Génard et Bruchou, 1993; Volz *et al.*, 1995). Le calibre individuel des fruits, ou leur distribution, sont des variables plus pertinentes (Lescourret *et al.*, 1998).

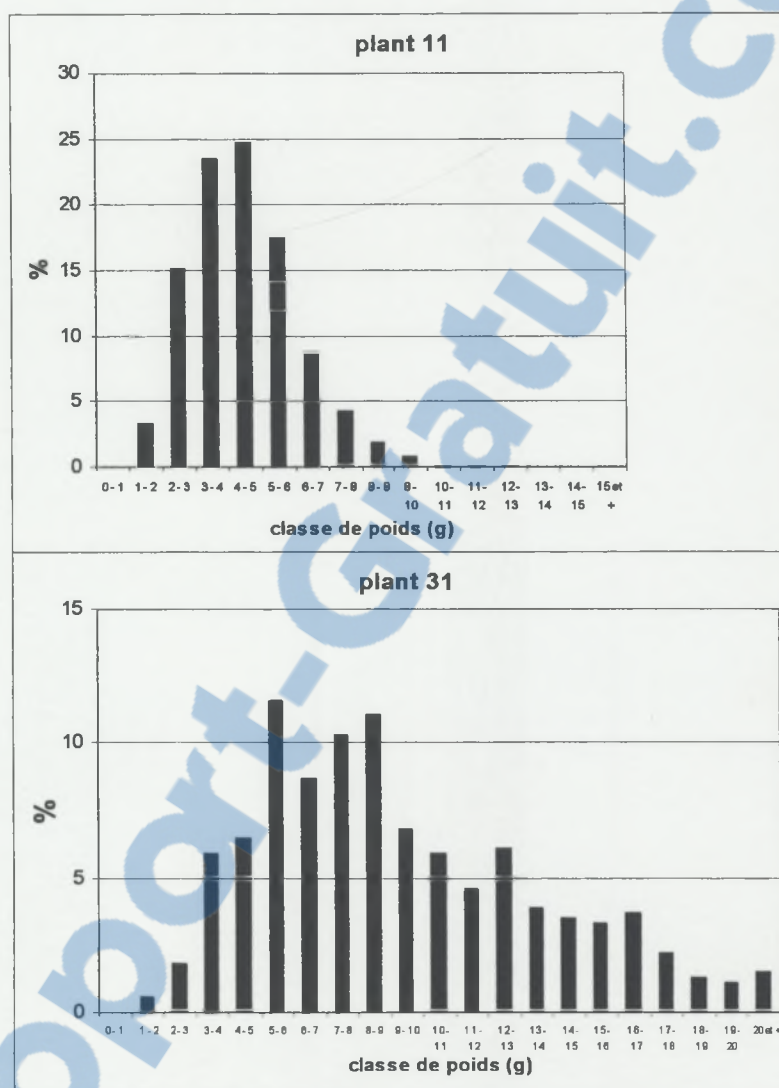


Figure 1: Distribution en classes de 1 g du calibre de l'ensemble des fruits récoltés sur 2 plants d'une même parcelle et un même cycle de production. Plant 11: 2564 fruits; plant 31: 544 fruits.

Deux processus non indépendants doivent donc être distingués dans notre démarche: l'élaboration du nombre de fruits et l'élaboration de leur calibre individuel. Ces deux variables sont le résultat d'une succession de mécanismes biologiques complexes et dépendants, de l'émergence des pousses florifères à la récolte. La modélisation est un outil qui peut permettre de représenter ce système complexe et les effets de divers facteurs, dont l'action de l'agriculteur, sur ses sorties. L'agriculteur peut en effet modifier par des opérations techniques le déroulement d'un mécanisme biologique, et donc les sorties du système.

Un minimum de connaissances est nécessaire pour appréhender cette complexité. Face au manque de références bibliographiques dans ce domaine sur le goyavier-fraise, des expérimentations ont été développées en vue de la construction d'un modèle d'élaboration du nombre de fruits et de leur calibre. Elles s'inspirent de la démarche d'élaboration du rendement conçue initialement sur les cultures annuelles (Meynard et Sébillote, 1994). Cette démarche, qui a été adaptée à une autre culture pérenne, le kiwi (Agostini, 1995; Lescourret *et al.*, 1998), consiste à identifier les étapes déterminantes de l'élaboration du nombre de fruits et du calibre, et les facteurs qui les influencent. Les étapes et les facteurs limitants peuvent ainsi être déterminés. Pour une espèce pérenne, le cycle phénologique est une entrée privilégiée pour identifier ces étapes: mise en place des boutons floraux, floraison, pollinisation, fructification, et croissance du fruit.

L'objectif de ce travail est de contribuer à la réalisation d'un modèle d'élaboration du nombre de fruits et de leur calibre pour le goyavier-fraise, en identifiant les processus et les facteurs limitants, et en proposant des modèles pour certaines étapes du cycle de production.

Dans un premier chapitre, une analyse bibliographique fait le point sur les connaissances existant sur le goyavier-fraise. Cette synthèse paraît nécessaire afin d'avoir une vue d'ensemble la plus complète possible sur cette espèce peu étudiée et peu connue. Dans un deuxième chapitre, son cycle phénologique en milieu naturel à la Réunion est décrit. Puis il est montré que des cycles de production peuvent être déclenchés en culture par un apport d'azote, ce qui a des retombées agronomiques et économiques importantes. Ces éléments permettent de préciser l'organisation générale d'un modèle d'élaboration du nombre de fruits et de leur calibre dans un objectif de prédiction et d'aide à la décision pour l'agriculteur. Les chapitres suivants sont consacrés aux composantes de ce modèle global qui ont pu être traitées dans le cadre de ce travail de thèse: la floraison, la fécondation-fructification et la croissance du fruit. Un dernier chapitre est consacré à la discussion et aux conclusions. Des études particulières, dont les résultats sont utilisés dans ce travail, sont présentées en annexes 1 et 2.

Chapitre I

Analyse bibliographique sur le goyavier-fraise

La première partie de cette analyse bibliographique est consacrée à la connaissance générale de la plante et de sa biologie. Dans la seconde partie sont présentés les éléments relatifs à la valorisation, agronomique et après-récolte, du goyavier-fraise.

I- Connaissance de la plante

1- Situation taxonomique

Le goyavier-fraise, *Psidium cattleianum* Sabine, appartient à la famille des Myrtaceae, sous-famille des Myrtoideae (32 genres) caractérisée par des fruits charnus. Cette sous-famille comporte quelques espèces d'importance économique dont la goyave, *Psidium guajava*. La seconde sous-famille, les Leptospermoideae (40 genres), est caractérisée par des espèces australiennes avec des capsules sèches comme fruit. Elle comprend en particulier les *Eucalyptus*.

P. cattleianum, ou *P. cattleyanum*, a plusieurs synonymes, et la nomenclature n'est pas clairement établie (voir Fosberg, 1941; Diong, 1998): *P. littorale* Raddi, *P. indicum* Bojer, *P. variabile* Berg, *P. sinense* D.C., *P. obovatum* Mart., *P. coriaceum* Mart. ex. Berg, *P. araca* hort. not Raddi, *Guayava cattleyana* Kuntze, *G. obovata* Kuntze (Fosberg, 1941; Fouqué, 1972; Lavergne, 1981; Maggs, 1985).

Deux variétés botaniques sont identifiées, principalement en fonction de la couleur des fruits:

- *P. cattleianum* var. *lucidum* (*P. littorale* Raddi var. *littorale* Bailey) à fruits de couleur jaune soufre
- *P. cattleianum* var. *cattleianum* (*P. littorale* Raddi var. *longipes* (Berg) Fosb.) à fruits de couleur rouge pourpre, plus petits que ceux de la variété précédente.

Dans sa zone d'origine au Brésil, des populations de chaque variété botanique provenant de régions différentes se distinguent sur des caractères végétatifs et du fruit (Raseira et Raseira, 1996). Les chercheurs brésiliens ont sélectionné dans cette diversité une variété horticole, 'Ya-Cy', à port bas et fruits jaunes (Raseira et Raseira, 1994, 1996). Popenoe (1974) et Scott (1990) citent des types différenciés par la taille et la forme des fruits: ronds, ovoïdes ou pyriformes. A la Réunion, seul le type à fruit rond est présent.

2- Origine, distribution, écologie

Le goyavier-fraise est originaire du sud-est du Brésil et du nord de l'Uruguay. Il a été introduit très tôt en Chine, probablement par les portugais (Popenoe, 1974). Il a été cultivé avec succès en Angleterre au début du 19^{ème} siècle par le botaniste Cattley. Des différentes étapes de sa

dispersion, il a gardé son nom scientifique et quelques noms vernaculaires: goyave de Chine, goyave de Cattley (Diong, 1998). Il a été peu à peu introduit dans toutes les zones tropicales et subtropicales du globe où il s'est facilement naturalisé.

Au sein du genre *Psidium*, le goyavier-fraise est l'espèce la plus résistante au froid. Introduit avec succès dans des régions à climat de type méditerranéen, il a supporté des températures négatives, mais il ne résisterait pas à des températures inférieures à -7°C (Popenoe, 1974; Wutscher et Shull, 1975; Vogel, 1982). Il préfère les climats humides et supporte mal la sécheresse. C'est une espèce héliophile. Son aire géographique de développement correspond à celle de l'oranger (Popenoe, 1974).

Le goyavier-fraise peut s'adapter à une large gamme de sols (Diong, 1998), mais il préfère les sols argilo-sableux riches et bien drainés (Popenoe, 1974). A la Réunion, il s'installe comme arbuste pionnier sur les laves des coulées modernes du Grand-Brûlé en zone très humide (Cadet, 1980).

3- Botanique, génétique

Le goyavier-fraise est un arbuste ou petit arbre pouvant atteindre 8m de hauteur (Scott, 1990). Son écorce est lisse, de couleur brune, verdâtre ou rosâtre, et se délite régulièrement. Les nouvelles pousses sont cylindriques.

Les feuilles sont simples, entières, pétiolées, et disposées opposées décussées. Le limbe est obovale, de 4.5 à 12 cm de longueur sur 2 à 6 cm de largeur (Scott, 1990), obtus à légèrement acuminé au sommet, coriace, vert foncé et brillant sur la face supérieure, vert plus clair sur la face inférieure. Les jeunes feuilles sont anthocyanées.

Les fleurs sont situées sur les nouvelles pousses. Elles sont axillaires et solitaires. Elles sont pédonculées. Les bractéoles sont ovales et caduques. Les boutons floraux globuleux sont rouges à leur apparition puis deviennent verts ensuite. Le calice est composé de 4 à 5 sépales obtus, persistants, longs de 5 à 6 mm. La corolle est composée de 4 à 5 pétales blancs, obovales. Les étamines sont nombreuses, de 256 à 480 en moyenne (Raseira et Raseira, 1996). L'ovaire est infère, en général tétraloculaire et possède un style long de 6 à 10 mm. Le stigmate est entier, discoïde.

Le fruit est une baie globuleuse ou obovoïde de 2.5 à 5 cm de diamètre, à épicarpe fin et coloré. Les graines dures, nombreuses, jaunâtres, réniformes, longues de 2 à 3 mm, sont incluses dans une pulpe molle et blanche. Les fruits sont parfumés et sucrés à maturité et leur arôme rappelle alors celui de la fraise, d'où le nom vernaculaire de l'espèce.

Le goyavier-fraise se reproduit fidèlement par semis (Fouqué, 1972; Popenoe, 1974; Hoyos, 1989). Les travaux de Raseira et Raseira (1996) sur la reproduction du goyavier-fraise suggèrent une apomixie partielle ou totale, qui expliquerait cette multiplication fidèle par semis.

Le goyavier-fraise est fortement polyploïde. Le niveau de ploïdie varie selon la variété botanique, mais aussi selon les auteurs. Le nombre de chromosomes de base de l'espèce diploïde *P. guajava* est $x=11$ (Atchison, 1947; Hirano et Nakasone, 1969a). *P. cattleianum* var. *cattleianum* serait heptaploïde, $2n=77$ (Hirano et Nakasone, 1969a; Srivastava, 1978;

Singhal *et al.*, 1985 (matériel végétal: *P. coriaceum*)) ou octaploïde, $2n=88$ (Atchison, 1947; Smith-White, 1948). *P. cattleianum* var. *lucidum* serait hexaploïde, $2n=66$ (Hirano et Nakasone, 1969a) ou octaploïde (Atchison, 1947). Une explication à cette disparité de résultats pourrait être apportée par les résultats de Raseira et Raseira (1996) qui ont trouvé des niveaux de ploïdie différents, mais toujours élevés (hexa- à octa-), sur différentes populations des deux formes botaniques au Brésil, alors que les études précédentes ont été réalisées sur un nombre limité d'individus. Hirano et Nakasone (1969a) s'interrogent sur la bonne capacité reproductrice de *P. cattleianum* var. *cattleianum* et l'uniformité des semis alors que cette espèce est fortement polyploïde et que l'heptaploïdie est un caractère instable. Quelques cas d'aneuploïdie ont été rapportés (Raseira et Raseira, 1996). Les nombreuses anomalies observées au moment de la méiose ont été attribuées à la forte polyploïdie (Singhal *et al.*, 1985; Raseira et Raseira, 1996).

4- Phénologie

En milieu naturel, le goyavier-fraise n'émet en général qu'une seule pousse florifère dans l'année, au début de l'été (Huenneke et Vitousek, 1990; Raseira et Raseira, 1996). A Hawaï, un repos dans l'activité végétative et reproductrice est noté en hiver. Le rythme de développement est indépendant de l'altitude dans la zone étudiée, entre 150 et 762 m (Huenneke et Vitousek, 1990). En parcelle cultivée, le goyavier-fraise a en général deux pousses florifères dans l'année (Teaotia *et al.*, 1970; Subramanyam *et al.*, 1992; Raseira et Raseira, 1996), mais leur rythme varie suivant les cas. Au Brésil, trois floraisons successives sont observées certaines années (Raseira et Raseira, 1996). Le goyavier-fraise ne serait pas sensible à la photopériode en hiver, poussant de façon équivalente avec une photopériode de jour long ou de jour court (Broschat et Donselman, 1983).

Teaotia *et al.* (1970) observent 59 jours entre l'apparition du bouton floral et la floraison. Subramanyam *et al.* (1992) indiquent une durée de 36 jours entre le bourgeon floral de 2 mm et la floraison.

5- Biologie florale et reproduction

Plusieurs études ont été menées, surtout en Inde, sur la biologie florale, le déroulement de la floraison et la pollinisation chez plusieurs espèces de *Psidium* comme préalable à des programmes de croisements contrôlés avec la goyave (Hirano et Nakasone, 1969b; Seth, 1970; Teaotia *et al.*, 1970; Chezhiyan et Shanker, 1983; Chezhiyan, 1988; Subramanyam *et al.*, 1992). La variété *lucidum* a été plus étudiée que la variété *cattleianum* dans ces travaux. Les résultats sont souvent variables d'un auteur à l'autre. L'anthèse a lieu le matin. Les anthères s'ouvrent entre 9 et 11 heures. Le stigmate serait réceptif dès l'ouverture de la fleur, et jusqu'à 24 heures après (Teaotia *et al.*, 1970), ou de 12 à 48 heures après l'anthèse (Chezhiyan, 1988).

Le pollen est de couleur jaunâtre, composé de grains triangulaires tricolpores de petite taille, environ 23 μm . Sa viabilité, mesurée par un test à l'acétocarmin, est généralement faible, 20 à 32% (Hirano et Nakasone, 1969b; Singhal *et al.*, 1985; Raseira et Raseira, 1996), ce que Raseira et Raseira (1996) et Singhal *et al.* (1985) rapprochent des anomalies observées à la méiose, en relation avec la forte polyploïdie. Teaotia *et al.* (1970) rapportent 75% de viabilité

par ce même test. La longévité du pollen est courte dans les conditions du champ, de l'ordre de 8 heures, et il reste viable jusqu'à 90 jours à 0% d'humidité et 4.5 °C (Seth, 1970).

Bien que Chezhiyan (1988) conclue à l'auto-incompatibilité de *P. cattleianum* var. *lucidum*, les deux variétés botaniques de goyavier-fraise sont considérées auto-compatibles par plusieurs auteurs (Hirano et Nakasone, 1969b; Teatota *et al.*, 1970; Raseira et Raseira, 1996). Ces deux derniers auteurs notent toutefois que la fructification est plus élevée en pollinisation libre qu'en autopolinisation passive. Hirano et Nakasone (1969b) obtiennent des fruits en croisant les deux variétés botaniques, mais relativisent l'occurrence de tels croisements car il n'a jamais été fait mention d'hybrides dans les populations naturelles où cohabitent les deux variétés.



6- Le goyavier-fraise, plante envahissante

Le goyavier-fraise s'est naturalisé dans de nombreux milieux insulaires subtropicaux du globe où il a un comportement de peste végétale (Devine, 1977; Lorence et Sussman, 1986; Huenneke et Vitousek, 1990; MacDonald *et al.*, 1991; Fleischmann, 1997). Sa production importante de graines au taux de germination élevé, son adaptation à une large gamme écologique, sa facilité de dissémination par les oiseaux et les mammifères, et l'absence de ravageur en font une espèce agressive et envahissante qui colonise différents milieux et empêche le développement de la flore indigène (Huenneke et Vitousek, 1990). Une meilleure efficacité de la photosynthèse favoriserait également le développement rapide du goyavier-fraise par rapport à des espèces indigènes (Dulai *et al.*, 1998).

Des programmes de lutte chimique contre *P. cattleianum* ont été mis en œuvre dans plusieurs pays, mais l'espèce est résistante à la plupart des herbicides classiques. Le 2,4-D, efficace contre *P. guajava*, est inefficace contre *P. cattleianum* (Anonyme, 1981). L'utilisation de piclorame a donné des résultats mitigés (Devine, 1977). L'absorption de cette molécule par les feuilles est cependant améliorée par adjonction de sulfate d'ammonium (Wilson et Nishimoto, 1975). Le glyphosate concentré, appliqué sur des blessures faites sur le tronc, ou le tébuthiuron en épandage de 2 à 6 kg.ha⁻¹ donnent de bons résultats pour l'éliminer (Motooka *et al.*, 1983, 1989).

La lutte biologique est envisagée depuis une dizaine d'années, en particulier à Hawaï. Des ravageurs spécifiques sont recherchés dans sa zone d'origine au Brésil (Wikler *et al.*, 1994). Deux hyménoptères ainsi identifiés semblent prometteurs pour un contrôle biologique de l'espèce (Wikler *et al.*, 1996; Pedrosa-Macedo, 2000). Compte tenu de son importance économique locale, il n'est pas envisagé de lutte biologique à la Réunion.

II- Agronomie, valorisation

1- Multiplication

Le semis est le moyen de multiplication le plus employé. Les graines fraîches germent facilement en 2 à 3 semaines (Fouqué, 1972; Sweet, 1986; Raseira et Raseira, 1996). Les taux de germination obtenus en laboratoire sont moyens à élevés (60 - 80%) en conditions chaudes (18 et 22°C) et humides. A 37°C, la germination est presque nulle. La scarification ou le traitement au froid des graines donne des taux de germination plus faibles que sur des graines

non traitées. Les taux de germination de graines provenant de différentes altitudes (150 à 762 m) sont significativement différents, mais sans relation avec l'altitude du plant ou le poids des graines (Huenneke et Vitousek, 1990). Les graines sont tolérantes à la dessiccation jusqu'à une humidité de 12% du poids frais. Une fois déshydratées en dessous de 19% d'humidité, elles restent viables après un passage de 24 heures dans l'azote liquide (Becwar *et al.*, 1983).

Le bouturage donne en général de bons résultats, de 66 à 95% de reprise, sur des boutures ligneuses d'au moins 1 cm de diamètre (Normand, 1993). Les boutures semi-ligneuses de quelques millimètres de diamètre sont plus difficiles à réaliser et les taux de réussite n'excèdent pas 70% (Nachtigal et Fachinello, 1995). L'ombrage des plants-mères augmente la teneur en amidon dans les feuilles et les branches et diminue les teneurs en sucres réducteurs et en composés phénoliques (Casagrande Junior *et al.*, 1999). Ces effets seraient à l'origine de l'amélioration de l'enracinement des boutures semi-ligneuses venant de plants-mères ombrés (Voltolini et Fachinello, 1997). D'autres facteurs comme le substrat d'enracinement ou l'utilisation d'hormone (Nachtigal et Fachinello, 1995), ou le stade phénologique du plant-mère au moment du prélèvement (Normand, 1993) ont également un effet sur le bouturage.

Le greffage semble très difficile. Raseira et Raseira (1996) font état de taux de réussite de moins de 5% dans différents essais.

2- Techniques culturales

Le goyavier-fraise n'est pas cultivé, au sens agricole du terme, et est surtout considéré comme une plante ornementale ou comme un arbre de cour (Sweet, 1986; Hoyos, 1989). De ce fait, il existe très peu de renseignements concernant sa culture et les quelques données publiées sont en général adaptées d'autres espèces proches comme la goyave, mais ne reposent pas sur des expérimentations. A notre connaissance, seules Salinero Corral et Aguin Casal (1993, 1996) présentent des résultats issus d'essais agronomiques sur le comportement du goyavier-fraise en Galicie, Espagne. Les résultats des travaux sur l'agronomie du goyavier-fraise menés à la Réunion ne sont pas détaillés ici. Ils sont synthétisés par ailleurs dans une fiche technique (Normand et Huet, 2000). Ils montrent l'intérêt agronomique de cette espèce, en particulier grâce à la gestion des cycles de production par des apports d'azote (voir chapitre II-II).

Les distances de plantation recommandées correspondent soit à la haie fruitière, (2.5m - 3.5m) x (0.5m - 1m) (Raseira et Raseira, 1996), soit à la plantation au carré de 3 à 5 m en tous sens (Fouqué, 1972; Popenoe, 1974; Salinero Corral et Aguin Casal, 1996). L'apport de fumier ou d'engrais à la plantation est conseillé. Pour la fumure d'entretien, les doses en éléments fertilisants varient selon les auteurs. Par exemple, Raseira et Raseira (1996) recommandent, pour des plants de 1 à 4 ans, 150 g de sulfate d'ammonium apporté en deux fois, et 100 g de chlorure de potassium.

La croissance est lente. La pratique de la taille n'est guère nécessaire (Popenoe, 1974), sauf pour maintenir les dimensions de la plante (Sweet, 1986) ou enlever les rameaux trop bas (Raseira et Raseira, 1996).

Les premières récoltes apparaissent dès la deuxième ou troisième année et sont alors faibles. Le potentiel de production est atteint vers 5 ou 6 ans (Popenoe, 1974). Salinero Corral et Aguin Casal (1996) ont enregistré en Espagne des productions de 0.2 à 4.3 t.ha⁻¹ entre 3 et 9 ans après la plantation, avec une moyenne d'environ 2.5 t.ha⁻¹ (soit 2.9 kg par plant). Raseira

et Raseira (1996) signalent des productions de 1.8 à 4.9 kg par plant à 3-4 ans. Les fruits arrivent à maturité entre la fin de l'été et l'automne (Popenoe, 1974; Salinero Corral et Aguin Casal, 1996). Comme chez la goyave, les fruits tombent à pleine maturité en laissant le pédoncule attaché au rameau.

P.cattleianum a été testé comme porte-greffe du goyavier, *P. guajava*, car il est résistant au wilt et tolère bien les basses températures. Par rapport à d'autres espèces de *Psidium* testées comme porte-greffe, il confère au plant greffé un meilleur développement en début de plantation (Teaotia et Phogat, 1971). Sur une période de 13 ans, les différences de développement des plants liées au porte-greffe se sont révélées faibles et non significatives (Singh *et al.*, 1976). Les rendements sont toujours plus élevés avec le goyavier-fraise comme porte-greffe, et il n'a pas d'effet sur le calibre du fruit (Teaotia et Phogat, 1971; Singh *et al.*, 1976).

3- Ravageurs et maladies

Peu de ravageurs sont signalés sur le goyavier-fraise. Les principaux sont les mouches des fruits: *Ceratitis capitata* (Vargas *et al.*, 1983), *C. rosa* et *C. catovirii* (Etienne, 1982), *Bactrocera dorsalis* (Vargas *et al.*, 1990), *B. curvipennis* et *B. psidii* (Brun et Chazeau, 1986), *Anastrepha fraterculus* (Raseira et Raseira, 1996), *A. suspensa* (Nguyen *et al.*, 1992). Les populations naturelles de goyavier-fraise attirent de grandes quantités de mouches des fruits durant la période de maturité des fruits (Vargas *et al.*, 1983; Normand *et al.*, 2000).

Diverses espèces de cochenille ont également été relevées sur le goyavier-fraise (Brun et Chazeau, 1986). *Cratopus* sp. peut causer des dégâts aux jeunes pousses et feuilles en altitude à la Réunion (Normand et Huet, 2000).

La sensibilité du goyavier-fraise aux nématodes du genre *Meloïdogyne* est controversée. A Cuba, Cuadra et Quincosa (1982) le trouvent fortement attaqué par des *Meloïdogynes*, en particulier *M. arenaria* et *M. acrita*. En Malaisie, Nik-Masdek et Ahmad-Kamil (1994) le trouvent résistant à *M. incognita*.

Le goyavier-fraise est résistant au wilt du goyavier (Teaotia et Phogat, 1971). Raseira et Raseira (1996) signalent des attaques sporadiques d'antracnose sur fruits.

4- Composition du fruit, utilisations

Les résultats des différentes analyses physico-chimiques réalisées sur le goyavier-fraise sont rassemblées dans le tableau 1. Contrairement à ce qui est couramment pensé, les fruits du goyavier-fraise ne sont pas très riches en vitamine C, avec des teneurs qui varient entre 11 et 50 mg.100g⁻¹ (Wilson, 1980; Gebhardt *et al.*, 1982; Paniandy, 1999). Ces teneurs varient cependant avec les populations et peuvent dépasser 100 mg.100g⁻¹ (Raseira et Raseira, 1996).

Les fruits sont climactériques (Akamine et Goo, 1979; Brown et Wills, 1983). Leur durée de conservation est très limitée, au plus 2 jours à température ambiante. Cette durée peut être allongée à 12 jours par un stade de récolte adapté et une température de conservation de 12 à 14°C (Paniandy *et al.*, 1999). Les fruits sont mangés en frais, ou destinés à la transformation: gelées, confiture, boissons et sorbets (Fouqué, 1972; Popenoe, 1974; Wilson, 1980; Hoyos,

1989). Un "vin" est également élaboré à partir de ce fruit (Lavergne, 1981; Hoyos, 1989). Les anthocyanes présentes dans l'épicarpe des fruits de la sous-espèce *cattleianum* leur confèrent un fort pouvoir colorant rouge (Paniandy, 1999).

Les fruits, comme le reste de la plante, sont utilisés en médecine naturelle. Ils ont des propriétés pectorale, tonique et antidiarrhéique (Lavergne, 1981). Les feuilles contiennent plusieurs principes actifs et sont exemptes d'alcaloïdes et de coumarines (Cecy et Yassumoto, 1973). Elles sont utilisées comme remèdes dans différents pays, contre les hémorragies, les coliques ou les diarrhées (Cecy et Yassumoto, 1973; Lavergne, 1981). La plante serait une source potentielle d'antioxydants (Bahorun, 1998).

Caractéristiques	<i>P. cattleianum</i> var. <i>lucidum</i> ¹	<i>P. cattleianum</i> var. <i>cattleianum</i> ¹	<i>P. cattleianum</i> var. <i>cattleianum</i> ²	<i>P. cattleianum</i> ^{3,z}
Poids du fruit (g)				6.0
Pulpe+peau (%poids total)	89.7	94.0		
Graines (% poids total)	10.3	6.0		
Eau (g.100g ⁻¹)	94.5	94.5	79.4	80.7
Cendres (g.100g ⁻¹)	0.75	0.75	0.77	0.80
Fibres (g.100g ⁻¹)			6.6	6.4
Glucides totaux (g.100g ⁻¹)				17.4
Sucres réducteurs (g.100g ⁻¹)	1.3	1.5		
Sucres non réducteurs (g.100g ⁻¹)	3.0	2.9		
Sucres totaux (g.100g ⁻¹)	4.3	4.2	5.1	
Autres glucides (g.100g ⁻¹)			6.5	
Protéines (g.100g ⁻¹)	0.80		0.88	0.58
Acides aminés (mg.100g ⁻¹)	384			361
Lipides (g.100g ⁻¹)			0.80	0.60
Vitamine C (mg.100g ⁻¹)	11 - 50	11 - 50		37

¹: Wilson, 1980; ²: Popenoe, 1974; ³: Gebhardt *et al.*, 1982.

^z: variété botanique non précisée

Tableau 1: Analyses physico-chimiques de fruits de goyavier-fraise selon différents auteurs.

Chapitre II

Premiers éléments sur la phénologie du goyavier-fraise et son comportement en culture. Conséquences pour la construction d'un modèle d'élaboration du nombre de fruits et de leur calibre.

Le cycle phénologique est une entrée privilégiée pour aborder les étapes de l'élaboration du nombre de fruits et de leur calibre chez une espèce fruitière pérenne. L'étude du cycle phénologique du goyavier-fraise à la Réunion est présentée dans une première partie (Normand, F. et Habib, R. (2001). Phenology of strawberry guava (*Psidium cattleianum*) in Réunion Island. *Journal of Horticultural Science & Biotechnology*, 76, 541–545). Le choix du milieu naturel pour ce suivi a été guidé d'une part par l'absence de parcelle cultivée au début de l'opération de diversification, et d'autre part par l'intérêt d'acquérir des références en milieu naturel auxquelles peuvent être ensuite comparées les données obtenues en parcelle cultivée.

Dans une seconde partie, on montre que l'apport d'azote modifie le cycle phénologique annuel décrit précédemment en déclenchant une pousse florifère (Normand, F. et Habib, R. (2001). Nitrogen fertilisation induces floriferous flush in strawberry guava (*Psidium cattleianum*). *Agronomie*, 21, 735–742).

Ces résultats ont des conséquences agronomiques et économiques intéressantes pour la culture du goyavier-fraise à travers la possibilité de faire deux récoltes par an, et nous amènent à préciser dans une troisième partie la structure du modèle d'élaboration du nombre de fruits et de leur calibre.

I- La phénologie du goyavier-fraise en milieu naturel à la Réunion

Phenology of strawberry guava (*Psidium cattleianum*) in Réunion Island.

By F. NORMAND¹ and R. HABIB²

¹ Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Département des Productions Fruitières et Horticoles, BP 180, 97455 Saint Pierre cedex, Ile de la Réunion, France

² Institut National de la Recherche Agronomique, Laboratoire d'Ecophysiologie et Horticulture, Domaine St Paul, Site Agroparc, F-84914 Avignon cedex 9, France

Journal of Horticultural Science & Biotechnology, 76 (5), (2001), 541–545.

SUMMARY

The phenology of wild stands of strawberry guava (*Psidium cattleianum* Sabine) was studied at 100, 480 and 720 m over 2 years in Réunion Island. The shoots bore flower buds at the basal nodes, and vegetative buds at the apical nodes. Trees produced vegetative growth from September to December, flowers from November to January and fruit from February to June. Shoot growth was earlier at 100 and 720 m the first year, but synchronised in the second year. Flowering and fruiting were also earlier at 100 m. The intervals between shoot emergence and flowering, and flowering and fruit maturity were strongly related to temperature, with base temperatures of 5.5° and 2.6°C, and heat units requirements of 675 and 2,552°C.d, respectively. The numbers of shoots, flowers and fruit per branch varied across sites and years, unrelated to temperature. However, production was significantly higher at 720 m. Production was much lower after a hurricane. Flowering (y) was related to shoot production per branch (x) ($\log(y) = 1 + 1.33 \cdot \log(x)$; $r^2 = 0.78$), whereas the fruit set rate was highly variable between trees (26 – 100%). The number of fruit per branch (1 – 6) was possibly related to poor shoot growth or fruit set, but unrelated to the timing of the phenological events. It is expected that commercial orchards would be more productive than feral stands.

The strawberry guava, *Psidium cattleianum* Sabine, is an evergreen shrub of the Myrtaceae family from South East Brazil and North Uruguay which has been naturalized throughout the tropics and subtropics (Popenoe, 1974). It was introduced to Réunion Island, a French overseas department located in the south-west Indian Ocean (lat. 21°06' S, long. 55°32' E) in the early 1800s, and rapidly colonised disturbed areas and forests up to 1,300 m, and now covers about 12,000 ha (Raunet, 1991).

Strawberry guava produces a round berry from 2 to 4 cm in diameter with a dark red (*P. cattleianum* f. *cattleianum*) or sulphur yellow (*P. cattleianum* f. *lucidum*) thin skin and a white soft and juicy flesh containing several hard seeds. Fruit are eaten fresh and are sweet and aromatic. Fruit with dark red skin give after crushing an aromatic and pink coloured pulp which can be processed into jam, jelly, juice and sorbet.

Little information is available on the phenology of this species, although flower buds are borne on new shoots as in common guava, *P. guajava*. In Hawaii, peak vegetative growth occurred from June to August and peak flowering and fruiting from June to October, and was independent of elevation from 150 to 762 m (Huenneke and Vitousek, 1990). Two growth flushes bearing flowers occurred in northern and southern India (Teaotia *et al.*, 1970; Subramanyan *et al.*, 1992). In Réunion Island, flowering occurred in November and December, with fruit harvested from March at sea level to August at elevation.

We report on the pattern of vegetative growth, flowering and fruiting of strawberry guava in Réunion. Data was collected from natural stands of trees growing at three elevations over two years. This information could be used to determine the production cycles within Réunion and the potential for commercial cultivation.

MATERIALS AND METHODS

Phenology

Phenology was monitored from April 1992 to July 1994 in feral stands located at Grand Brûlé (100 m), Piton Armand (480 m) and Bras Pistolet (720 m) on the windward wet side of Réunion (Table I). On each site, five adult trees in the full sun were chosen ranging from 1.8 to 3.0 m high. Trees were taller and denser at 720 m than at 100 and 480 m. In the first year, phenology was measured every 10 day on five arbitrarily selected and tagged terminal branches per tree, and on their offspring for the secondary growth. In the second year, phenology was measured on the shoots that appeared in the first year.

TABLE I
Description of the three experimental sites in Réunion Island

Site	Elevation (m)	Plant community	Soil ^z	Actual mean annual rainfall (mm)	Mean air temperature (°C)			
					Coldest months Aug' 92	Jul' 93	Warmest months Jan 93'	Feb' 94
Grand Brûlé	100	forest	recent basaltic lava flows	4,000 – 5,000	20.7	19.8	27.3	25.9
Piton Armand	480	fallow	perhydrated soil of the ando group	6,500 – 7,500	16.3	16.5	23.0	22.9
Bras Pistolet	720	fallow	perhydrated soil of the ando group	6,500 – 7,500	14.8	14.8	21.5	21.5

^z Raunet (1991).

Strawberry guava leaves are opposite and the nodes of a branch were numbered in a growing order from the subapex (node 1) to the base (Figure 1). The apical bud, at the extremity of the branch, was noted AB, and all leaf axillary buds were qualified as reproductive or vegetative. Three phenological stages were noted: 1. shoot emergence, composed of bud break and beginning of the new shoot elongation; 2. flowering: full opening of the flower, which lasts only 1 day; and 3. natural drop of the fruit, indicating full maturity. The frequency of a phenological stage on a site was calculated as the occurrence, summed over the five trees at this site, of this stage at one date divided by the total occurrence of this stage for the whole season on the site.

The number of shoots, flowers and fruit per terminal branch and the number of nodes per shoot were averaged each year at the tree level. To avoid possible effects of branch age or phenological stage, only terminal branches of the same age up to the peak vegetative growth were considered, excluding the few terminal branches with secondary growth activity and their offspring. Fruit set rate was calculated from the number of fruitlets three weeks after anthesis and the number of flowers.

Heat units calculation

Mean daily air temperatures of Grand Brûlé were directly determined. The daily thermic gradient g between the weather at Saint Benoit (40 m) and Plaine des Palmistes (1,025 m) was used to estimate mean daily air temperatures at 480 and 720 m:

$$[I] \quad \theta_x = \theta_{StBenoit} + g \cdot (x - 40)$$

where θ_x and x are respectively the mean daily air temperature (°C) and the elevation (m),

$$\text{and } g = \frac{\theta_{Plaine des Palmistes} - \theta_{StBenoit}}{1,025 - 40}.$$

These estimations were possible because of the regularity of the slope between Saint Benoit and Plaine des Palmistes, the same exposure of the two sites and their close neighbourhood with this transect.

Heat units were used to explore the relationships between the temperature and the development phase from shoot emergence (flower buds apparition) to flowering (SF), and from flowering to fruit maturity (FM). Daily heat units are defined as:

$$[II] \quad dti = \left(\frac{T_{\min i} + T_{\max i}}{2} \right) - T_B = T_{mi} - T_B$$

where dti is heat units (degree days) of day i , $T_{\min i}$, $T_{\max i}$ and T_{mi} are respectively the minimum, maximum and mean air temperature (°C) of day i and T_B is the base temperature (°C) of the development phase.

Heat units sum TT for a development phase where the daily mean temperature is always greater than the base temperature is defined as:

$$[III] \quad TT = \sum_{i=1}^{i=n} dti = \sum_{i=1}^{i=n} (T_{mi} - T_B) = n \cdot (T_m - T_B)$$

where dti is heat units (degree days) of day i , T_{mi} is the mean air temperature (°C) of day i , T_m , n and T_B are respectively the mean air temperature (°C), the duration in days and the base temperature (°C) of the development phase.

Heat units sums for flower bud development phase (SF) were calculated on each flower bud from shoot emergence to flowering. Heat units sums for fruit development phase (FM) were calculated on each fruit from flowering to fruit maturity.

Data analysis

Base temperatures for SF and FM were calculated on the complete data set (year 1 and 2) according to four classical methods described by Arnold (1959): identification of the base temperature 1) which minimised the standard deviation in days corresponding to the heat units sums of the data set, 2) which minimised the coefficient of variation of heat units sums of the data set, 3) which set to zero the regression coefficient between heat units sums and mean temperature during the phase for the data set, and 4) by calculating the x-intercept of the linear relationship between the development rate ($1 / \text{phase duration in days}$) and the mean temperature of the phase for the data set. For the first three methods, a program of minimisation of nonlinear function was used. For the fourth method, the relationship between development rate and mean temperature was analysed by regression analysis. Base temperatures were also directly calculated using mathematical formulae derived from the three first preceding methods (Yang *et al.*, 1995). The standard error and the 95% confidence interval of the base temperature calculated by minimising the coefficient of variation of heat units sums were estimated by bootstrap. Jackknife after bootstrap was used to detect observations with large influence on the result.

The numbers of shoots, flowers and fruit per branch, and the number of nodes per shoot were analysed by two-way analysis of variance (3 sites x 2 years) with $n = 5$ trees as replicates. Data set was unbalanced with $n = 4$ for the second year at Piton Armand (one tree was lost during a hurricane) and Bras Pistolet (on one tree, shoots of the second year emerged on branches from the first year secondary growth). Counts were transformed by log, and rates angular transformed prior to analysis. Because of unbalanced data set, differences between treatments were assessed using Tukey's multiple comparison tests on adjusted means (Dagnelie, 1998). Non-transformed adjusted means are presented in the results section. The relationship between the number of flowers and the number of shoots per branch was analysed by regression analysis after transformation by log of these two variables. Statistical analysis were performed using S-Plus 2000 statistical package (Mathsoft, 1999 a, b).

RESULTS AND DISCUSSION

Shoot structure

The number of nodes on a shoot (2 – 7) was determined at the beginning of its elongation after bud burst. Leaves, vegetative buds and flower buds were visible at this time. Across site and year, about 90% of the shoots had 4 to 6 nodes. Flower buds were borne singly in the leaf axils at the base of the shoots.

Vegetative buds were located on the upper nodes if flower buds were absent. Apical bud died on about 30% of the shoots. New shoots mainly originated from the top two buds. No shoots emerged on older wood. These results show a polarity of strawberry guava shoot with basal nodes flowering and apical nodes mainly vegetative (Figure1).

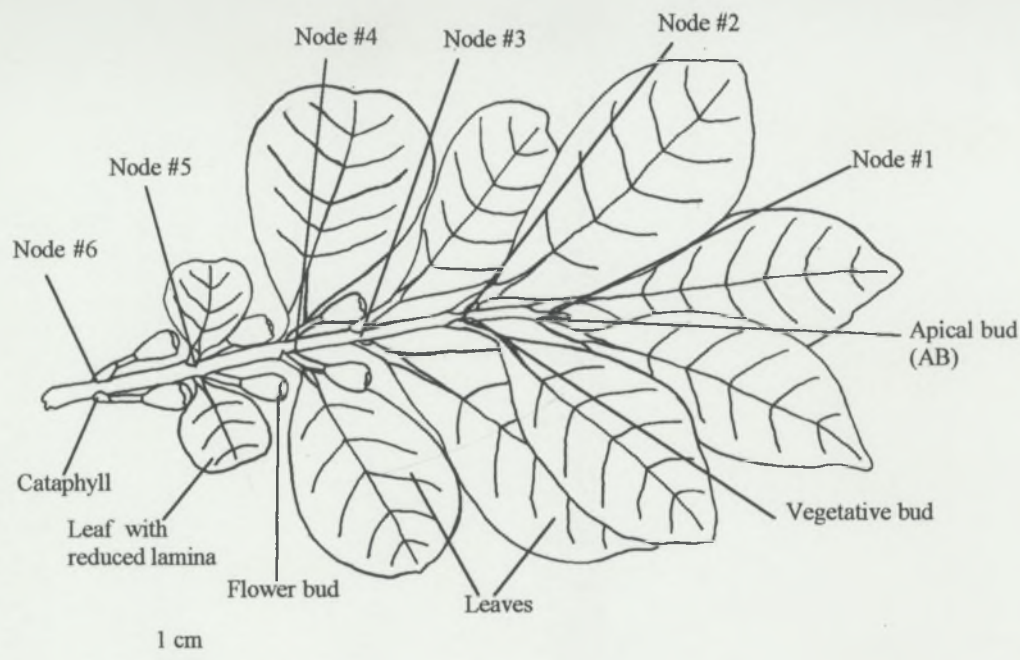


FIG. 1

Flowering shoot of strawberry guava, indicating the numbering used for nodes (see Materials and methods).

Vegetative and reproductive growth

There was a peak vegetative growth at each site at the end of the dry/cool season, between September and November. It was earlier at Grand Brûlé and Bras Pistolet than at Piton Armand in 1992/93 (Figure 2), but more synchronised in 1993/94 (Figure 3). Flowering reflected shoot emergence, with earlier flowering at low elevation. The main harvest was from February (Grand Brûlé) to June (Piton Armand and Bras Pistolet) in 1993 and from March (Grand Brûlé) to June (Bras Pistolet) in 1994. At each site, the harvest lasted 8-10 weeks. Secondary vegetative growth sometimes occurred later in the season, but contributed to few flower or fruit.

The pattern of shoot growth, flowering and fruiting described here is consistent with previous results on feral strawberry guava stands dominated by the forma *lucidum* (Huennecke and Vitousek, 1990), indicating that the two botanical groups have similar phenology. However, Teatitia *et al.* (1970) and Subramanyan *et al.* (1992) reported two distinct shoot growths on cultivated *P. cattleianum* f. *lucidum* in India. It is not known if these differences are due to environment or genotype. Differences in phenology also appear in cultivated guavas, with 1, 2 or 3 flowerings per year (Menzel and Paxton, 1986; Subramanyam *et al.*, 1992).

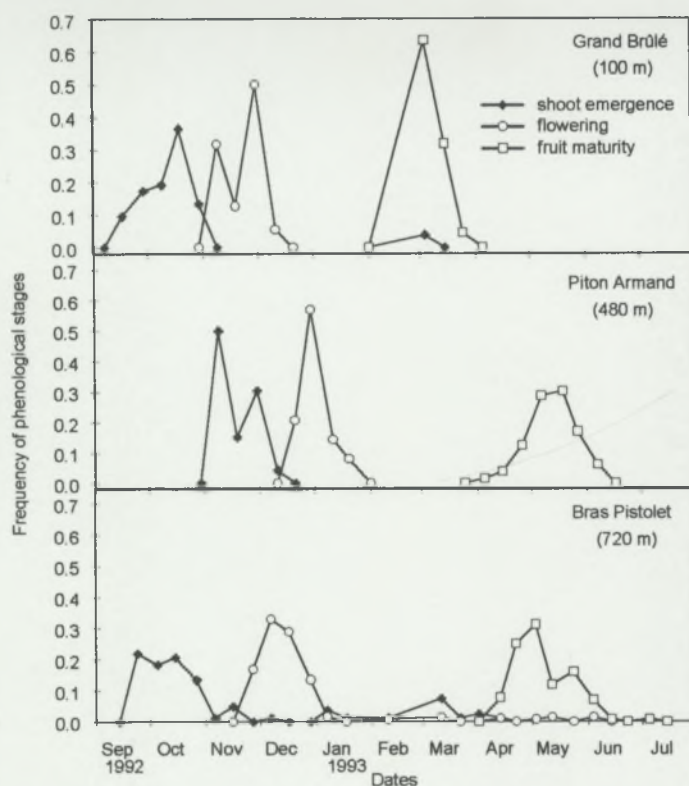


FIG. 2
Phenology of strawberry guava during 1992/93 at Grand Brûlé, Piton Armand and Bras Pistolet. The frequency of phenological stages is the occurrence of a phenological stage, summed over all the trees at a site at one date divided by the total occurrence of this phenological stage for the whole season.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

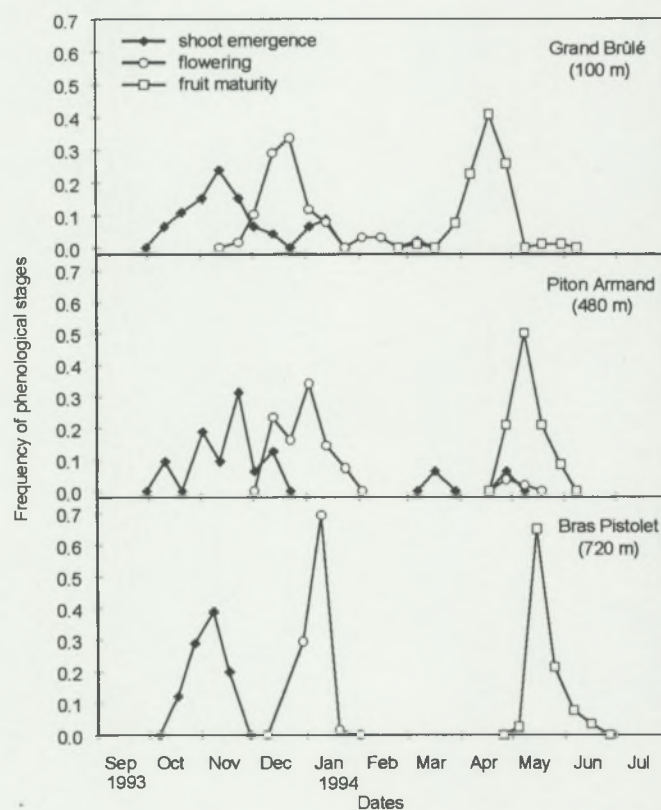


FIG. 3
Phenology of strawberry guava during 1993/94 at Grand Brûlé, Piton Armand and Bras Pistolet. The frequency of phenological stages is the occurrence of a phenological stage, summed over all the trees at a site at one date divided by the total occurrence of this phenological stage for the whole season.

The intervals between shoot emergence and flowering (SF), and flowering and fruit maturity (FM) increased with the elevation (Table II), suggesting a strong effect of temperature on plant development. Mean temperatures during SF varied from 17.5° to 25.5°C and during FM from 19.6° to 26.6°C. The four methods used for analysis gave similar base temperatures: 5.6°, 5.5°, 5.4°, and 5.3°C respectively for SF and 2.8°, 2.7°, 2.5° and 2.4°C respectively for FM. Identical values were given by mathematical formulae for the first three methods of calculation, indicating the absence of inconsistent data (Yang *et al.*, 1995), which was corroborated by the jackknife after bootstrap results. Means of 5.5° and 2.6°C were selected for SF and FM respectively. Bootstrap estimates of mean, standard error and 95% confidence interval for base temperature calculated by minimising the coefficient of variation of heat units were 5.5, 0.32, [4.8; 6.1] for SF and 2.7, 0.56, [1.6; 3.8] for FM. Similar precision were observed with the other methods of calculation, indicating that the base temperatures are well estimated with our data set. Mean heat unit sums calculated with these base temperatures were similar across years and sites and unrelated to elevation, validating the model (Table III). Mean heat units sums for SF and FM were 675° and 2,552°C.d, respectively.

TABLE II

Days between shoot emergence and flowering, and between flowering and fruit maturity of strawberry guava at three sites over two seasons in Réunion. Data are means \pm s.d. of individual flower buds and fruit. Sample sizes are given in parenthesis

Site	Elevation (m)	Shoot emergence to flowering		Flowering to fruit maturity	
		1992/93	1993/94	1992/93	1993/94
Grand Brûlé	100	39 \pm 4 (124)	36 \pm 3 (119)	105 \pm 7 (59)	116 \pm 5 (71)
Piton Armand	480	42 \pm 3 (156)	45 \pm 4 (46)	137 \pm 7 (39)	129 \pm 8 (21)
Bras Pistolet	720	52 \pm 4 (285)	55 \pm 4 (260)	147 \pm 12 (142)	136 \pm 8 (115)

TABLE III

Degree.days between shoot emergence and flowering, and between flowering and fruit maturity of strawberry guava at three sites over two seasons in Réunion. Data are means \pm s.d. of individual flower buds and fruit. Sample sizes are given in parenthesis

Site	Elevation (m)	Shoot emergence to flowering		Flowering to fruit maturity	
		1992/93	1993/94	1992/93	1993/94
Grand Brûlé	100	725 \pm 58 (124)	640 \pm 52 (119)	2,493 \pm 162 (59)	2,628 \pm 119 (71)
Piton Armand	480	635 \pm 46 (156)	645 \pm 45 (46)	2,636 \pm 125 (39)	2,535 \pm 150 (21)
Bras Pistolet	720	661 \pm 55 (285)	711 \pm 52 (260)	2,597 \pm 199 (142)	2,455 \pm 126 (115)

The strong effect of temperature on shoot and fruit development is consistent with data obtained from temperate and subtropical species (Munoz *et al.*, 1986; Fleisch and Bartholomew, 1987; Bittenbender and Kobayashi, 1990). But the base temperatures appear low for a subtropical species. Chou *et al.* (1973) observed that cambial activity in common guava was accelerated by temperature above 15°C. This physiologically related temperature threshold was considered as the lower temperature threshold for guava vegetative growth (Menzel and Paxton, 1986), or was used as base temperature to calculate heat units sum from

cycling (pruning and fertilising) to harvest in Hawaii (Bittenbender and Kobayashi, 1990). However, in a linear heat units system, the calculated base temperature has a statistical and not a physiological value (Arnold, 1959; Yang *et al.*, 1995; Bonhomme, 2000), is dependent of the calculation method and is generally lower than the minimum temperature for development (Yang *et al.*, 1995; Bonhomme, 2000). Strawberry guava is a subtropical cold resistant species which withstands frost without severe injury (Popenoe, 1974). The minimum temperatures for development phases in this species are thus expected to be lower than in the tropical common guava, and low calculated base temperatures are possible. However, further validation of these results are required.

Shoot growth, flowering and fruiting

The numbers of shoots, flowers and fruit per terminal branch were significantly lower in 1993/94 (Table IV). Shoot production was not affected by site, whereas flower and fruit production were higher at 720 m than at 480 m, and intermediate at 100 m. The number of nodes per shoot was higher at 720 m than at 100 and 480 m, but not affected by year. Fruit set was similar across years and sites. The significant interaction site x year was due to the constant number of fruit per branch for both year and the higher fruit set in 1993/94 at 100 m whereas these parameters were lower in 1993/94 at 480 and 720 m. There was a strong relationship between the number of shoots (x) and the number of flowers (y) per branch ($\log(y) = 1 + 1.33 \cdot \log(x)$; $r^2 = 0.78$; $n = 27$; $p < 0.001$).

Table IV

Numbers of shoots, flowers and fruit per branch, number of nodes per shoot and fruit set of feral strawberry guava on three sites over two years in Réunion. Data are adjusted non-transformed means of five plants per site x year combination, except for Piton Armand and Bras Pistolet in 1993/94 (n = 4). For each factor, means in column with different superscript letters are significantly different (P<0.05)

Factor	Treatment	No. shoots	No. nodes	No. flowers	No. fruit	Fruit set
Site	Grand Brûlé	1.3	4.5 ^b	4.0 ^{ab}	2.5 ^{ab}	70
	Piton Armand	1.5	4.3 ^b	3.9 ^b	1.9 ^b	60
	Bras Pistolet	1.9	5.1 ^a	7.6 ^a	4.5 ^a	57
	signif. ^z	n.s.	**	**	**	n.s.
Year	1992/93	2.0 ^b	4.6	7.6 ^b	4.4 ^b	64
	1993/94	1.1 ^a	4.6	2.7 ^a	1.6 ^a	61
	signif. ^z	***	n.s.	***	***	n.s.
Interaction	signif. ^z	n.s.	n.s.	n.s.	*	*

^z F test signification for the factor: n.s., *, **, *** nonsignificant, or significant at $P < 0.05$, 0.01 or 0.001 respectively.

The fewer shoots, flowers and fruit in 1993/94 was probably related to hurricane Colina on January 19, 1993. High winds bent trees and induced severe leaf drop at Grand Brûlé and Piton Armand. There was only partial defoliation at Bras Pistolet. Poor shoots and flowers production during the second year could be the consequence of the loss of leaves (Goldschmidt and Golomb, 1982; Buwalda and Smith, 1990), and of the probable damages to the root system a few months before.

The rate of plant development was similar at the three sites when temperature was taken into account. Flowering and fruiting were also heavier at Bras Pistolet than at Grand Brûlé and Piton Armand in year 1. Differences in year 2 were smaller after a hurricane. The trees were taller and denser at high elevation. Differences in growth and production amongst the sites could be due to soil, weather (especially temperature on dry matter production) or genotype. It was not possible to separate these factors in the present study.

Average yield was below 4 fruit per branch, except for trees at high elevation in year 1. Considering the strong relationship between flowering and shoot production, poor yields seem to be related to poor shoot growth. Cultivated yellow strawberry guava produced about 4.7 shoots per branch in India (Subramanyam *et al.*, 1992), generally higher than in our study. Fruit set varied with tree (26 – 100%), site and year, with low rates probably also a limiting factor in low yield. Reproductive biology is poorly documented in this species (Normand, 1994) and the factors affecting the fruit set are not well described. Fruit retention was high, with only a few fruit dropping, caused by external factors (pest, shock), once fruit had set. As found with seven guava cultivars in Australia (Menzel and Paxton, 1986), differences in yield could not be related to differences in phenology.

Prospects for commercial production in Réunion.

We provide the first report on strawberry guava phenology and production in feral stands in Réunion. Climatic conditions in the humid areas of the island are generally suited for growth and production for this species. However, information is required on the cause of the poor shoot growth and flowering.

These results are hopeful for the development of commercial orchards in Réunion as better tree management and plant growth are expected to provide higher yields. The production season could be extended with orchards located at different elevations. The quality of the fruit also needs to be evaluated under commercial cultivation.

REFERENCES

- ARNOLD, C. Y. (1959). The determination and significance of the base temperature in a linear heat unit system. *Proceedings of the American Society for Horticultural Science*, **74**, 430–45.
- BITTENBENDER, H.C. and KOBAYASHI K.D. (1990). Predicting the harvest of cycled 'Beaumont' guava. *Acta Horticulturae*, **269**, 197–204.
- BONHOMME, R. (2000). Bases and limits to using 'degree.day' units. *European Journal of Agronomy*, **13**, 1–10.

- BUWALDA, J. G. and SMITH, G. S. (1990). Effects of partial defoliation at various stages of the growing season on fruit yields, root growth and return bloom of kiwifruit vines. *Scientia Horticulturae*, **42**, 29–44.
- CHOU, T., HWA, S. AND CHIANG, T. (1973). Seasonal changes in cambial activity in the young branch of *Psidium guajava* Linn. *Taiwania*, **18**, 35–41.
- DAGNELIE, P. (1998). *Statistique théorique et appliquée. tome 2, inférence statistique à une et à deux dimensions*. De Boeck & Larcier S.A., Brussels, Belgium.
- FLEISCH, H. and BARTHOLOMEW, D.P. (1987). Development of a heat unit model of pineapple ('Smooth Cayenne') fruit growth from field data. *Fruits*, **42**, 709–15.
- GOLDSCHMIDT, E. E. and GOLOMB, A. (1982). The carbohydrate balance of alternate-bearing citrus trees and the significance of reserves for flowering and fruiting. *Journal of the American Society for Horticultural Science*, **107**, 206–8.
- HUENNEKE, L. F. and VITOUSEK, P. M. (1990). Seedling and clonal recruitment of the invasive tree *Psidium cattleianum*: Implications for management of native Hawaiian forests. *Biological Conservation*, **53**, 199–211.
- MATHSOFT, 1999a. *S-Plus® 2000 Guide to Statistics, volume I*. Data Analysis Products Division, Mathsoft, Seattle, WA.
- MATHSOFT, 1999b. *S-Plus® 2000 Guide to Statistics, volume II*. Data Analysis Products Division, Mathsoft, Seattle, WA.
- MENZEL, C.M. and PAXTON, B.F. (1986). The pattern of growth, flowering and fruiting of guava varieties in subtropical Queensland. *Australian Journal of Experimental Agriculture*, **26**, 123–8.
- MUNOZ, C., SEPULVEDA, G., GARCIA-HUIDOBRO, J. and SHERMAN, W. B. (1986). Determining thermal time and base temperature required for fruit development in low-chilling peaches. *HortScience*, **21**, 520–2.
- NORMAND, F. (1994). Le goyavier-fraise, son intérêt pour l'île de la Réunion. *Fruits*, **49**, 217–27.
- POPENOE, W. (1974). Fruits of the Myrtle family. In: *Manual of tropical and subtropical fruits, excluding the banana, coconut, pineapple, citrus fruits, olive and fig*. (originally published in 1920 by Macmillan Publishing Co, New York) Hafner Press, New York., 272–311.
- RAUNET, M. (1991). *Le milieu physique et les sols de l'île de la Réunion. Conséquences pour la mise en valeur agricole*. (CIRAD, Ed.), Montpellier, France.
- SUBRAMANYAM, M. D., DINESH, M. R. and BRAGANZA, M. (1992). Varietal evaluation and floral biology studies in the genus *Psidium*. *Acta Horticulturae*, **321**, 211–9.
- TEAOTIA, S. S., PHOGAT, K. P. S. and SRIVASTAVA, V. S. (1970). Blossom biology studies in *Psidium* species. *Progressive Horticulture*, **2**, 101–12.
- YANG, S., LOGAN, J. and COFFEY, D.L. (1995). Mathematical formulae for calculating the base temperature for growing degree days. *Agricultural and Forest Meteorology*, **74**, 61–74.

II- L'azote déclenche l'émission de pousses florifères chez le goyavier-fraise

Nitrogen fertilisation induces floriferous flush in strawberry guava (*Psidium cattleianum*)

Frédéric Normand^{a*}, Robert Habib^b

^a CIRAD-FLHOR, BP 180, 97455 Saint Pierre cedex, Ile de la Réunion, France

^b INRA, Plantes et Systèmes de culture Horticoles, Domaine St Paul, Site Agroparc, F-84914 Avignon cedex 9, France

Agronomie, 21(8), (2001), 735–742

Abstract

Empirical observations seemed to indicate that NPK fertilisation induced a rapid floriferous flush in strawberry guava (*Psidium cattleianum*), thereby modifying its natural phenological cycle. Factors affecting the plant response to fertilisation were investigated, including: identification of the nutrient(s) responsible of the flush; effect of the amount of fertiliser and of the plant phenological stage. Complete fertilisation was followed by the emergence of a floriferous flush 30 to 50 days after fertilisation. Nitrogen was the trigger of this flush, enhancing the proportion of flushing branches. The amount of fertiliser or the leaf nitrogen content had no effect on the flush characteristics. The phenological stage of the plant had a strong influence on the response to fertiliser, with a more intense and less variable flush when fertilisation was applied after a 3 months resting period than just after a harvest. This suggested that the carbohydrate availability within the plant was involved in the response.

Psidium cattleianum / nitrogen / phenology / crop cycling / Reunion Island

Résumé

Des observations empiriques semblent indiquer qu'une fertilisation NPK induit rapidement une pousse florifère chez le goyavier-fraise (*Psidium cattleianum*), modifiant ainsi son cycle phénologique naturel. Les facteurs affectant la réponse de la plante à la fertilisation ont été étudiés: identification du (des) élément(s) minéraux responsable(s) de la pousse, effet de la quantité d'engrais et du stade phénologique de la plante. Une fertilisation complète est suivie par l'émergence d'une pousse florifère 30 à 50 jours après l'apport. L'azote est le déclencheur de cette pousse; il augmente le pourcentage de branches émettant des pousses. La quantité d'engrais ou la teneur en azote dans les feuilles n'a pas d'effet sur les caractéristiques de la pousse. Le stade phénologique de la plante a par contre une grande influence sur la réponse à la fertilisation, avec une pousse plus intense et moins variable quand la fertilisation est

apportée après une période de repos de 3 mois plutôt que juste après une récolte. Cela suggère que la disponibilité en carbone dans la plante est impliquée dans sa réponse à l'azote.

***Psidium cattleianum* / azote / phénologie / cycle de culture / Ile de la Réunion**

1. INTRODUCTION

The strawberry guava, *Psidium cattleianum* Sabine, is an evergreen shrub of the Myrtaceae family native from South America. It was introduced at the beginning of the 19th century in the Réunion Island (Indian Ocean; lat. 21°06'S, long. 55°32'E) where it rapidly colonised the humid disturbed areas and forests, covering up about 12 000 ha [19]. It has not been of agricultural importance until now.

It is a potential new fruit species for fresh market and processing industry in Réunion Island. Very little information about cropping techniques being available from the literature, a domestication program has been started [16] with the objective to improve yield and fruit quality by the way of classical agronomic factors (tree spacing, fertilisation, pruning). There is little variability among plants [18] so that there was no need to start with a selection program.

Growing common guava, *P. guajava* L., with several production cycles each year is practised in several producing countries. The objectives are to produce a regular crop throughout the year, or to shift the harvest period to a more favourable cropping season. Production cycles are triggered after harvest by pruning and/or defoliation, combined with fertilisation or rewatering after drought [22, 9, 23, 14]. However, factors affecting yield components of triggered cycles are poorly documented.

Strawberry guava phenology in natural conditions in Réunion Island consists of a single flush of floriferous shoots in October (spring), followed by flowering in December and fruit maturity from March to July depending on the elevation. No vegetative activity occurs between fruit maturity and the next floriferous flush [17]. The same phenological pattern has been described in feral thickets in Hawaii [8]. But empirical observations on strawberry guava grown in orchards in Réunion Island suggest a positive effect of complete fertilisation (NPK) on production cycle triggering.

The objective of this study was to specify the factors affecting the strawberry guava floriferous flush consecutive to fertilisation by i) determining which nutrient(s) of the complete fertilisation is involved in the floriferous flush, ii) quantifying the effect of increasing amount of fertiliser on the shoot and flower emergence, and iii) evaluating the effect of the plant phenological stage at fertilisation on the plant reaction intensity. Two experiments were conducted in strawberry guava orchards located on the humid East coast of the island (average annual rainfall = 6 700 mm) with tree spacing of 3 m (inter row) x 1 m (on the row).

2. MATERIALS AND METHODS

2.1. Determination of the nutrient(s) responsible of the induced floriferous flush

The objective of this experiment was to determine the nutrient(s) responsible of the floriferous flush observed after a complete fertilisation. It was conducted on a five and a half year-old orchard located at 480 m. Plants were grown from seed and were thus all genetically different. The last harvest was from March to May 2000 and plants had no vegetative nor reproductive activity since May 2000. The last fertilisation they received was in March 1999. The following five treatments were applied on July 4th, 2000, two months after the end of the harvest. The amount of each individual nutrient was calculated to bring the same quantity of this nutrient than the complete fertilisation.

- control: no nutrient
- NPK treatment: 240 g 15.12.24 per plant, i.e. 36 g N (13.2 g N as NO_3^- and 22.8 g N as NH_4^+); 28.8 g P_2O_5 ; and 57.6 g K_2O per plant;
- N treatment: 138.5 g NH_4NO_3 , i.e. 36 g N, per plant;
- P treatment: 27.5 ml per plant of a H_3PO_4 85% solution (1050 g $\text{P}_2\text{O}_5\cdot\text{L}^{-1}$) diluted in 1 L pure water, i.e. 28.8 g P_2O_5 per plant;
- K treatment: 115.2 g K_2SO_4 , i.e. 57.6 g K_2O , per plant.

A complete randomised block design with 4 replicates was used. The plants studied were homogeneous in size and vigour, selected on their stem section at the soil level. Each treatment was applied to 3 consecutive plants on the row, but only the central plant reaction was measured. Fertilisers were brought evenly on the ground area between the stem and the plumb of the tree edge. Regular rainfalls (146 mm during the fortnight following the application) assured a good dissolution of the nutrient and maintained the soil wet. Treated plants were at a distance of at least 4 m so that there was no interaction between treatments. Fifteen terminal branches of 9 month-old, identified by the dry peduncle of the fruit from the last harvest, were randomly chosen and labelled on each studied plant. The numbers of shoots and flower buds produced by the labelled branches were recorded 2.5 months after the treatment application. The number of flowers produced was recorded just before flowering.

Leaf nitrogen content was determined just before fertilisation, and then at 10, 20, 30, 40 and 50 days after fertilisation (DAF). On each plant, the second pair of leaves from the apex was collected on five, 9 month-old arbitrarily chosen terminal branches. Samples from plants receiving the same treatment in blocks 1-2 and 3-4 were mixed. Leaf samples were collected in the morning, put in plastic bags in an ice-box and brought to the laboratory for analysis. Samples were weighted and oven dried at 75°C until weight stabilisation. Nitrogen was analysed through segmented flow analysis according to the modified Berthelot reaction after a Kjeldahl digestion without nitrate reduction [21]. Results were expressed on a leaf dry weight basis.

2.2. Amount of fertiliser and phenological stage effect on the induced floriferous flush

The objective of this experiment was to study the effect of the fertilisation date and of the amount of fertiliser on the shoots and flowers emergence. It was conducted on a four year-old orchard located at 520 m. The following treatments were applied:

- control: no fertilisation
- X treatment: 18 g N per plant (120 g 15.12.24)
- 2X treatment: 36 g N per plant (240 g 15.12.24)
- 3X treatment: 54 g N per plant (360 g 15.12.24)
- 4X treatment: 72 g N per plant (480 g 15.12.24)
- 5X treatment: 90 g N per plant (600 g 15.12.24)

Treatments were applied on different plants of homogeneous size at two dates corresponding to different plant phenological stages: on April 4th, 1997, at the end of a heavy crop; and on July 24th, 1997, three months after the end of the harvest. A complete randomised block design with 4 replicates was used for each date. Each treatment was applied to a single plant per block. Fertiliser was brought as in the previous experiment. The fortnights after fertilisation were rainy (337 and 388 mm after April and July fertilisations respectively), which assured a good dissolution of the nutrient and maintained the soil wet. Fertilised plants were at a distance of at least 3 m.

The numbers of shoots and flower buds were recorded as in the first experiment. For the April fertilisation, the number of flowers was not recorded, and leaf nitrogen content was measured on the plants of one block, every 10 day from fertilisation to 50 DAF, following the leaf sampling and analysis procedures described above.

2.3. Data analysis

Two complementary approaches were used to quantify the floriferous flush following fertilisation. The first one consisted of calculating the number of flowers produced from basic parameters separating the vegetative and the reproductive intensities of the flush:

$$Nbf = NbB \cdot \%FB \cdot sh/FB \cdot \%fsh \cdot fb/fsh \cdot (1 - \%fbD)$$

Where: Nbf: number of flower produced;

NbB: number of labelled terminal branches;

%FB: proportion of flushing terminal branches, i.e. the ratio of the number of terminal branches for which at least one bud gives a new shoot on the number of labelled terminal branches, NbB;

sh/FB: number of new shoots per flushing terminal branch;

%fsh: proportion of floriferous shoots, i.e. the ratio of the number of new shoots bearing at least one flower bud on the total number of new shoots;

fb/fsh: number of flower buds per floriferous shoot;

%fbD: proportion of dropped flower buds;

These parameters were calculated from the data collected in the experiments. The treatments effects were analysed by analysis of variance (ANOVA) on each parameter. Missing data were numerous when flushing was scarce (sh/FB, %fsh and fb/fsh did not exist when %FB is null) or not too floriferous. Then ANOVAs were performed on unbalanced data sets.

The second approach was the analysis of the treatments effects on the total number of shoots (totsh), on the total number of floriferous shoots (totfsh) and on the total number of flower buds (totfb) produced, calculated over the 15 labelled branches per plant. It did not

distinguish as clearly as previously vegetative and reproductive intensities of flushing, but provided balanced data sets for ANOVA.

One way ANOVA was performed with the S-Plus 2000 software [11] for testing treatments effects on the variables in each experiment. In the second experiment, ANOVAs were realised with and without the control to reveal differences between amounts of fertiliser in the latter case. Proportions were angular transformed before analysis to stabilise the variance, but untransformed means are presented in the results section. Differences between treatments were assessed using Tukey's multiple comparison test.

3. RESULTS

3.1. Determination of the nutrient(s) responsible of the induced floriferous flush

Plants which received NPK or N treatment rapidly emitted a floriferous flush after nutrient application: buds swelled at 30 DAF and burst from 50 DAF. Controls and plants which received P or K treatment had no vegetative nor reproductive activity after nutrient application. Consequently, they were not considered for statistical analysis. Plant reaction following the N treatment were lower than for the NPK treatment, particularly for %FB, fb/fsh and the number of shoots, floriferous shoots and flower buds produced (*table 1*). Differences were however not significant.

The leaf nitrogen content in plants which received nitrogen increased from fertilisation and reached a maximum at 20 DAF whereas that in plants which did not received nitrogen slightly decreased. Then it steeply decreased for all treatments between 20 and 30 DAF and increased from 30 to 50 DAF (*figure 1*). The nitrogen application (NPK and N treatments) induced a significantly higher leaf nitrogen content from 20 to 50 DAF.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

Table I

Characteristics (mean \pm s.d.) of the strawberry guava flush consecutive to the application of 36 g N per plant (N treatment) or 240 g 15-12-24, i.e. 36 g N, per plant (NPK treatment) two months after a resting period following harvest.

Treat.	%FB	sh/FB	%fsh	fb/fsh	%fbD	totsh	totfsh	totfb
NPK	0.90 \pm 0.08	2.14 \pm 0.49	0.96 \pm 0.03	4.36 \pm 1.27	0.22 \pm 0.13	29.2 \pm 9.2	28.2 \pm 9.2	130.5 \pm 81.1
N	0.67 \pm 0.24	2.05 \pm 0.39	0.91 \pm 0.06	3.74 \pm 0.86	0.18 \pm 0.07	19.9 \pm 5.4	18.1 \pm 4.8	70.0 \pm 31.1
Signif. ¹	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

¹: n.s., non significant

%FB: proportion of flushing terminal branches; sh/FB: number of new shoots per flushing terminal branch; %fsh: proportion of floriferous shoots; fb/fsh: number of flower buds per floriferous shoot; %fbD: proportion of dropped flower buds; totsh: total number of shoots produced, totfsh: total number of floriferous shoots produced; totfb: total number of flower buds produced.

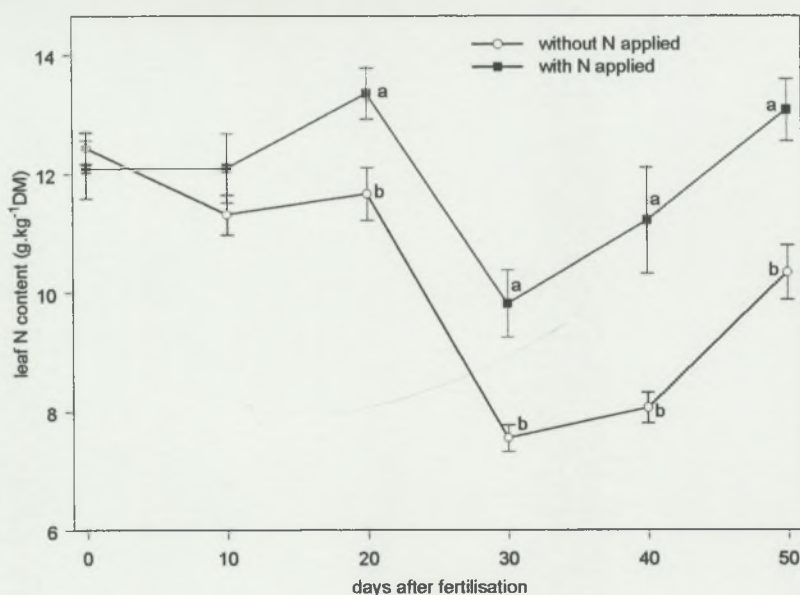


Figure 1: The effect of fertilisation on the leaf nitrogen content of strawberry guava. Plants which received nitrogen (NPK –36 g N; 28.8 g P₂O₅; 57.6 g K₂O per plant– and N – 36 g N per plant– treatments) are separated from plants which received no nitrogen (non fertilised control, P – 28.8 g P₂O₅ per plant– , and K – 57.6 g K₂O per plant– treatments). Vertical bars indicate standard error of each mean. For each date, means with different letters are significantly different (Tukey's test at $P=0.05$).

3.2. Amount of fertiliser and phenological stage effect on the induced floriferous flush

The April 1997 fertilisation, just after harvest, induced a rapid plant growth. Non-fertilised controls also showed a slight growth. Buds swelled after 20 DAF and burst from 30 to 50 DAF. The flush was weak and not very floriferous irrespective of the amount of fertiliser (*table II*). For a given treatment, the variability in plant response was high, particularly for variables related to the floriferous status of the shoots. In spite of marked differences between means, differences between treatments were not significant. The proportion of flushing branches, %FB, was low on control and increased with the amount of fertiliser. Highly significant differences occurred between control and 4X and 5X treatments. Differences in the total number of shoots produced, totsh, were related to the differences in %FB as sh/FB was not affected by treatments. The other parameters were not significantly affected by the treatments. No floriferous shoot was observed on the control and on the X treatment. ANOVA without the control gave non significant differences between amounts of fertiliser for all parameters. There was a tendency of increasing %FB means with the amount of fertiliser.

Table II

Characteristics (mean \pm s.d.) of the strawberry guava flush consecutive to the application of different amounts of NPK fertiliser (X to 5X treatments) or without fertilisation (control) at the end of harvest in April 1997. ANOVA results are first presented including all the treatments, and then without the control to test the effect of the amount of fertiliser.

Treat.	N applied (g per plant)	%FB	sh/FB	%fsh	fb/fsh	totsh	totfsh	totfb
control	0	0.07 \pm 0.08 b	1.75 \pm 0.35	0	-	1.7 \pm 2.1 b	0	0
X	18	0.22 \pm 0.23 ab	1.89 \pm 0.19	0	-	6.2 \pm 6.8 ab	0	0
2X	36	0.58 \pm 0.24 ab	1.81 \pm 0.30	0.08 \pm 0.10	2.41 \pm 2.00	16.5 \pm 9.3 ab	1.7 \pm 2.9	6.0 \pm 11.3
3X	54	0.35 \pm 0.35 ab	1.64 \pm 0.50	0.17 \pm 0.33	3.05 ¹	9.5 \pm 11.7 ab	4.5 \pm 9.0	13.7 \pm 27.5
4X	72	0.73 \pm 0.24 a	2.09 \pm 0.18	0.18 \pm 0.22	3.78 \pm 1.07	23.2 \pm 8.8 a	4.7 \pm 4.7	16.0 \pm 16.5
5X	90	0.72 \pm 0.26 a	1.78 \pm 0.36	0.16 \pm 0.19	1.78 \pm 0.95	20.0 \pm 10.1 ab	3.0 \pm 2.6	6.2 \pm 7.6
mean	-	0.443	1.83	0.11	2.73	12.9	2.3	7.0
signif. ²	-	**	n.s.	n.s.	n.s.	*	n.s.	n.s.
Without the control:								
mean	-	0.519	1.84	0.24	2.73	15.1	2.80	8.40
signif. ²	-	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

¹ : n = 1 value, s.d. is not calculated

²: signif.: n.s.: non significant; *, **: significant respectively at $P=0.05$ and $P=0.01$. Means followed by different letters are significantly different (Tukey's multiple comparison test).

%FB: proportion of flushing terminal branches; sh/FB: number of new shoots per flushing terminal branch; %fsh: proportion of floriferous shoots; fb/fsh: number of flower buds per floriferous shoot; totsh: total number of shoots produced, totfsh: total number of floriferous shoots produced; totfb: total number of flower buds produced.

No significant correlation was found between the leaf nitrogen content at 0, 10, 20 or 30 DAF and the parameters of the plant response (data not presented). Correlations with leaf nitrogen content at 40 and 50 DAF were not tested as bud burst had occurred by this time, and the number of shoots and flower buds were already fixed.

There was no relation between the amount of nitrogen applied in April 1997 and the leaf nitrogen content during 50 DAF. Consequently, data of X to 5X treatments were pooled and compared with the leaf nitrogen content of the control (figure 2). Changes in the leaf nitrogen content of control and fertilised plants had the same pattern, but values were higher in fertilised plants from 10 DAF, indicating a rapid uptake of nitrogen. The maximum leaf nitrogen content was reached at 20 DAF, followed by a slow decrease until 50 DAF. Leaf nitrogen content of control and fertilised plants were higher than in the first experiment.

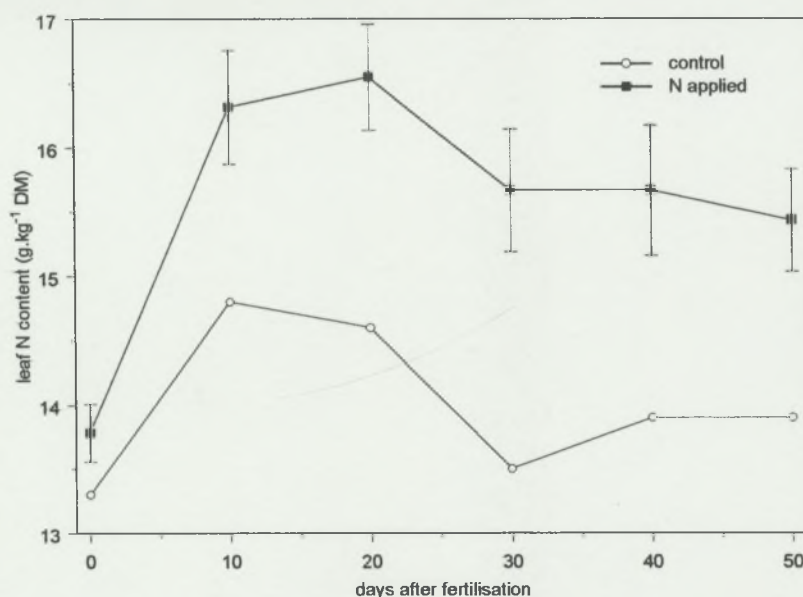


Figure 2: The effect of fertilisation on the leaf nitrogen content of strawberry guavas fertilised with different amounts of complete fertiliser in April 1997. Plants which received nitrogen (18, 36, 54, 72 and 90 g N per plant) were pooled and the vertical bars indicate the standard error of mean. The unique control plant was not fertilised.

The July 1997 fertilisation, after a 3 months resting period following harvest, induced a rapid plant growth. Buds swelled at 30 DAF and burst from 50 DAF. The flush was intense and floriferous (*table III*). A slight flushing activity was observed on the control. For a given treatment, the variability of plant response was lower than after the fertilisation of April 1997. The fertilised plants had all a high proportion of flushing branches, %FB, significantly higher than for the control. A high variability in %FB was observed in this latter. The number of new shoots per flushing branch (sh/FB) was lower for the control and increased with amounts of fertiliser. It was significantly different between the control and the 5X treatment. Almost all shoots were floriferous, whatever the presence or absence and the amount of fertiliser. The number of flower buds per floriferous shoot (fb/fsh) was lower for the control and increased by 1 to almost 2 units on fertilised plants, with significant differences between the extreme mean values (control and 4X treatment). The proportion of dropped flower buds (%fbD) was not affected by the presence or absence and the amount of fertiliser. Highly significant differences were found on totsh, totfsh and totfb between the control and the fertilised plants. They resulted from the differences in %FB as this parameter was then multiplied by parameters unaffected or slightly affected by treatments. The number of shoots, floriferous shoots and flower buds produced increased with the amount of fertiliser, but differences were not significant.

4. DISCUSSION

4.1. Nutrient(s) responsible of the floriferous flush induced by fertilisation

The results of both experiments clearly illustrates the rapid emergence of a floriferous flush after a complete fertilisation of strawberry guava. Fertilisation, therefore, modifies the natural phenological cycle of strawberry guava in Réunion Island by triggering a production cycle.

Nitrogen was the nutrient involved in the emergence of a floriferous flush following a complete fertilisation. In the first experiment, the plant reaction was observed only on plants which received nitrogen (NPK and N treatments). Phosphorus and potassium were not involved as nutrients. The number of shoots and flower buds formed with a pure nitrogen application or with a complete fertilisation were not significantly different. This suggests that no interaction occurs between nitrogen and other nutrients on the shoot emergence. Nitrogen stimulates the growth of strawberry guava with the proportion of flushing branches as the most affected parameter.

The amount of fertiliser, and thus of nitrogen, has no significant effect on the induced flush, whatever the strawberry guava phenological stage at the fertilisation date. In particular, there is no relation between the leaf nitrogen content from fertilisation to bud swell and the plant response parameters, suggesting that factors other than leaf nitrogen content are involved in the vegetative and reproductive intensity of the plant response.

4.2. Effect of the plant phenological stage at fertilisation

The plant phenological stage when fertilisation was applied had a strong effect on the plant response. When fertilisation was applied in April, just after harvest, the induced flush was weak and produced few flowers, and the variability in the response was high. When fertilisation was applied in July, after a 3 month resting period following harvest, the induced flush was intense and floriferous, and the variability in plant reaction was low. These differences could be related to the plant phenological stage when fertilisation was applied, or to climatic factors such as temperature, water availability or photoperiod, which could enhance shoot or flower initiation. Elevation, and consequently the temperature, has no clear effect on the strawberry guava growth [17]. Minimum monthly rainfall during the experiments was 222 mm and the soil was always wet. Moreover, strawberry guava is not a photoperiodic species for growth [2]. It appears therefore that the differences in the flush characteristics were related to the plant phenological stage when fertilisation was applied.

Table III

Characteristics (mean \pm s.d.) of the strawberry guava flush consecutive to the application of different amounts of NPK fertiliser (X to 5X treatments) or without fertilisation (control) in July 1997, after a 3 months resting period following harvest. ANOVA results are first presented including all the treatments, and then without the control to test the effect of the amount of fertiliser.

Treat.	N applied (g per plant)	%FB	sh/FB	%fsh	fb/fsh	%fbD	totsh	totfsh	totfb
control	0	0.33 \pm 0.30 b	2.07 \pm 0.12 b	0.98 \pm 0.03	5.02 \pm 1.32 b	0.04 \pm 0.05	10.2 \pm 9.1 b	10.0 \pm 8.7 b	47.5 \pm 36.6 b
X	18	0.97 \pm 0.07 a	2.38 \pm 0.18 ab	1.00 \pm 0.00	6.21 \pm 0.78 ab	0.10 \pm 0.07	34.5 \pm 2.4 a	34.5 \pm 2.4 a	214.7 \pm 33.7 a
2X	36	0.98 \pm 0.03 a	2.37 \pm 0.15 ab	1.00 \pm 0.00	6.39 \pm 0.47 ab	0.13 \pm 0.04	35.0 \pm 3.2 a	35.0 \pm 3.2 a	223.5 \pm 22.2 a
3X	54	0.98 \pm 0.03 a	2.64 \pm 0.49 ab	1.00 \pm 0.00	6.45 \pm 0.28 ab	0.12 \pm 0.04	39.0 \pm 8.2 a	39.0 \pm 8.2 a	253.2 \pm 57.1 a
4X	72	0.98 \pm 0.03 a	2.79 \pm 0.30 ab	1.00 \pm 0.00	6.92 \pm 0.62 a	0.08 \pm 0.01	41.2 \pm 5.4 a	41.2 \pm 5.4 a	286.7 \pm 53.3 a
5X	90	0.96 \pm 0.07 a	3.00 \pm 0.47 a	0.99 \pm 0.01	6.57 \pm 0.61 ab	0.14 \pm 0.13	43.2 \pm 6.1 a	43.0 \pm 6.4 a	282.7 \pm 52.8 a
mean	-	0.87	2.56	1.00	6.13	0.10	33.9	33.8	218.1
signif. ¹	-	***	*	n.s.	*	n.s.	***	***	***
Without the control:									
mean	-	0.98	2.64	1.00	6.51	0.11	38.6	38.5	252.2
signif. ¹	-	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

¹: signif.: n.s.: non significant; *, ***: significant respectively at $P=0.05$ and $P=0.001$. Means followed by different letters are significantly different (Tukey's multiple comparison test).

%FB: proportion of flushing terminal branches; sh/FB: number of new shoots per flushing terminal branch; %fsh: proportion of floriferous shoots; fb/fsh: number of flower buds per floriferous shoot; %fbD: proportion of dropped flower buds; totsh: total number of shoots produced, totfsh: total number of floriferous shoots produced; totfb: total number of flower buds produced.

Our results support the hypothesis that the characteristics of the flush triggered by nitrogen are in relation with the carbohydrate availability. On one hand, carbohydrate availability changes in fruit trees during the year with respect to the phenological stages [20, 10, 24]. On the other hand, relations between the carbohydrate availability and shoots or flower buds formation are documented for several fruit tree species [25, 7, 3]. Growing strawberry guava fruit are strong sinks for carbohydrates: a study of the dry matter balance made on 3 plants of the orchard of the second experiment during the 1997 fruiting period indicated that immature fruit bore about 25% of the total plant dry weight, 2 months before harvest (unpublished data). We can consequently suspect a low carbohydrate availability at harvest, as in avocado [20], tangerine [10] or mango [24]. Plant carbohydrate level is conditioned by the reserves, the leaf area and the fruit load. It is, therefore, related to the source-sink balance at the tree level, hence probably the strong variability in plant reaction when fertilisation was applied during or just after harvest. Carbohydrate levels are restored after harvest in different fruit species [20, 10, 24]. Higher carbohydrate levels in the plant probably led to the intense and floriferous flush consecutive to a fertilisation after a 3 month resting period. The variability was lower as the differences in plant carbohydrate level induced by the fruit load had vanished and all plants were supposed to have recovered a normal carbohydrate level. Under this hypothesis, shoots and flower buds formation appears sensible to carbohydrate availability and strawberry guava behaves as an alternate bearing species when floriferous flushes are triggered with nitrogen [13].

Although significant difference occurred only between the control and the 4X treatment, the number of flower buds per floriferous shoot was higher on the fertilised plants than on the non-fertilised plants when fertilisation was applied after a resting period (*table III*). This suggests that fertilisation stimulated the flower buds formation. This stimulation was rapid after the fertilisation as flower buds were formed and visible when the buds burst and the new shoots began to elongate 4 to 6 weeks after fertilisation, depending on the temperature. There is often a good correlation between leaf nitrogen content and leaf photosynthetic rate in fruit trees such as peach [5] or apple [4] and in herbaceous species [12, 6, 15]. Assuming such a relation in strawberry guava, the higher leaf nitrogen content after fertilisation could lead to a higher carbohydrate availability in fertilised plants. The higher number of flower buds in fertilised plants, particularly when high amounts of nitrogen were applied after a resting period, could be in relation with these changes in carbohydrate availability [7, 1].

4.3. Agronomic consequences

The possibility of triggering production cycles by nitrogen has interesting agronomic consequences as cycles could be superimposed on the plants. Two to three harvests are possible on the same plant in a year, which would increase the annual yield (unpublished data). Triggering dates could be chosen to harvest fruit with respect to the market demand or during the cooler season when fruit fly populations are low and fruit quality is better. Nitrogen fertilisation programmes must, therefore, be adapted to the growth cycle of strawberry guava. The cycles are easier to trigger than in the guava cultivation where a combination of pruning/defoliation and fertilisation is needed [22, 9, 23, 14]. However, the suspected dependence of the plant response to the carbohydrate availability is a limitation that should be taken into account. To our knowledge, this aspect has not been found for common guava.

REFERENCES

- [1] Bodson M., Bernier G., Is flowering controlled by the assimilate level?, *Physiol. Veg.* 23 (1985) 491–501.
- [2] Broschat T.K., Donselman H.M., Effect of photoperiod on growth of seven species of tropical and subtropical ornamental trees, *HortScience* 18 (1984) 571.
- [3] Buwalda J.G., Smith G.S., Effects of partial defoliation at various stages of the growing season on fruit yields, root growth and return bloom of kiwifruit vines, *Sci. Hortic.* 42 (1990) 29–44.
- [4] Cheng L., Guak S., Fuchigami L.H., The relationship between leaf nitrogen content and photosynthesis in apple leaves, *HortScience* 31 (1996) 578.
- [5] DeJong T.M., Leaf nitrogen content and CO₂ assimilation capacity in peach, *J. Am. Soc. Hortic. Sci.* 107 (1982) 955–959.
- [6] Gastal F., Saugier B., Relationships between nitrogen uptake and carbon assimilation in whole plants of tall fescue, *Plant Cell Environ.* 12 (1989) 407–418.
- [7] Goldschmidt E.E., Golomb A., The carbohydrate balance of alternate-bearing citrus trees and the significance of reserves for flowering and fruiting, *J. Am. Soc. Hortic. Sci.* 107 (1982) 206–208.
- [8] Huenneke L.F., Vitousek P.M., Seedling and clonal recruitment of the invasive tree *Psidium cattleianum*: Implications for management of native Hawaiian forests, *Biol. Conserv.* 53 (1990) 199–211.
- [9] Lötter J. de V., Vegetative and reproductive habit of the guava ('*Psidium guajava*' cv. Fan Retief) in relation to pruning methods, *Acta Hortic.* 275 (1990) 229–237.
- [10] Marchal J., Folliot M., Formation et utilisation des réserves en sucres et en amidon dans les parties aériennes du clémentinier en Corse, *Fruits* 47 (1992) 228–240.
- [11] Mathsoft, S-Plus® 2000 Guide to Statistics, volume 1, Data Analysis Products Division, Mathsoft, Seattle, 1999.
- [12] Migus W.N., Hunt L.A., Gas exchange rates and nitrogen concentration in two winter wheat cultivars during the grain-filling period, *Can. J. Bot.* 58 (1980) 2110–2116.
- [13] Monselise S.P., Goldschmidt E.E., Alternate bearing in fruit trees, *Hort. Rev.* 4 (1982) 128–173.
- [14] Nakasone H.Y., Paull R.E., Tropical fruits, Crop production science in horticulture 7, CAB International, Oxon, United Kingdom, 1998.
- [15] Nijs I., Behaeghe T., Impens I., Leaf nitrogen content as a predictor of photosynthetic capacity in ambient and global change conditions, *J. Biogeo.* 22 (1995) 177–183.
- [16] Normand F., Le goyavier-fraise, son intérêt pour l'île de la Réunion, *Fruits* 49 (1994) 217–227.
- [17] Normand F., Habib R., Phenology of strawberry guava (*Psidium cattleianum*) in Réunion Island, *J. Hort. Sci. Biotech.* 76 (2001) 541–545.
- [18] Popenoe W., Fruits of the Myrtle family, in: Hafner Press, (Eds), Manual of tropical and subtropical fruits, Macmillan Publishing Co, New York, 1974, 272–311.
- [19] Raunet M., Le milieu physique et les sols de l'île de la Réunion. Conséquences pour la mise en valeur agricole, CIRAD, Montpellier, 1991.
- [20] Scholefield P.B., Sedgley M., Alexander D. McE., Carbohydrate cycling in relation to shoot growth, floral initiation and development and yield in the avocado, *Sci. Hortic.* 25 (1985) 99–110.
- [21] Searle P.L., The Berthelot or indophenol reaction and its use in the analytical chemistry of nitrogen. A review, *Analyst* 109 (1984) 549–568.

[22] Shigeura G.T., Bullock R.M., Flower induction and fruit production of guava (*Psidium guajava* L.), *Acta Hortic.* 57 (1976) 247–251.

[23] Singh G., Pandey D., Rajan S., Singh A.K., Crop regulation in guava through different crop regulating treatments, *Fruits* 51 (1996) 241–246.

[24] Stassen P.J.C., Janse van Vuuren B.P.H., Storage, redistribution and utilization of starch in young bearing “Sensation” mango trees, *Acta Hortic.* 455 (1996) 151–159.

[25] Weinberger J.H., Cullinan F.P., Further studies on the relation between leaf area and size of fruit, chemical composition, and fruit bud formation in Elberta peaches, *Proc. Am. Soc. Hortic. Sci.* 29 (1932) 3–7.

III- Un modèle général de l'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise.

1- Conséquences des résultats précédents

Le goyavier-fraise a un cycle phénologique constitué d'un cycle de production annuel en milieu naturel à la Réunion (Fig. 2). En culture, l'apport d'azote entraîne l'émission d'un nouveau cycle de production (Fig. 3) dont l'intensité dépend du stade phénologique de la plante au moment de l'apport. Cette caractéristique a des conséquences au niveau agronomique pour l'agriculteur, et au niveau de la structure du modèle d'élaboration du nombre de fruits et de leur calibre.

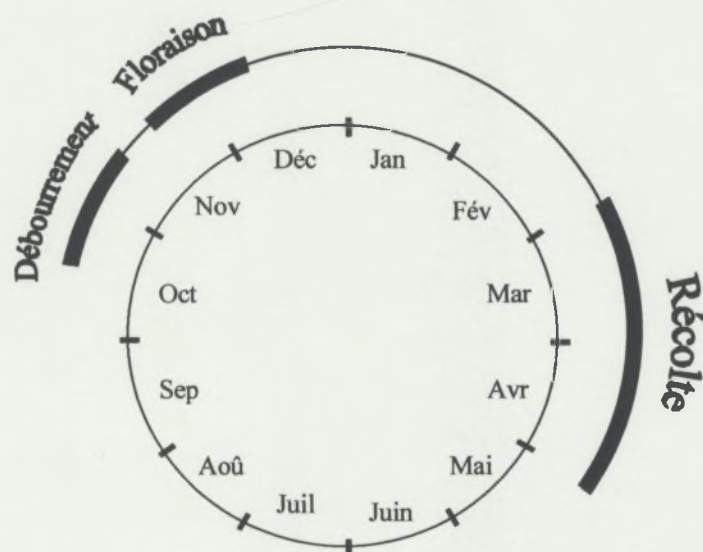


Figure 2: Le cycle phénologique naturel du goyavier-fraise à la Réunion (500 m d'altitude).
Un seul cycle de production, suivi d'un repos végétatif et reproducteur entre la récolte et le débourrement du cycle suivant.

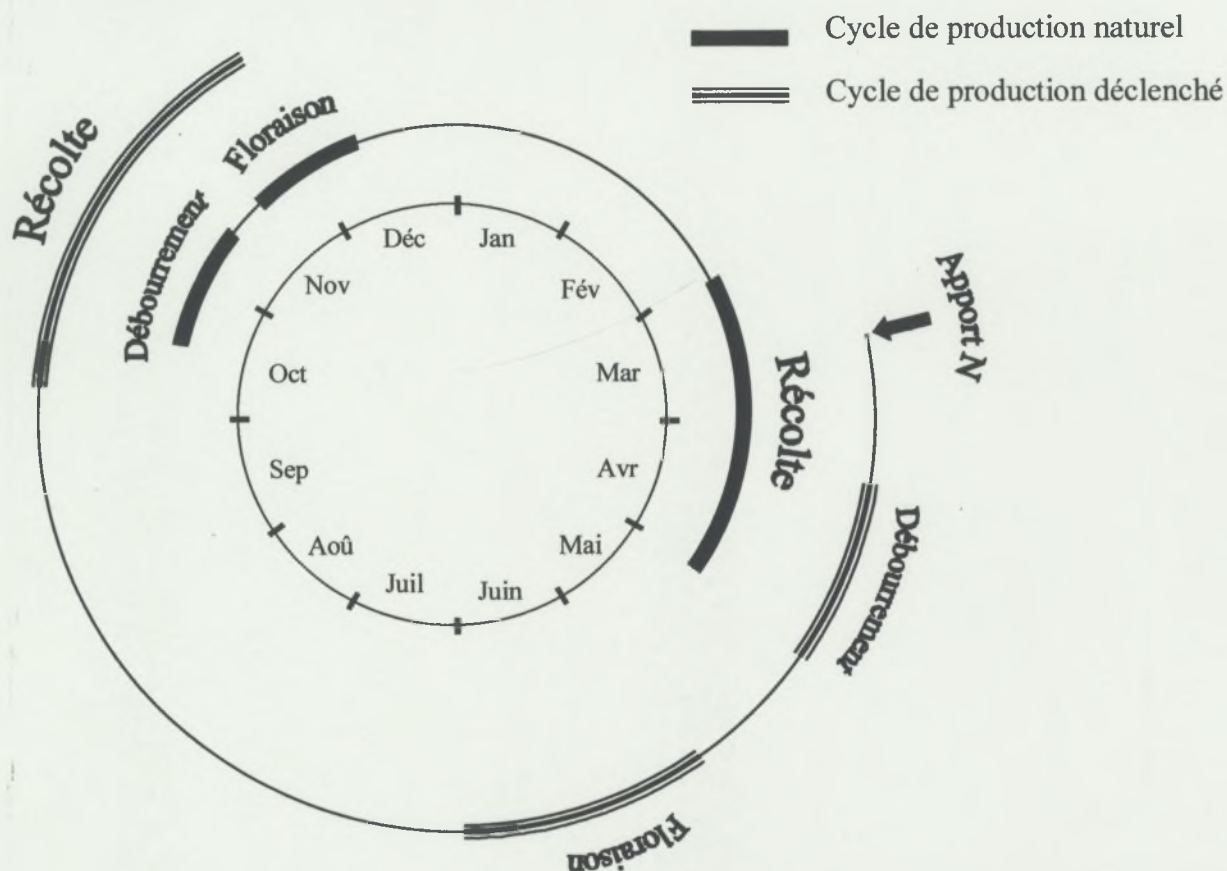


Figure 3: Phénologie du goyavier-fraise en parcelle cultivée (500 m d'altitude). Un cycle de production spontané correspondant au cycle de production naturel, et un cycle de production déclenché par un apport d'azote au cours de la récolte naturelle.

1.1- Conséquences agronomiques

Le déclenchement d'un cycle de production supplémentaire en culture permet à l'agriculteur de faire deux récoltes dans l'année: la récolte correspondant à la pousse naturelle annuelle, qui est émise spontanément, et la récolte correspondant au cycle de production déclenché par l'azote. Ainsi, la production annuelle par plant est augmentée, + 36 à 69% selon le niveau d'éclaircissage des plants (Normand et Huet, 2001). Il est possible, par une date d'apport de l'azote bien choisie, de produire à des périodes où la commercialisation est intéressante, en particulier lorsqu'il n'y a plus de fruits en milieu naturel et qu'il existe une demande de la part des transformateurs (Lapierre, 1996). On pourrait ainsi produire du goyavier-fraise toute l'année, mais l'effet négatif des pluies et des températures élevées sur la qualité des fruits limite en pratique les périodes de production d'avril à décembre.

Une forte variabilité de la production et du calibre moyen des fruits, à l'échelle du plant (Fig. 4) et de la parcelle, est observée sur les cycles de production successifs. Cette variabilité, ajoutée à celle du calibre individuel des fruits (Fig. 1), a des conséquences négatives sur la

rentabilité de la culture: baisse du rendement commercial avec les récoltes de fruits de petit calibre et difficultés parfois à honorer les engagements commerciaux, les contrats étant conclus en fin d'année pour l'année suivante.

Le déclenchement de cycles de production par de l'azote est donc un atout fort pour la domestication du goyavier-fraise, à condition de bien contrôler ces cycles pour en récupérer les retombées économiques positives et en limiter les conséquences négatives. Ce contrôle nécessite de pouvoir prévoir les périodes de récolte, le nombre de fruits produits et leur calibre, pour intervenir au besoin par des opérations techniques adaptées. Le modèle devra donc répondre à ces objectifs de prédiction.

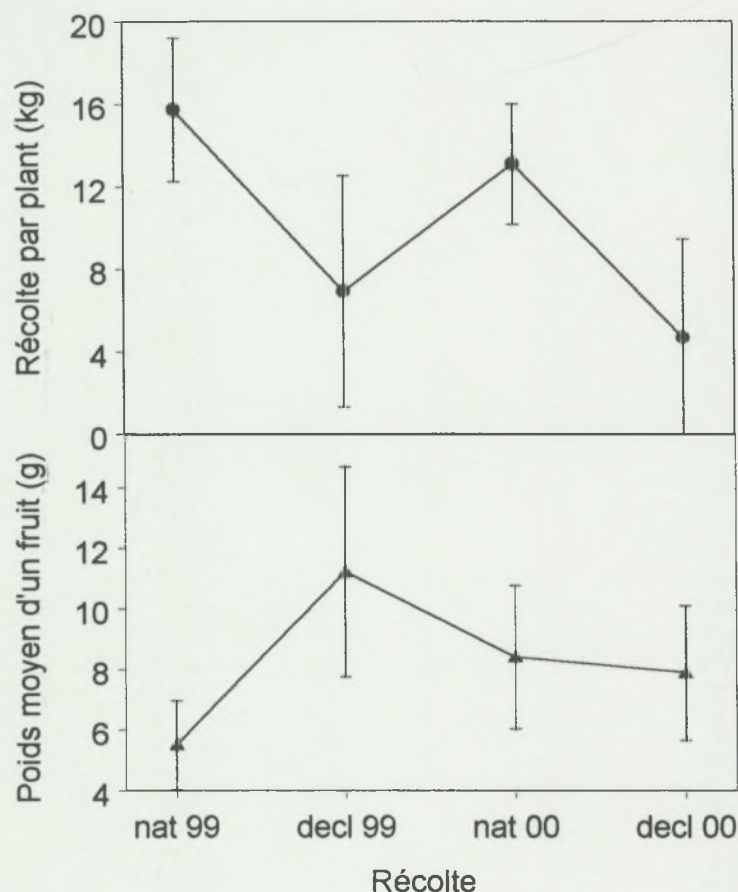


Figure 4: Variabilité de la production et du poids moyen d'un fruit par plant lors de 2 cycles naturels (nat) et 2 cycles déclenchés par fertilisation (decl) en 1999 et 2000. Moyenne \pm sd; $n = 4$ pour le premier cycle; $n = 8$ ensuite (les mêmes plants, y compris les 4 plants du premier cycle).

1.2- Conséquences sur la structure du modèle

La gestion des cycles de production du goyavier-fraise par l'agriculteur nous amène à raisonner la structure du modèle à l'échelle du cycle de production, et non plus à l'échelle du cycle phénologique annuel. Le ou les facteur(s) inducteur(s) du cycle naturel étant inconnu(s), on considère que ce dernier va du débourrement des bourgeons à la récolte. Le cycle de production déclenché va de la date d'apport de l'azote à la récolte correspondante.

La pratique de deux cycles de production par an en verger amène à des cycles consécutifs rapprochés, ou parfois superposés. Les caractéristiques d'un nouveau cycle, déclenché ou naturel, sont affectées par les caractéristiques des plants au moment de son déclenchement ou de son induction respectivement. A l'inverse, l'émission d'un nouveau cycle de production sur

un plant portant un cycle de production en cours peut affecter le déroulement de ce dernier à travers l'apport d'azote, la mise en place de nouveaux puits (débourrement), puis de nouvelles sources (nouvelles feuilles). Il est donc nécessaire de prendre en compte les effets d'un cycle de production sur le suivant dans le cas de cycles consécutifs, ce que l'on pourrait rapprocher de la "sensibilité du suivant" en grandes cultures (Sébillotte, 1990), et les effets réciproques dans le cas de deux cycles de production superposés.

Dans ce travail, nous choisissons de ne pas considérer ces effets. Pour cela, nous travaillerons sur le cycle de production naturel de la plante en culture, sans cycle de production déclenché précédent, ni déclenchement d'un nouveau cycle de production avant la fin de la récolte. Nous supposons ainsi que la plante a récupéré durant la période de repos apparent suivant la récolte précédente un niveau de réserves carbonées non limitant pour la pousse naturelle qui est étudiée. Les effets entre cycles seront pris en compte ultérieurement et ne font pas partie de notre étude.

L'étude de la phénologie du goyavier-fraise a montré que la durée des intervalles débourrement-floraison et floraison-récolte est liée à la température. Il est donc indispensable de dérouler le cycle de production sur une échelle de temps thermique. En effet, la gamme altitudinale étendue de la zone de culture du goyavier-fraise à la Réunion (0-1200 m) entraîne une gamme de températures large, qui est de plus accentuée par le déroulement possible de cycles de production en saison chaude ou fraîche sur chaque site. Les modèles thermiques statistiques présentés dans l'étude de la phénologie ont été déterminés en conditions naturelles, et non en culture. De plus, ils ne concernent pas l'intervalle apport d'azote (fertilisation) - floraison. Ils ne seront donc pas utilisés pour l'échelle de temps thermique.

La détermination de la température de base (TB) et de la somme de températures (ST) pour les intervalles fertilisation-demi floraison (50% des fleurs ouvertes) et demi floraison-demi récolte (50% des fruits récoltés) a été réalisée à partir de relevés sur des cycles de production ayant eu lieu à différentes saisons sur des parcelles à différentes altitudes (Annexe 1). Les résultats sont: $TB = 8.1^{\circ}\text{C}$ et $ST = 914^{\circ}\text{C.j}$ pour fertilisation-demi floraison, et $TB = -4^{\circ}\text{C}$ et $ST = 3134^{\circ}\text{C.j}$ pour demi floraison-demi récolte, quel que soit le cycle de production, naturel ou déclenché.

2- Structure du modèle général de l'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise

Le modèle global d'élaboration du nombre de fruits et de leur calibre que l'on souhaite construire a un objectif très appliqué de prévision du nombre de fruits produits, des calibres et des périodes de récolte, afin que l'agriculteur puisse préparer la récolte et la commercialisation et au besoin intervenir par des actes techniques adaptés. L'échelle spatiale du modèle est donc la parcelle qui est le niveau de référence de l'agriculteur. Les entrées et les sorties doivent donc être relatives à la parcelle.

Mais le modèle doit considérer différents niveaux d'échelle suivant les processus en jeu: organe, plante, parcelle. Un niveau est constitué par l'ensemble des objets du niveau inférieur ayant chacun ses caractéristiques propres. Une source de variabilité est donc explicitement liée aux différences de caractéristiques propres des objets constituant un niveau si celles-ci ont une influence sur les processus de l'élaboration du rendement. Le matériel végétal étant de semis, une première source de variabilité est liée au plant, et la parcelle doit être considérée

comme une population d'individus tous différents. Les causes de cette variabilité ne sont pas toujours explicitées. Elle peut être d'origine génétique et/ou micro-environnementale et se traduit au niveau du développement du plant, de son fonctionnement, de sa physiologie. Cette approche est une généralisation de la modélisation de vergers monoclonaux classiques (clone producteur uniquement quand il y a des pollinisateurs) où la variabilité interplant n'est souvent pas prise en compte, le verger étant génétiquement homogène et considéré comme un environnement homogène.

Des résultats préliminaires ont montré que le poids du fruit à maturité est déterminé en grande partie par son nombre de graines et le rapport feuilles:fruits du plant (nombre de feuilles / nombre de fruits). La croissance du fruit doit donc être abordée à l'échelle du fruit (nombre de graines) sachant l'arbre auquel il appartient (rapport feuilles:fruits). Comme les chutes de fruits sont rares après la fructification, les conditions de croissance des fruits sont fixées dès la fin de la floraison par les conditions de pollinisation, la fécondation et la fructification. La pollinisation dépend elle-même du déroulement de la floraison, sur chaque plant (autopollinisation) et sur les autres plants de la parcelle (allopollinisation), en faisant l'hypothèse que la quantité de pollen disponible un jour donné est proportionnelle au nombre de fleurs ouvertes ce jour, et fonction des conditions climatiques journalières (Lescourret *et al.*, 1998). Le déroulement de la floraison dépend du nombre de fleurs portées par chaque plant et est affecté par la température. Pour un cycle de production naturel, le nombre de fleurs et le nombre de feuilles d'un plant dépendent en partie de son développement spatial, dont un bon indicateur peut être la section du tronc (e.g. Shinozaki *et al.*, 1964 a,b).

Cet enchaînement, construit à partir de la connaissance de facteurs affectant la croissance du fruit, constitue l'architecture du modèle global qui est ainsi composé de 5 modèles décrivant l'intégralité du cycle de production: élaboration du nombre de fleurs, floraison, pollinisation, fécondation-fructification et croissance du fruit (Fig. 5). Les entrées de chaque modèle sont des valeurs des facteurs de variabilité qui ont été identifiés pour les processus considérés. Les sorties de chaque modèle sont des valeurs des facteurs de variabilité identifiés pour les processus du modèle suivant. Le modèle global comprend une composante quantitative, commençant par la détermination du nombre de fleurs du plant et aboutissant à son nombre de fruits et à leur calibre, et une composante temporelle qui représente le déroulement du cycle de production, en particulier la floraison et la récolte.

Compte tenu des objectifs du modèle, en particulier son utilisation pratique sur le terrain, et du peu de connaissances sur l'espèce, nous avons orienté notre travail vers des modèles descriptifs basés sur des relations statistiques. Une limite de ces modèles est leur extrapolation (Marcelis *et al.*, 1998), et il est important de préciser le domaine sur lequel il est prévu d'utiliser les modèles afin de construire les expérimentations en conséquence: représentativité des parcelles; gamme de valeurs des facteurs testés. Le domaine d'utilisation est la zone de culture potentielle du goyavier-fraise à la Réunion, caractérisée par des altitudes, et donc des températures, différentes. La prise en compte de la variabilité est un aspect central du travail de modélisation. Les facteurs *a priori* principaux affectant les différentes étapes de l'élaboration du rendement ont été envisagés à cause du peu de connaissances sur l'espèce. Il est évident qu'une partie plus ou moins importante de la variabilité ne sera pas expliquée, et il est nécessaire cependant de la considérer afin 1) de rendre compte de la variabilité globale observée à chaque étape, et 2) de pouvoir déterminer des régions de confiance pour les prédictions. Une approche stochastique de modélisation a donc été privilégiée.

Une approche déterministe d'acquisition et de gestion du carbone n'a pas été retenue à cause de la diversité génétique du matériel végétal, de la longueur des travaux nécessaires au paramétrage d'un tel modèle, et de son utilisation demandant de nombreux paramètres climatiques journaliers pas forcément disponibles chez l'agriculteur. De plus, certains processus du cycle de production (floraison, pollinisation) ne sont *a priori* pas liés au carbone.

Le modèle présenté démarre et se termine à l'échelle du plant. Afin de conserver le niveau d'intérêt pratique qu'est la parcelle, il est proposé de caractériser cette dernière par la distribution des sections de tronc de ses plants, obtenue par exemple par une mesure sur un échantillon de plants. Cette distribution intègre la variabilité entre plants propre à la parcelle et liée aux différences génétiques et environnementales locales. Une section de tronc est ensuite tirée dans cette distribution pour chaque plant. Pour la récolte, les calibres et les dates de récolte sont agrégés à l'échelle de la parcelle afin de calculer les sorties finales attendues: nombre de fruits produits, distribution des calibres et des dates de récolte.

Les principaux enjeux scientifiques de ce travail sont l'étude et la modélisation de l'élaboration du nombre de fruits et de leur calibre chez une espèce fruitière pérenne, et la caractérisation et la prise en compte de différents niveaux de variabilité dans la modélisation. Ce travail fournit aussi des données sur une espèce fruitière peu connue et peu étudiée mais présentant un potentiel agricole intéressant. Les résultats obtenus permettent de faire une analyse critique de l'approche agronomique de la domestication et d'en préciser les avantages et les limites.

Des retombées pratiques sont attendues, par une aide à une meilleure gestion des cycles de production pour les producteurs de goyavier-fraise. Ce modèle peut aussi aider à mettre au point des itinéraires techniques en simulant les effets d'actes techniques sur les caractéristiques de la récolte.

Par rapport à ce modèle global d'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise, seules les étapes de la floraison, de la fécondation-fructification, et de la croissance du fruit ont été abordées dans le cadre de cette thèse. Ces travaux sont présentés dans les trois chapitres suivants.

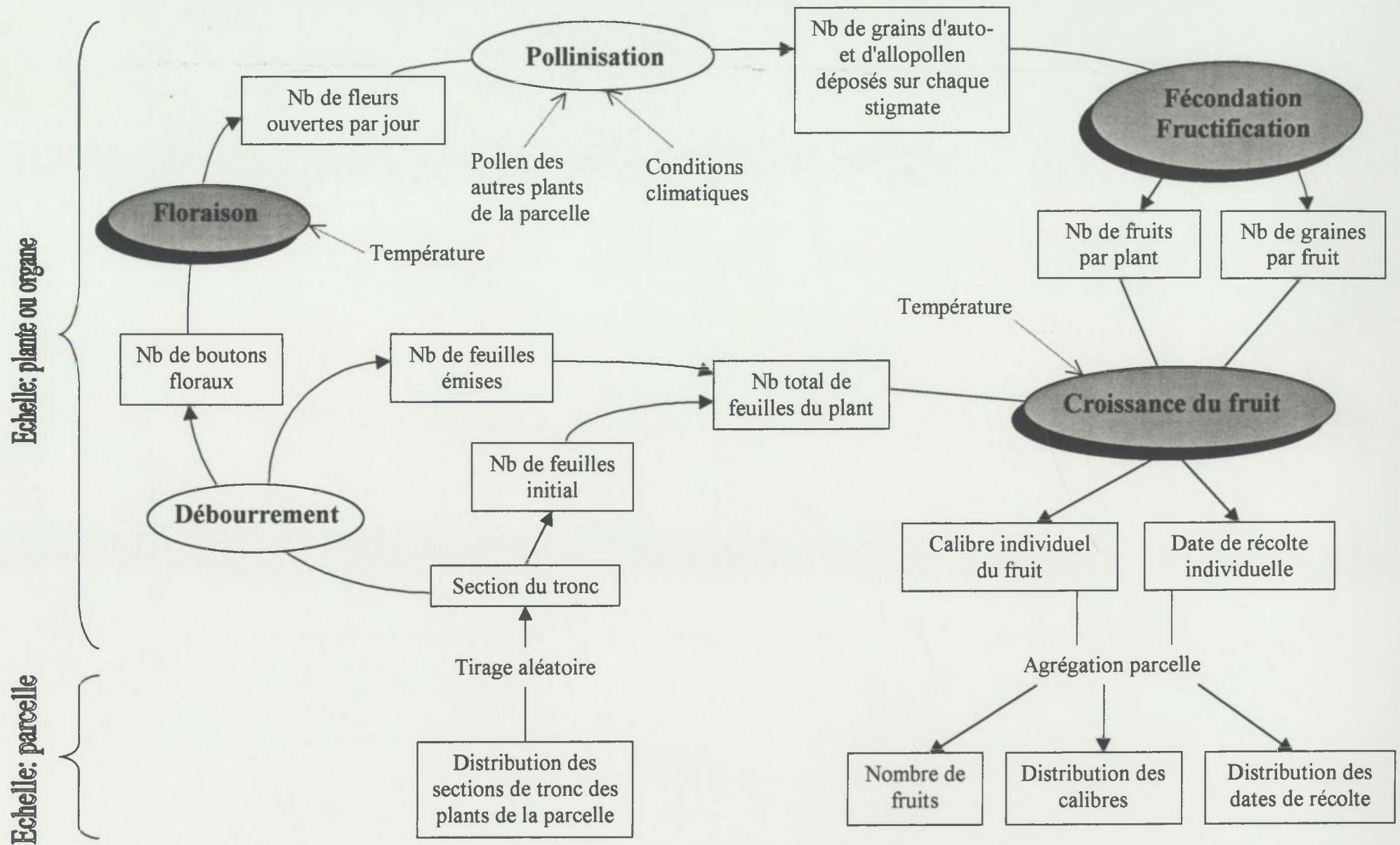


Figure 5: Organisation du modèle global d'élaboration du nombre de fruits, de leurs calibres et de leurs dates de récolte pour un cycle de production naturel chez le goyavier-fraise. Les ellipses représentent des processus et les rectangles des composantes. Les ellipses grisées sont les processus qui ont été abordés dans ce travail.

Chapitre III

Etude et modélisation de la floraison du goyavier-fraise

Le déroulement de la floraison, c'est à dire le nombre de fleurs ouvertes chaque jour, détermine en partie les conditions quotidiennes de pollinisation sous l'hypothèse que la quantité de pollen disponible est proportionnelle au nombre de fleurs ouvertes. L'étude de la phénologie du goyavier-fraise a montré que le développement des pousses florifères et des boutons floraux est sous l'influence de la température, facteur principal des modèles phénologiques. Nous faisons donc l'hypothèse que le déroulement de la floraison est affecté par la température.

L'autre facteur affectant le déroulement de la floraison est le nombre de boutons floraux. Dans le cadre du modèle global d'élaboration du nombre de fruits et de leur calibre, ce dernier est déterminé par l'étape de l'élaboration du nombre de fleurs.

La biologie de la reproduction du goyavier-fraise est caractérisée par une auto-incompatibilité partielle et variable entre plants (Chapitre IV). Dans l'objectif de la modélisation de la pollinisation et de la fécondation-fructification, le déroulement de la floraison doit être prédit à l'échelle de chaque plant (source d'autopollen) par rapport aux autres plants de la parcelle (sources d'allopollen). Un modèle qui prédit globalement la floraison à l'échelle de la parcelle n'est pas suffisant.

Dans une première partie, un modèle stochastique de floraison conçu pour un verger de kiwi est testé sur le goyavier-fraise (Normand, F. et Habib, R. Application of an existing flowering model to the strawberry guava. Analysis of discrepancies between model and data. Communication présentée au *6th International Symposium on Computer Modeling in Fruit Research and Orchard Management*, 15-18 juillet 2001, Davis, USA. A paraître dans *Acta Horticulturae*). Il reproduit mal la floraison à l'échelle de l'arbre et de la parcelle à cause de l'asynchronisme de floraison entre les plants mis en évidence dans cette étude.

Dans une seconde partie est proposé un modèle général de déroulement de la floraison tenant compte de l'asynchronisme de floraison au sein du plant et entre les plants (Normand, F., Habib, R. et Chadœuf J. A stochastic flowering model describing an asynchronously flowering set of trees. Manuscript accepté sous réserve de modifications dans *Annals of Botany*). Des hypothèses sont proposées pour simplifier le modèle. Deux modèles sont dérivés: l'un néglige l'asynchronisme de floraison au sein du plant, et l'autre néglige de plus l'asynchronisme de floraison entre les plants. Ces modèles sont ensuite appliqués au goyavier-fraise et comparés. Une démonstration complémentaire utilisée pour calculer la vraisemblance d'une floraison observée est placée en annexe 3.

Une troisième partie synthétise les travaux précédents et conclut sur le choix d'un modèle de floraison pour le goyavier-fraise dans le cadre d'un modèle global d'élaboration du rendement.

I- Application d'un modèle de floraison au goyavier-fraise. Analyse des écarts entre les sorties du modèle et les données

Application of an Existing Flowering Model to the Strawberry Guava. Analysis of Discrepancies Between the Model and Data

F. Normand
CIRAD-FLHOR
Centre de Coopération Internationale en
Recherche Agronomique pour le Développement
BP 180, 97455 Saint Pierre cedex
La Réunion
France

R. Habib
INRA - PSH Bât. A
Domaine Saint Paul
F-84914 Avignon cedex 9
France

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES 

Communication présentée au *6th International Symposium on Computer Modeling in Fruit Research and Orchard Management*, 15-18 juillet 2001, Davis, USA.

A paraître dans *Acta Horticulturae*

Keywords: *Psidium cattleianum*, flowering model, stochastic model, temperature, phenology, Réunion Island

Abstract

A precise model of flowering at both the individual and the orchard level is an important prerequisite in determining final production for fruit species with particular pollination requirements such as the strawberry guava (*Psidium cattleianum*). Most flowering models have been designed for orchards while implicitly assuming that all the trees have the same flowering pattern. Their accuracy may be affected by flowering variability among trees, linked to genetic and/or environmental factors. A stochastic, temperature driven, flowering model developed for kiwi orchards was tested on strawberry guava. This model had two parameters: a flowering threshold a and a parameter of variance relative to flowering σ^2 . Data were obtained from the flowering of 3 sets of trees, all genetically different, located at 200, 520 and 890 m of elevation on Réunion Island. Data collected at 520 m were used to fit the model, while data obtained at 200 and 890 m were used to validate it. The model described the flowering at 520 m well, but failed to account for flowering at 200 and 890 m and to represent the individual flowering patterns on each site. Analysis of the discrepancies between the model and the data showed that factors that vary among trees should be considered: number of flowers per tree, time and duration of individual flowering. The previous stochastic model has thus been re-estimated at the tree level and the over-all flowering was considered as the superposition of the individual flowerings weighted by their respective number of flowers. The model then had $2K$ (K = number of trees) parameters. It described the individual flowering pattern well, and more accurately represented the over-all flowering per site. However, a more general model with a reduced number of parameters needs to be derived from it.

INTRODUCTION

Flowering phenology is a regulating factor of yield in species for which fruit set or fruit quality depends on cross-pollination (e.g. Harrison et al., 1994; Rodrigo and Herrero 1996) and/or fruit weight depends on pollination intensity (e.g. George et al., 1986; Pyke and Alspach, 1986). Yield and fruit size in such species rely on the overlapping of the flowering of pollinizer and producer cultivars. Consequently, an accurate model which simulates the whole flowering pattern, rather than only determining the 50% flowering date (mid-bloom), is an important prerequisite in determining yield and fruit quality for these species. Some empirical functions have been proposed to describe the typical bell-shaped flowering pattern at the plant level (Medan and Bartoloni, 1998; Hof et al., 1999). At the level of a set of plants, a stochastic approach was proposed by Osawa et al. (1983) for balsam fir bud phenology, and adapted to the kiwifruit flowering by Agostini et al. (1999).

The strawberry guava (*Psidium cattleianum* Sabine) is a shrub of the Myrtaceae family native to South East Brazil, now naturalized throughout the tropics and subtropics. It has been recognized as a promising new fruit species for humid areas in Réunion Island (Indian Ocean, 21°36' S, 55°32' E). There is little phenotypic variability among plants, so that strawberry guava trees are grown from seedlings, all being genetically different. Previous unpublished results on fruit set and fruit growth showed that flowering was an important step in determining fruit size and yield as i) strawberry guava trees have variable levels of self compatibility; ii) seed set is necessary for fruit set; and iii) fruit weight is positively correlated with the number of seeds. Moreover, harvest dates can be estimated from flowering dates as the flowering to fruit maturity duration is temperature dependent (Normand and Habib, 2001). Consequently, a model accurately simulating flowering both at the orchard and at the tree level is a prerequisite to estimate fruit size, yield and harvest dates which are elements of decision for cultural practices during fruit growth (fruit thinning, fruit pest control, preparation of harvest).

This paper presents the test of a stochastic flowering model developed for kiwifruit orchard (Agostini et al., 1999) on strawberry guava.

MATERIALS AND METHODS

Data collection

Flowering of strawberry guava was recorded on three sites located at 200, 520 and 890 m elevation on the windward east coast of Réunion Island. Ten (at 520 and 890 m) or eight (at 200 m) similar sized trees (1.5 – 2 m high) were sampled. Strawberry guava flower buds are borne on new shoots emerging from the terminal branches. Twenty terminal branches were randomly sampled on each tree before bud burst. Flowers, which remain open one day, were counted daily on the shoots that emerged from the sampled terminal branches. The data set obtained at 520 m was used for parameter estimation, and the data sets at 200 m and 890 m were used for model validation. The total number of flower buds observed at 200, 520 and 890 m were 695, 2370 and 972 respectively. Differences in flower number were related to different flushing and flowering intensities of the trees among the sites. Data were expressed as a flowering probability distribution.

Mean daily air temperatures $(T_{max} + T_{min})/2$ were directly measured at 200 m, and were estimated at 520 and 890 m using the daily thermic gradient between temperatures recorded at 40 and 1025 m, as described by Normand and Habib (2001). In previous experiments, we found that the delay between the initiation of a flowering shoot and mid-bloom was 914°C.d with a base temperature $T_b = 8.1^\circ\text{C}$. Although the flowering data were collected on natural, non-triggered flowering, heat units were calculated using 8.1°C as the base temperature, and the starting date t_0 for the heat unit sum was determined on each site so

that the over-all mid-bloom occurred the day when the calculated degree-days were the closest to 914°C.d.

Model description and parameters estimation

The flower buds on a site were assumed to form an homogeneous population with respect to development time accumulation. For a flower bud, the probability of flowering between t_{i-1} and t_i is (Osawa et al., 1983; Agostini et al., 1999)¹:

$$P(t_i) = \Phi \left[\frac{a - t_i}{\sigma \cdot \sqrt{t_i}} \right] - \Phi \left[\frac{a - t_{i-1}}{\sigma \cdot \sqrt{t_{i-1}}} \right] \quad (1)$$

where $\Phi(u) = \int_u^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \cdot dx$ is the cumulative probability between u and $+\infty$ of the

standard normal distribution.

This model has two parameters related to the stochastic process: the flowering threshold a and the parameter of variance relative to flowering σ^2 (the amount of development time a flower bud has accumulated by actual time t is assumed to be normally distributed with mean t and variance $\sigma^2 t$ (Osawa et al., 1983)). The other parameters are linked to the thermal time scale: the base temperature T_b and the starting date t_0 . Parameters were estimated using the maximum likelihood method (Dacünha-Castelle and Duflo, 1982). Calculation were performed with a minimisation program for nonlinear functions of the Splus 2000 statistical package using a general quasi-Newton optimiser (Mathsoft, 1999). Comparisons among models were performed with a test maximum of likelihood (ML test, Lehman 1983). Empirical and theoretical distributions were compared using a bilateral Kolmogorov-Smirnov's test (KS test, Sprent 1989).

RESULTS AND DISCUSSION

Application of the model to the strawberry guava

The parameters estimated on the data set obtained at 520 m were $a = 914.4^\circ\text{C.d}$ and $\sigma = 3.12^\circ\text{C.d}$, with a log-likelihood $L = -8244.0$. The model fit the observed flowering pattern well (Fig. 1). Simulated and observed flowering distributions were not significantly different (KS test, $P = 0.10$). But the model failed to fit the flowering patterns at 200 m (KS test, $P < 0.001$) and 890 m (KS test, $P < 0.001$) where it did not fit the broader flowering peak observed around mid-bloom (Fig. 2 and 3). The duration of simulated flowering was quite similar on the 3 sites (592, 649 and 472°C.d respectively, at 200, 520 and 890 m), indicating that flowering was mostly driven by temperature, and the model reproduced this aspect satisfactorily on each site. The comparison of the flowering of individual trees with the model showed that the model also failed to fit them correctly. KS tests were not significant ($P > 0.05$) on 1 tree out of 8, on 1 tree out of 10 and on 2 trees out of 10 at 200, 520 and 890 m, respectively.

Thus, the model gave a good representation of the over-all flowering pattern on the site where it was fitted, but failed on the other sites and with flowering of individual trees. It was thus unacceptable for strawberry guava as it stands. Discrepancies between model and data have been studied to try to improve the model.

¹ : La justification de cette formule est présentée au chapitre suivant III-II, pages 61-62

Analysis of discrepancies between model and data

The flowering duration and the onset and mid-bloom dates of each tree on the three sites are presented in Table 1. Both time and duration of flowering showed variability among trees on a site. The differences, in days, among the dates of flowering onset of the earliest and the latest trees were 18, 23 and 20 d, respectively, at 200, 520 and 890 m. For mid-bloom, these differences were 8, 11 and 14 d, respectively. Individual flowering durations varied widely from one tree to another, and were shorter than the over-all flowering duration on the site. At the site level, there was no significant relationship between the number of sampled flowers and the individual flowering duration. However, the variability in flowering duration was somewhat explained by the number of sampled flowers when trees of the 3 sites were pooled ($n = 28$, $r^2 = 0.25$, $P = 0.006$). The flowering of individual trees was therefore, not synchronous on each site and had different patterns, which invalidates our implicit assumption of identical and synchronous flowering.

Improvement of the model

The model should be improved if we would explicitly consider the variability of flowering patterns of individual trees which are characterized by three components: the time and duration of flowering, as well as the number of flowers per tree which indicates the contribution of the tree to the over-all flowering. The flower buds on a site would therefore not be considered as an homogeneous population with respect to the development time accumulation. The differences could be linked to genetic factors (Primack, 1980; Pors and Werner 1989; Hof et al., 1999) as the strawberry guava trees were all genetically different, or to environmental factors (Primack, 1980), leading to differential development time accumulation. Each tree could thus be characterized by its own flowering threshold, a_k , which represents the mean flower date for tree k and whose difference with the over-all mid-bloom date on the site indicates the advance or the delay of the tree k flowering compared with the over-all flowering. Likewise, the individual parameter of variance relative to flowering, σ_k^2 , expresses the variation in the flowering process at the tree level and is related to the flowering duration. We thus consider the flower buds of a single tree as an homogeneous population with respect to development time accumulation, and the model defined by equation 1 is applied at the tree level. The probability of flowering between t_{i-1} and t_i for a flower bud of the tree k is:

$$P_k(t_i) = \Phi \left[\frac{a_k - t_i}{\sigma_k \cdot \sqrt{t_i}} \right] - \Phi \left[\frac{a_k - t_{i-1}}{\sigma_k \cdot \sqrt{t_{i-1}}} \right] \quad (2)$$

where $\Phi(u)$ is defined as in equation 1.

At the level of a set of K trees, the event $\{a \text{ flower opens between } t_{i-1} \text{ and } t_i\}$ is:

$$\{a \text{ flower opens between } t_{i-1} \text{ and } t_i\} = \bigcup_{k=1}^{k=K} \{ \text{the flower belongs to tree } k \} \cap \{ \text{the flower opens between } t_{i-1} \text{ and } t_i \text{ on tree } k \}$$

These K events are exclusive. If each tree bears n_k flower buds and $\sum_{k=1}^{k=K} n_k = N$, then the probability for a flower bud belonging to these K trees to open between t_{i-1} and t_i is:

$$P(t_i) = \sum_{k=1}^{K} \frac{n_k}{N} \cdot P_k(t_i) = \sum_{k=1}^{K} \frac{n_k}{N} \cdot \left[\Phi \left[\frac{a_k - t_i}{\sigma_k \cdot \sqrt{t_i}} \right] - \Phi \left[\frac{a_k - t_{i-1}}{\sigma_k \cdot \sqrt{t_{i-1}}} \right] \right] \quad (3)$$

If the total flowering of the K trees spread from t_l to t_f , we verify that $\sum_{t=t_l}^{t=t_f} P(t) = 1$.

This model has $2K$ parameters, with the number of flower buds per tree being known. It was fit to the data on each site and the estimated parameters are presented in Table 2. Flowering thresholds a_k were close to the mid-bloom dates (Table 1). The parameters of variance relative to flowering, σ_k^2 , were variable. As expected from the data analysis, their coefficient of variation at the site level was similar to that of individual flowering duration (Table 1).

The model fit the individual flowering date well (Fig. 4) as the KS tests were not significant on 7 trees out of 8, 9 trees out of 10 and 10 trees out of 10 at 200, 520 and 890 m, respectively. The model fit the overall flowering well at 520 m (KS test, $P = 0.16$), but not at 200 m ($P = 0.002$) and 890 m ($P = 0.04$). The fits were however better than with the previous model (Fig. 2, 3).

Comparison of the models

On each site, the goodness of fit of the over-all flowering given by the two models, the one designed at the site level ('site' model, eq.1) and the one designed at the tree level ('tree' model, eq.3), were compared with a ML test. As the 'tree' model was fitted on each site (Table 2), the two parameters of the 'site' model were estimated at 200 and 890 m prior to the ML tests in order to compare on each site models with parameters estimated on the same data. The tests were highly significant ($P < 0.001$) on each site, indicating that the 'tree' model gave a better goodness of fit of the over-all flowering.

At the level of individual flowering, the two models were compared in the same way than previously. They had then the same number of parameters: the two parameters a_k and σ_k^2 relative to each tree (Table 2) for the 'tree' model, and the two parameters a and σ^2 estimated on each site for the 'site' model. The likelihood was always higher with the 'tree' model, indicating that this model gave a better goodness of fit of individual flowering.

CONCLUSIONS

A stochastic flowering model designed for an orchard of kiwi vines was tested on strawberry guava. It failed to simulate individual and over-all flowering of a set of trees as time and duration of individual flowering, and the number of flowers per tree, were variable. A model simulating flowering both at the plot and at the tree level was needed for this species. The stochastic model has been improved by estimating the individual flowering and considering the over-all flowering as the superposition of individual flowering, weighted by the respective number of flower buds. A better fit of the data, both at the individual and at the plot level, was then obtained. As it stands, the model has many parameters and is only descriptive. Further work is needed to derive a more general model, reduce the number of parameters and give it predictive ability. The variability in individual flowering was probably linked to genetic factors. Flowering variability induced by environmental heterogeneity could be approached in the same way.

The model applied at the tree level did not always fit individual flowering patterns well. Empirical flowering patterns showed strong irregularities (Fig. 4), even on trees where about half of the flower buds were sampled. This indicates that flowering is not a smooth process on the tree. A more accurate model could be built by considering these irregularities, but their origin needs to be specified.

Literature Cited

- Agostini D., Habib R. and Chadoeuf J., 1999. A stochastic approach for a model of flowering in kiwifruit 'Hayward'. *J. Hort. Sci. Bio.* 74:30–38.
- Dacūnha-Castelle D. and Duflo M., 1982. Probabilités statistiques. I. Problèmes à temps fixe. Masson, Paris, 177–179.
- George A.P., Nissen R.J. and Ko H.L., 1986. Productivity of custard apple (*Annona atemoya* Hort.): factors affecting yield and fruit size. *Acta Hort.* 175:37–41.
- Harrison R.E., Luby J.J. and Ascher P.D., 1994. Pollen source affects yield components and reproductive fertility of four half-high blueberry cultivars. *J. Amer. Soc. Hort. Sci.* 119:84–89.
- Hof L., Keiser L.C.P., Elberse I.A.M. and Dolstra O., 1999. A model describing the flowering of single plants, and the heritability of flowering traits of *Dimorphotheca pluvialis*. *Euphytica* 110:5–44.
- Lehman E.L., 1983. Theory of point estimation. Wiley and Sons, New York.
- Mathsoft, 1999. S-Plus 2000 guide to statistics, volume 1. Mathsoft, Seattle, 459–500.
- Medan D. and Bartoloni N., 1998. Fecundity effects of dichogamy in an asynchronously flowering population: a genetic model. *Ann. Bot.* 81:373–383.
- Normand F. and Habib R., 2001. Phenology of strawberry guava (*Psidium cattleianum*) in Réunion Island. *J. Hort. Sci. Bio.* 76:541–545.
- Osawa A., Shoemaker C.A. and Stedinger J.R., 1983. A stochastic model of balsam fir bud phenology utilizing maximum likelihood parameter estimation. *For. Sci.* 29:478–490.
- Pors B. and Werner P.A., 1989. Individual flowering time in a goldenrod (*Solidago canadensis*): field experiment shows genotype more important than environment. *Am. J. Bot.* 76:1681–1688.
- Primack R.B., 1980. Variation in the phenology of natural populations of montane shrubs in New Zealand. *J. Ecol.* 68:849–862.
- Pyke N.B. and Alspach P.A., 1986. Inter-relationships of fruit weight, seed number and seed weight in kiwifruit. *NZ Agric. Sci.* 20:153–156.
- Rodrigo J. and Herrero M., 1996. Evaluation of pollination as the cause of erratic fruit set in apricot 'Moniqui'. *J. Hort. Sci.* 71:801–805.
- Sprenst P., 1989. Applied nonparametric statistical methods. Chapman and Hall, London, 48–57.

Tables

Table 1. Dates of flowering onset and mid-bloom, and flowering duration on individual strawberry guava trees at 3 sites. These data were obtained from a sample of flowers representative of the flower buds of each tree. The thermal time scale started on each site so that degree.days accumulation for the over-all mid-bloom, with a base temperature of 8.1°C, was the closest to 914 °C.d. C.V.: coefficient of variation (%).

Tree No.	200 m			520 m			890 m		
	Onset (°C.d)	Mid bloom (°C.d)	Duration (°C.d)	Onset (°C.d)	Mid bloom (°C.d)	Duration (°C.d)	Onset (°C.d)	Mid bloom (°C.d)	Duration (°C.d)
1	824.1	914.4	325.7	621.4	868.9	356.3	783.4	894.2	384.8
2	794.8	900.0	307.6	881.3	965.3	271.8	954.7	1007.4	149.0
3	779.6	854.9	240.9	721.0	868.9	344.4	840.1	923.5	339.1
4	870.1	930.0	166.3	821.3	952.8	319.3	792.8	894.2	225.4
5	763.6	959.2	501.1	721.0	929.4	495.3	755.8	894.2	262.3
6	703.5	930.0	332.9	743.1	929.4	348.9	755.8	871.5	284.8
7	672.9	914.4	429.5	710.2	881.3	329.3	860.7	1018.1	274.0
8	930.0	959.2	90.4	857.6	1002.5	397.9	783.4	933.6	309.5
9	-	-	-	721.0	881.3	549.4	860.7	954.7	243.0
10	-	-	-	832.9	977.7	422.7	851.5	997.3	376.5
Mean	792.3	920.3	299.7	763.1	925.8	383.5	823.9	938.9	284.8
C.V.	10.6	3.7	44.5	10.7	5.2	22.1	7.5	5.7	25.2

Table 2. Number of flowers and parameters of the stochastic process in the flowering model estimated for each tree on the three sites. C.V.: coefficient of variation (%).

Tree No.	200 m			520 m			890 m		
	No. of flowers	a_k	σ_k	No. of flowers	a_k	σ_k	No. of flowers	a_k	σ_k
1	104	918.7	1.60	287	837.7	2.60	67	913.8	3.21
2	147	902.8	1.70	185	962.5	1.51	24	1003.0	0.94
3	16	897.7	3.18	233	872.0	2.43	48	944.7	2.65
4	83	924.1	1.41	228	952.2	2.45	96	885.0	1.83
5	86	985.2	3.15	275	919.4	3.00	176	888.5	1.86
6	119	915.2	2.18	169	962.8	2.91	137	861.7	2.01
7	124	909.7	2.67	314	884.2	2.44	51	985.3	2.51
8	16	941.3	1.15	286	998.9	2.49	239	933.4	2.16
9	-	-	-	301	875.2	2.98	48	945.1	1.90
10	-	-	-	92	973.1	2.91	86	993.4	2.47
Mean	86.9	924.3	2.13	237.0	923.8	2.57	97.2	935.4	2.15
C.V.	55.7	3.0	37.2	29.8	5.8	17.3	69.7	5.2	28.3

Figures

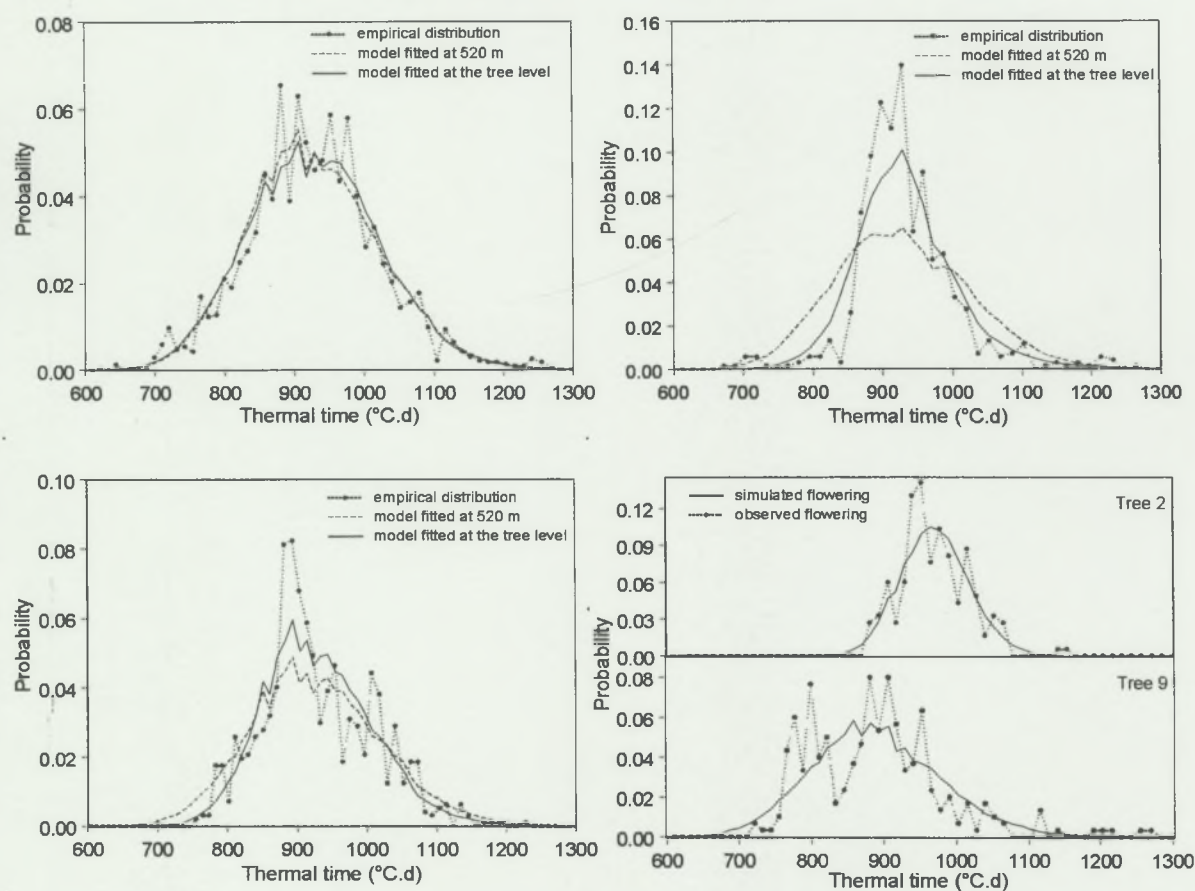


Fig. 1 (upper left). Comparison between the observed and simulated flowering probability distribution at 520 m with the model fitted on all the flowers and the model fitted at the tree level.

Fig. 2 (upper right). Comparison between the observed and simulated flowering probability distribution at 200 m with the model fitted on all the flowers at 520 m and the model fitted at the tree level.

Fig. 3 (lower left). Comparison between the observed and simulated flowering probability distribution at 890 m with the model fitted on all the flowers at 520 m and the model fitted at the tree level.

Fig. 4 (lower right). Flowering model fitted at the tree level on trees 2 and 9 at 520 m showing the differences in timing, duration and shape of flowering of individual trees.

II- Un modèle stochastique pour décrire la floraison de plants à floraison asynchrone

A stochastic flowering model describing an asynchronously flowering set of trees.

F. NORMAND

CIRAD-FLHOR, BP 180, 97455 Saint Pierre cedex, Réunion Island, France

tel: (262) 2 62 50 62 74 fax: (262) 2 62 50 58 44

R. HABIB

INRA, Plante et Systèmes de Culture Horticoles, Bât. A, Domaine St Paul, Site Agroparc, F-84914 Avignon cedex 9, France

tel: (33) 4 32 72 24 34 fax: (33) 4 32 72 24 32

J. CHADŒUF

INRA, Biométrie, Domaine St Paul, Site Agroparc, F-84914 Avignon cedex 9, France

tel: (33) 4 32 72 21 76 fax: (33) 4 32 72 21 82

Accepté sous réserve de modifications dans *Annals of Botany*

ABSTRACT

The paper presents a general stochastic model to simulate the time course of flowering of individual trees and of their population integrating the synchronisation of flowering both between and within trees. Under some hypotheses, a simplified expression of the model, called "shoot" model, is proposed, where both the synchronisation of flowering between and within trees are characterised by one parameter. Two derived models, the "tree" model and the "population" model, are presented. They neglect the asynchrony of flowering respectively within trees, and between and within trees. The models have been fitted and tested with data on *Psidium cattleianum* flowering observed at 200, 520 and 890 m of elevation in Réunion Island. The "shoot" model better fits the data and can reproduce the strong irregularities in flowering observed on empirical data. The asynchrony of flowering in *P. cattleianum* is more pronounced within than between the trees. Simulations show that various flowering patterns can be reproduced by the "shoot" model. The use of the different levels of organisation of the general model is discussed.

Key words: Flowering asynchrony, flowering model, stochastic model, degree.days, *Psidium cattleianum* Sabine, strawberry guava.

INTRODUCTION

The time course of the flowering of individual trees and of their population, i.e. the number of flowers daily open during the flowering duration, has important consequences on the reproductive success and the genetic structure of the population (Bawa, 1983; Ims, 1990; Murawski and Hamrick, 1991; Hof *et al.*, 1999). In particular, seed and fruit production of self incompatible species relies on the overlapping of the flowering of compatible genotypes, with important economical consequences in agriculture. The time course of flowering of a population could also be considered in relation to the sensitivity of flowers to pathogens in order to aid the development of integrated pest and disease management (Dodd *et al.*, 1992), or for the human health when the prediction of airborne pollens is used to forecast allergic risks (Frenguelli *et al.*, 1989; Ickovic *et al.*, 1989; Andersen, 1991; Belmonte and Roure, 1991).

Plants, in particular tropical trees and shrubs, display a large variety of flowering patterns (Gentry, 1974; Bawa, 1983). Asynchronous flowering is widespread (Primack, 1980; Bawa, 1983; Ims, 1990; Carthew, 1993), even in mass-flowering species (Augspurger, 1983). The time course of the population flowering, as the superposition of the time course of individual flowerings, is therefore affected by the time and duration of individual flowering, as well as by the individual number of flowers. Two main components can then be distinguished in the synchronisation of flowering (Hof *et al.*, 1999): the synchronisation between plants and the synchronisation within plants. It was shown for several woody and non-woody species that the flowering phenology is mainly under genetic control (Nienstaedt, 1974; Primack, 1980; Mosseler and Papadopol, 1989; Pors and Werner, 1989; Carthew, 1993; O'Brien and Calder, 1993; Boudry *et al.*, 1994; Mitchell-Olds, 1996; Hof *et al.*, 1999), implying to consider the individual level in a detailed study of the flowering phenology (e.g. Augspurger, 1983; Fripp *et al.*, 1987; Carthew, 1993). The flowering synchronisation improvement and the flowering period were important objectives for breeding and selection programs (Janick and Moore, 1975; Hof *et al.*, 1999; Citadin *et al.*, 2001). The flowering phenology is also affected by the environment (Murfet, 1977; Primack, 1980).

Temperature is recognised as being the main variable which drives the timing of budburst or flowering in woody plants, and different models have been proposed to predict the date of flowering of a plant population. They all assume that budburst, or flowering, occurs when a critical development threshold a is reached, the stage of development being a sum of daily rates of development. They differ in the expression of the rate of development as a function of temperature (e.g. Hänninen, 1987; Chuine *et al.*, 1998, 1999), and in the way they integrate chilling requirements for temperate trees (Cannell and Smith, 1983; Hänninen, 1987; Murray *et al.*, 1989; Kramer, 1994; Chuine *et al.*, 1998, 1999). The predicted date of flowering is the day when the critical development threshold is reached. It generally corresponds to the onset of bloom or to the mid-bloom date, when 50% of the flowers are open, or to the date of maximum concentration of airborne pollen (Boyer, 1973; Chuine *et al.*, 1998, 1999).

But very little attention has been paid to the modelling of the time course of flowering of a tree or a population. The time course of flowering at the tree level is the result of a within tree flowering variability. This variability may be *a priori* related to genetic factors, to physiological factors, to the time-lag in flower induction or in bud break, or to the buds' effective temperature. These factors and their influence are complex to determine and quantify, and their relative importance may vary among species and environment. An

approach for a general model of the time course of flowering is to reproduce the effects of the processes underlying the time course of flowering and not to explain these processes. Thus, Agostini *et al.* (1999) proposed a stochastic flowering model for a kiwifruit (*Actinidia deliciosa* (A. Chev.) C.F. Liang & A.R. Ferguson) orchard of female vines (see details in the Materials and Methods section). In such a clonal population, the time course of individual flowering is expected to be reasonably similar, with a strong synchronisation of flowering between plants. If flowering is a regular process on individuals and the environment is homogeneous, then the time course of flowering for the population and for individuals should not differ, but on the number of flowers open per day. The overlapping of the flowering of male (described as previously but with a different set of parameters) and female individuals is then needed to ensure pollination and fruit set (Lescouret *et al.*, 1999).

More generally, in a non clonal population, such a model would be inappropriate because of asynchronous flowering and differences in the time course of individual flowerings (e.g. Primack, 1980; Augspurger, 1983; Fripp *et al.*, 1987; Carthew, 1993). Simple empirical functions have been proposed to describe the time course of flowering at the individual level with plant specific parameters (Fripp *et al.*, 1987; Medan and Bertoloni, 1998; Hof *et al.*, 1999), but without integration at the population level.

The aim of this paper is to propose a theoretical stochastic model to simulate the time course of flowering of individual trees and of their population integrating the synchronisation of flowering both between and within trees. A general model is presented and hypotheses are proposed to simplify it and reduce the number of parameters. Derived models neglecting the within tree or the within and between trees variability are also presented. Procedure to estimate the parameters of the models are proposed. The models are then fitted and tested with experimental data obtained in the subtropical Reunion Island on populations of strawberry guava (*Psidium cattleianum* Sabine) growing in natural conditions at three elevations. The simplifying assumptions used for the model construction are tested on *P. cattleianum*. The structure and the uses of the models are then discussed.

MATERIALS AND METHODS

Model description

The heart of the model was developed by Osawa *et al.* (1983) to describe Balsam Fir (*Abies balsamea* (L.) Mill.) bud phenology. Dennis *et al.* (1986) adapted it to insect development, and Agostini *et al.* (1999) to kiwifruit flowering. We recall here its assumptions and formulation. It assumes that the development of a given bud, here a flower bud, is a stochastic process consisting in the accumulation of small increments of development u_d . This process is supposed to begin with the flower bud in a given stage at time $t = 0$. The process $S(t)$ is defined as the amount of development time a flower bud has accumulated by actual time t : $S(t) = \sum_{d \leq t} u_d$. The amount of development time $S(t)$ and t are expressed in degree-days. If the

increment of time Δd is small and the increments of development u_d are independent and identically distributed with an expectation $E(u_d) = \Delta d$, then $S(t)$ is normally distributed with mean t and variance $\sigma^2 t$ (Osawa *et al.*, 1983). Flowering occurs when the amount of development $S(t)$ passes through a threshold a . So the probability a given flower bud is flowering at time t is:

$$p(t) = \text{Prob}[S(t) > a] = \Psi \left[\frac{a - t}{\sigma \cdot \sqrt{t}} \right] \quad (1)$$

where $\Psi(u) = \int_u^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \cdot dx$ is the cumulated probability between u and $+\infty$ of

the standard normal distribution.

And the probability a given flower bud is flowering between t_{i-1} and t_i is:

$$P(t_i) = \Psi \left[\frac{a - t_i}{\sigma \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{a - t_{i-1}}{\sigma \cdot \sqrt{t_{i-1}}} \right] \quad (2)$$

The implicit assumption of this model is that the flower buds follow the same stochastic process, i.e. they accumulate increments of development in the same way and have the same flowering threshold.

Let now consider a population of K trees, and a unique development time scale at the population level, which is the most common case. The synchronisation of flowering between trees is related to different individual flowering patterns. We assume that the synchronisation of flowering within tree is related to the different flowering patterns of the tree flowering units. A flowering unit is a level of organisation whose flower buds are in a similar stage of development and experience a similar environment. They have consequently about the same pattern of development. The variability between flowering units could be linked to the micro-environment (e.g. effective temperature) or to the unit age (e.g. time-lag in bud break or in flower bud initiation). For convenience, we will later call the flowering units as flowering shoots. However, the architectural level of a flowering unit needs to be determined for each species (e.g. inflorescence for some species, ...). This variability results in a variable flowering threshold a and a variable parameter of variance relative to flowering, σ^2 , at the shoot and at the tree levels. Increasing values of the parameter of variance relative to flowering are expected when the level of organisation increases (shoot < tree < population).

In the population of K trees, each tree p has a given number FS_p of flowering shoots. A flowering shoot, noted $fs_{j,p}$ ($j \in [1, FS_p]$), bears $N_{j,p}$ flower buds and is characterised by a flowering threshold $a_{j,p}$ and a parameter of variance relative to flowering $\sigma_{j,p}^2$ common to all its flower buds. Eqn (2) can be applied on the flower buds of a flowering shoot with the specific parameters $a_{j,p}$ and $\sigma_{j,p}^2$ as they are supposed to have about the same pattern of development. Let consider the expectation of $a_{j,p}$, $a_p = E(a_{j,p})$ ($j \in [1, FS_p]$), as the tree flowering threshold for the tree p . The difference between the flowering threshold of the flowering shoot and the tree flowering threshold, $d_{j,p} = a_{j,p} - a_p$, expressed in thermal time units, represents the advance ($d_{j,p} < 0$) or the delay ($d_{j,p} > 0$) of the mid-bloom on the shoot j compared with the mid-bloom on all the shoots of the tree. By construction, the expectation of $d_{j,p}$ is zero.

Therefore, the probability of flowering between t_{i-1} and t_i for a flower bud on the shoot $fs_{j,p}$, $j \in [1, FS_p]$, $p \in [1, K]$, is:

$$P_{j,p}(t_i) = \Psi \left[\frac{(a_p + d_{j,p}) - t_i}{\sigma_{j,p} \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{(a_p + d_{j,p}) - t_{i-1}}{\sigma_{j,p} \cdot \sqrt{t_{i-1}}} \right] \quad (3)$$

This is the general form of the flowering model, where a_p , $d_{j,p}$ and $\sigma_{j,p}^2$ are parameters. It means that all flower buds on the shoot $fs_{j,p}$ have the same probability to flower between t_{i-1} and t_i . This probability depends only on time and on the model parameters. Therefore, it assumes that the flowering of a given flower bud on the shoot $fs_{j,p}$ is not affected by the other flower buds of the shoot. In other terms, the flower buds are independent as regards their probability to flower.

The model parameters are numerous and their number $(K + 2 \cdot \sum_{p=1}^{p=K} FS_p)$ depends on the number of shoots and trees. In this way, the general formulation is not easy to use for prediction and parameter values are specific of the data used to estimate them. We propose three hypotheses to simplify the model and reduce the number of parameters:

- 1- the differences $d_{j,p}$ follow the same normal distribution $N(0, \delta^2)$ on all the trees,
- 2- the parameter of variance relative to the flowering at the shoot level, $\sigma_{j,p}^2$, is independent of the shoot and the tree, and is called σ_s^2 ,
- 3- the tree flowering thresholds a_p follow a normal distribution $N(\alpha_s, \beta_s^2)$.

With the first and the second hypothesis, we can calculate from eqn (3) the probability of flowering between t_{i-1} and t_i for a flower bud on the tree p , unconditional on the shoot:

$$P_p(t_i) = \int_{-\infty}^{+\infty} \left[\Psi \left[\frac{(a_p + u) - t_i}{\sigma_s \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{(a_p + u) - t_{i-1}}{\sigma_s \cdot \sqrt{t_{i-1}}} \right] \right] \cdot g(u) \cdot du = h(a_p, t_i) \quad (4)$$

where $g(u) = \frac{1}{\delta \cdot \sqrt{2\pi}} \cdot e^{-\frac{u^2}{2\delta^2}}$ is the density function of the normal distribution $N(0, \delta^2)$.

With the third hypothesis, we can calculate from eqn (4) the probability of flowering between t_{i-1} and t_i for a flower bud, unconditional on the shoot and the tree:

$$P(t_i) = \int_{-\infty}^{+\infty} P_p(t_i) \cdot f(a) \cdot da = \int_{-\infty}^{+\infty} h(a, t_i) \cdot f(a) \cdot da \quad (5)$$

where $f(a) = \frac{1}{\beta_s \cdot \sqrt{2\pi}} \cdot e^{-\frac{(a-\alpha_s)^2}{2\beta_s^2}}$ is the density function of the normal distribution

$N(\alpha_s, \beta_s^2)$. Or:

$$P(t_i) = \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} \left[\Psi \left[\frac{(a + u) - t_i}{\sigma_s \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{(a + u) - t_{i-1}}{\sigma_s \cdot \sqrt{t_{i-1}}} \right] \right] \cdot g(u) \cdot du \right\} \cdot f(a) \cdot da \quad (6)$$

where $g(u)$ and $f(a)$ are as in eqn (4) and (5) respectively.

This formulation of the model is of general use and has only four parameters integrating the synchronisation of flowering between and within trees:

- σ_s^2 , parameter of variance relative to the flowering at the shoot level, which expresses the time course of the flowering process at the shoot level.
- δ^2 , variance of the central normal distribution of the differences $d_{j,p}$ on the tree, which expresses the synchronisation of shoots flowering within a tree. The higher δ^2 , the longer the flowering of a tree.
- α_s , mean of the normal distribution of the tree flowering thresholds which is the mean flowering threshold of the flowering shoots population when taking into account the flowering of individual shoots.
- β_s^2 , variance of the normal distribution of the trees flowering thresholds, which expresses the flowering synchronisation between trees. The lower β_s^2 , the more synchronous the flowering between trees.

To evaluate the relevance of this model called hereafter the "shoot" model, two simpler models have been derived from it for comparison. The first one, called "tree" model, considers that the within tree variability can be neglected (i.e., the flowering units of a tree are quasi-synchronous): $d_{j,p} = 0$. The effect of trees on flowering is then characterised by their respective tree flowering threshold a_p' , distributed according to a normal distribution $N(\alpha_t, \beta_t^2)$. We hypothesise that the parameter of variance relative to flowering, σ_t^2 , is independent of trees. This model has three parameters: α_t , β_t^2 and σ_t^2 and the probability of flowering between t_{i-1} and t_i for a flower bud, unconditional on the tree, is:

$$P(t_i) = \int_{-\infty}^{+\infty} \left[\Psi \left[\frac{a - t_i}{\sigma_t \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{a - t_{i-1}}{\sigma_t \cdot \sqrt{t_{i-1}}} \right] \right] \cdot f(a) \cdot da \quad (7)$$

where $f(a) = \frac{1}{\beta_t \cdot \sqrt{2\pi}} \cdot e^{-\frac{(a - \alpha_t)^2}{2\beta_t^2}}$ is the density function of the normal distribution

$N(\alpha_t, \beta_t^2)$.

The second one, called "population" model, considers that there is moreover no effect of the trees on flowering, i.e. that their respective flowering are quasi-synchronous at the population level. Trees have the same flowering threshold a , and there is a parameter of variance relative to the flowering of the whole population: σ^2 . This model, similar to that of Agostini *et al.* (1999), has two parameters a and σ^2 and is described by eqn (2).

Model simulation

The number of flower buds of each flowering shoot is necessary to perform simulation with the "shoot" model, so that their respective weight in the total flowering is taken into account. A development time scale is also required. Simulations are realised as follows. Firstly, a flowering threshold a_p is randomly sampled for each tree in the normal distribution $N(\alpha_s, \beta_s^2)$.

Then, a difference $d_{j,p}$ is randomly sampled for each flowering shoot of each tree in the normal distribution $N(0, \delta^2)$. Given a_p , the flowering threshold $a_{j,p}$ of each flowering shoot is calculated: $a_{j,p} = a_p + d_{j,p}$. Given $a_{j,p}$, the flowering probability distribution (FPD) of the $N_{j,p}$ flower buds of the shoot is determined by eqn (3), using $\sigma_{j,p}$ or σ_s according to the level of complexity of the model used. A flowering date is finally randomly sampled in FPD for each of the $N_{j,p}$ flower buds of the shoot. Simulated flowering dates are then aggregated at the shoot, at the tree or at the population level.

To illustrate the effect of between and within tree synchronisation on the flowering pattern, a simulation of the flowering of a five tree population has been carried out in four distinct cases of synchronisation:

- A- asynchrony strong between trees and weak within tree ($\beta_s = 250^\circ\text{C.d}$; $\delta = 2^\circ\text{C.d}$)
- B- strong asynchrony between and within trees ($\beta_s = 150^\circ\text{C.d}$; $\delta = 150^\circ\text{C.d}$)
- C- asynchrony weak between trees and medium within trees ($\beta_s = 20^\circ\text{C.d}$; $\delta = 100^\circ\text{C.d}$)
- D- weak asynchrony between and within trees ($\beta_s = 20^\circ\text{C.d}$; $\delta = 2^\circ\text{C.d}$)

The other parameters are set at $\alpha_s = 1000^\circ\text{C.d}$ and $\sigma_s = 0.7^\circ\text{C.d}$. The development time scale runs from 500 to 1500°C.d , with a 10°C.d daily increment. Each tree has the same weight: 100 flowering shoots, each with 15 flower buds, i.e. 1500 flower buds per tree.

Data

Flowering of *Psidium cattleianum*, strawberry guava tree, was recorded from November 1998 to January 1999 on three sites located at 200, 520 and 890 m of elevation on the windward East coast of the subtropical Réunion Island, Indian Ocean (lat. $21^\circ 06'$ S, long. $55^\circ 32'$ E). Sites at 200 and 890 m were fallow lands invaded by feral strawberry guava trees. At 520 m, trees were from an experimental five year-old seedling orchard. They did not receive any cultural care for the previous year and were considered as feral. All individuals were genetically different on all the sites.

On each site, ten trees of similar size (1.5 – 2 m high) were sampled. Flower buds are borne on new shoots emerging from the terminal branches at the end of the cool and dry season. Mostly one to three shoots emerge quasi simultaneously from a same terminal branch (Normand and Habib, 2001). These shoots and their flower buds have therefore the same age and experience the same micro-environment. Consequently, we considered all the flowering shoots emerging from a terminal branch as the flowering unit used in the model presentation, and it is called hereafter "flowering shoot". Terminal branches were randomly sampled in the tree canopy before bud burst: 20 on eight trees and 100 on two trees at 520 m, and 20 per tree at 200 and 890 m. On each tree with 100 observed shoots, 20 shoots were randomly sampled among the 100 shoots, and the parameter estimation was made using these 20 sampled shoots to avoid a higher weight of these trees in the likelihood estimation due to their higher number of observed flowers. On the other hand, the full 100-shoot data sets were used to estimate the distribution of parameters $d_{j,p}$ and $\sigma_{j,p}$ on a large sample in order to test the simplifying hypotheses made to obtain eqn (7). Each flower stays open one day: the flower bud bursts in the morning (7 – 9 a.m.) and anthers and petals dry and fall in the late afternoon. On each site, open flowers were daily counted on the shoots emerged from the sampled terminal branches. The data set obtained at 520 m was used for parameter estimation and model comparison, and the data sets at 200 m and 890 m were used for the model validation.

Mean daily air temperatures $(T_{max} + T_{min})/2$ were estimated at 520 and 890 m using the daily thermic gradient between temperatures recorded at 40 and 1025 m, as described by Normand and Habib (2001). Mean daily air temperatures were directly determined at 200 m. The development time scale used was a thermal time scale (Cannell and Smith, 1983). Data collected in orchards over several years and elevations showed that the delay between the triggering of a flowering shoot with fertilisation and mid-bloom was 914°C.d with a base temperature $T_b = 8.1^{\circ}\text{C}$. Although flowering data for this study were collected on natural non triggered flowering, heat units were calculated using 8.1°C as base temperature, and the starting date t_0 for the heat units sum was determined on each site so that mid-bloom occurred the day when the calculated degree-days were the closest to 914°C.d .

To facilitate comparisons between sites, empirical and simulated time course of flowering were expressed as flowering frequency distribution (FFD). The flowering frequency

at day t_i was: $F(t_i) = \frac{n(t_i)}{N}$ where $n(t_i)$ was the number of flowers open at day t_i and N was the total number of observed flowers.

Statistical analysis

For each model, parameters are estimated by the maximum likelihood method (Dacūnha-Castelle and Duflo, 1982) on data giving in the course of time the number of open flowers on a sample of observed flowering shoots from a sample of M trees. Each tree p has FSO_p flowering shoots, called $fso_{j,p}$, each bearing $NO_{j,p}$ flower buds ($j \in [1, FSO_p]$, $p \in [1, M]$). Observations are made from t_1 to t_D and cover the whole flowering of the M trees. At time t_i ($i \in [1, D]$), $n_{j,p}(t_i)$ flower buds burst into bloom on the shoot $fso_{j,p}$. Eqn (3) assumes that flower buds are independent with regard to their probability of flowering, so that the probability of the observed time course of flowering from t_1 to t_D on the shoot $fso_{j,p}$ follows a multinomial

distribution² and is $\frac{NO_{j,p}!}{\prod_{i=1}^D n_{j,p}(t_i)!} \prod_{i=1}^D [P_{j,p}(t_i)]^{n_{j,p}(t_i)}$. The multinomial coefficient depends

only on the sample size and does not influence the parameter estimation and the maximum likelihood tests. It is then removed from the following calculations.

Furthermore, we hypothesise that flowering units on a tree are independent with regard to their probability of flowering. Then the likelihood of the sample for the "shoot" model is given by:

$$L = \prod_{p=1}^{p=K} \left\{ \prod_{j=1}^{j=FSO_p} \left[\prod_{i=1}^{i=D} \left(\Psi \left[\frac{(a+u)-t_i}{\sigma_s \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{(a+u)-t_{i-1}}{\sigma_s \cdot \sqrt{t_{i-1}}} \right] \right)^{n_{j,p}(t_i)} \cdot g(u) \cdot du \right] \cdot f(a) \cdot da \right\} \quad (8)$$

$$\text{where } \Psi(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}} \cdot dy; \quad g(u) = \frac{1}{\delta \cdot \sqrt{2\pi}} \cdot e^{-\frac{u^2}{2\delta^2}} \quad \text{and} \quad f(a) = \frac{1}{\beta_s \cdot \sqrt{2\pi}} \cdot e^{-\frac{(a-a_s)^2}{2\beta_s^2}}$$

² : L'équivalence entre l'approche avec une loi multinomiale et celle avec les probabilités conditionnelles est montrée en annexe 3.

The likelihood of a sample for the "tree" and the "population" models are simpler expressions.

The likelihood estimates of the parameters σ_s^2 , δ^2 , α_s , and β_s^2 are the values which maximise L , or minimise $-\log(L)$. Calculations were performed with a minimisation program for nonlinear function of Splus 2000 statistical package using a general quasi-Newton optimiser (Mathsoft, 1999) on the $M=10$ trees (2370 flower buds) observed at 520 m.

Computing limitations did not permit the minimisation of $-\log(L)$ for the "shoot" model including two nested integrals [eqn (8)]. We have estimated the parameters (α_s , β_s , δ , σ_s) by selecting the values which minimised $-\log(L)$ in a range of values centred on a first estimation derived from the data (α_s , δ and σ_s were the means of these parameters estimated at the tree level, and β_s was the standard deviation of individual flowering thresholds). Although not optimal, this method leads to asymptotically unbiased parameters whose variance tends to zero when the number of flowers tends to infinity.

Test of maximum likelihood (Lehman, 1983) was used to compare the 3 models of increasing complexity "population", "tree" and "shoot", and to test the pertinence of some hypotheses made for the model construction. The test evaluates the increase of likelihood between two models compared with the difference in their numbers of parameters. For a pair comparison of n models, twice the difference of log-likelihood, calculated with the estimated parameters on a data set, were compared to a chi-square value for $P = 5/n \%$ and a number of degrees of freedom equal to the difference between their respective parameters numbers.

Empirical and theoretical distributions were compared using a bilateral Kolmogorov-Smirnov's test (Sprent, 1989).

RESULTS

Model simulations

Simulations performed with the "shoot" model in four situations of between and within tree synchronisation are presented in Fig. 1. The main flowering patterns are well represented by the model (Gentry, 1974; Bawa, 1983), in particular the case of a strong asynchrony between trees which could not be modelled by classical flowering models (Fig. 1A). These patterns may be moreover modified by different numbers of flower buds on the shoots and trees. The flowering of individual trees are also given by the simulation (data not shown), and can be useful to quantify the degree of overlap in flowering among trees (Primack, 1980, Augspurger, 1983).

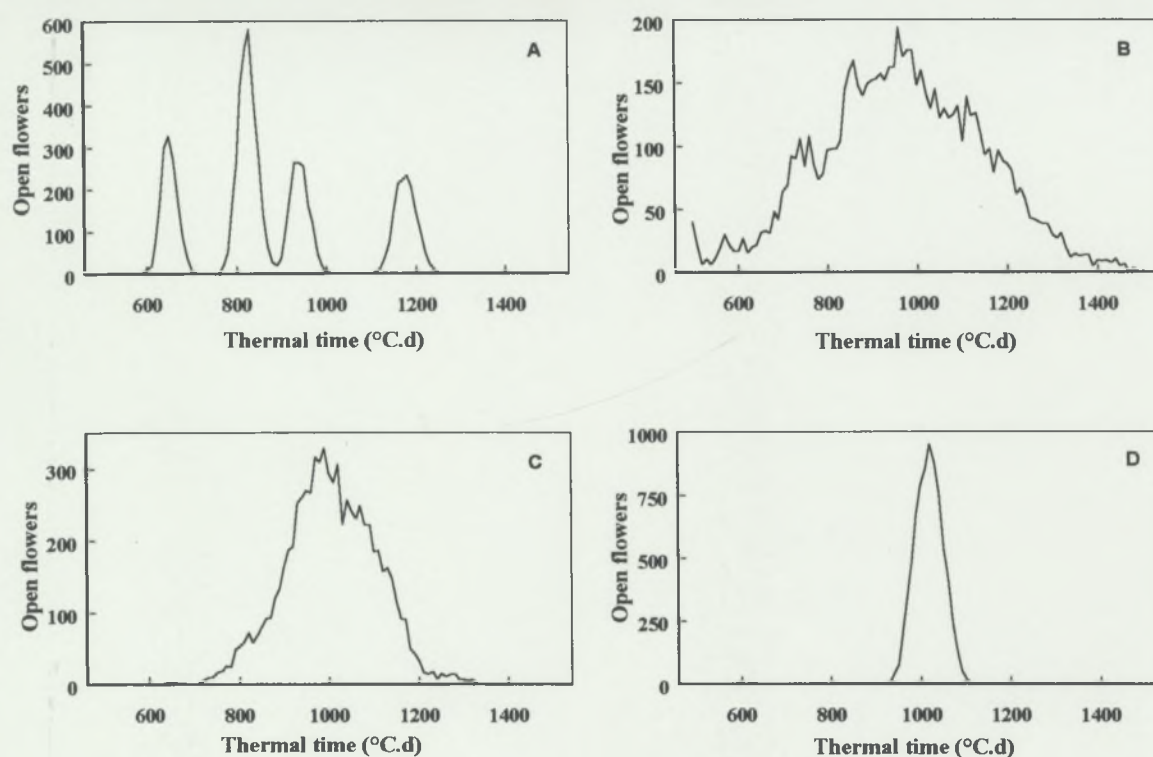


FIG. 1. Simulated flowering patterns with the "shoot model" of a set of five trees bearing each 100 shoots with 15 flower buds per shoot for different levels of synchronisation of flowering between (β_s) and within (δ) trees: A: $\beta_s = 250^\circ\text{C.d}$; $\delta = 2^\circ\text{C.d}$, B: $\beta_s = 150^\circ\text{C.d}$; $\delta = 150^\circ\text{C.d}$, C: $\beta_s = 20^\circ\text{C.d}$; $\delta = 100^\circ\text{C.d}$, D: $\beta_s = 20^\circ\text{C.d}$; $\delta = 2^\circ\text{C.d}$. The other parameters are set at $\alpha_s = 1000^\circ\text{C.d}$, $\sigma_s = 0.7^\circ\text{C.d}$.

Model fitting and model comparison

The 520 m data set is summarised in Table 1. Not all the labelled terminal branches gave flowering shoots. The number of flower buds per flowering shoot varied from 1 to 50. The total number of flower buds observed was 4151, and 2370 flower buds were taken into account for parameters estimation after a 20-shoot sampling among the 100 shoots of two trees (Table 1). At 200 and 890 m, 695 and 972 flower buds respectively were observed. The differences were related to different flushing and flowering intensities of the trees between the sites. Only eight trees flowered at 200 m.

The parameters of the three models fitted on the data at 520 m were:

- "shoot" model: $\alpha_s = 927.3^\circ\text{C.d}$; $\beta_s = 40.6^\circ\text{C.d}$; $\delta = 95.4^\circ\text{C.d}$; $\sigma_s = 0.96^\circ\text{C.d}$
- "tree" model: $\alpha_t = 923.7^\circ\text{C.d}$; $\beta_t = 50.8^\circ\text{C.d}$; $\sigma_t = 2.61^\circ\text{C.d}$
- "population" model: $a = 914.4^\circ\text{C.d}$; $\sigma = 3.12^\circ\text{C.d}$

As expected, the value of the parameter of variance relative to flowering increased with the level of organisation of the model (shoot < tree < population). It reflected the dispersion of flowering at each level. β_t and β_s were high, indicating the asynchrony of

flowering between trees (Table 1). Individual tree flowering thresholds a_p' , estimated with the "tree" model, varied from 837.7 °C.d to 998.9 °C.d. The difference represented 14 days. Likewise, δ was high, indicating the asynchrony between shoots on the trees. The largest difference between shoot flowering thresholds $a_{j,p}$ on a tree was 445 °C.d, or 38 days.

The test of maximum likelihood was highly significant for all pair comparisons of the three models (Table 2). The "shoot" model, with the higher likelihood, fitted the best the data. The large gain of likelihood between the "shoot" model and the two other models indicated a large gain of accuracy in describing the data when the shoot level was considered.

TABLE 1. *Number of flowering shoots and flower buds, dates of onset and end of flowering of the flower sample observed on 10 Psidium cattleianum trees (1 – 10) at 520 m in Réunion Island. Trees 3s and 6s correspond to a 20-shoot random sampling among the 100 tagged shoots of the trees 3 and 6 respectively.*

Tree	Flowering shoots	Flower buds	Onset of flowering (°C.d)	End of flowering (°C.d)
1	18	287	621.4	977.7
2	18	185	881.3	1153.2
3	92	1112	710.2	1203.2
4	19	228	821.3	1140.6
5	19	275	721.0	1216.2
6	94	1071	731.8	1190.3
7	20	314	710.2	1039.5
8	17	286	857.6	1255.6
9	20	301	721.0	1270.3
10	13	92	832.9	1255.6
3s	20	233	721.0	1078.3
6s	17	169	743.1	1190.3

TABLE 2. *Pair comparisons of the "shoot", "tree" and "population" flowering models by tests of maximum likelihood, each at $P = 1.7\%$, with the flowering data of 10 Psidium cattleianum trees obtained at 520 m in Réunion Island.*

Model 1	Model 2	$2 \cdot \Delta LL$	df	$\chi^2_{0.017}$	P
"shoot"	"tree"	3896.8	1	5.7	0
"shoot"	"population"	4695.0	2	8.1	0
"tree"	"population"	798.2	1	5.7	0

$2 \cdot \Delta LL$: two times the difference of log-likelihood of the models; df : difference between the number of parameters of the models; $\chi^2_{0.017}$: test's threshold (0.983 quantile of the cumulated chi-square distribution with df degrees of freedom); P : 1 – cumulated probability of $2 \cdot \Delta LL$ on a chi-square distribution with df degrees of freedom.

Simulated FFDs by the three models are compared to empirical FFD at 520 m in Fig. 2. Empirical FFD showed strong irregularities, particularly during the peak flowering, indicating that flowering was not a smooth process. These irregularities were due to the flowering at the same time of one or several shoots bearing a lot of flower buds. Models gave a good general fit of the flowering pattern. However, the irregularities were not well simulated as model FFDs are averages and consequently smoothed the curves. Fig. 1 showed that the "shoot" model could reflect such strong irregularities with a particular simulation.

Model validation

When expressed in thermal time, the duration of over-all empirical flowering on the 3 sites were almost similar (592.8°C.d at 200m; 648.9°C.d at 520m; 472.2°C.d at 890m), indicating that flowering was mostly driven by temperature, which was our basic hypothesis. But flowering patterns were different, with a straight flowering peak during the mid-bloom at 200 m (Fig. 3), and just before mid-bloom at 890 m (Fig. 4). These peaks are related to heavy flowering trees with close flowering thresholds. At 200m, four trees (on eight) bore 71.1% of the observed flower buds and had very close flowering thresholds (902.7 to 918.7°C.d). At 890 m, three trees (on ten) bore 56.8% of the observed flower buds and had also close flowering thresholds (861.6 to 933.4°C.d).

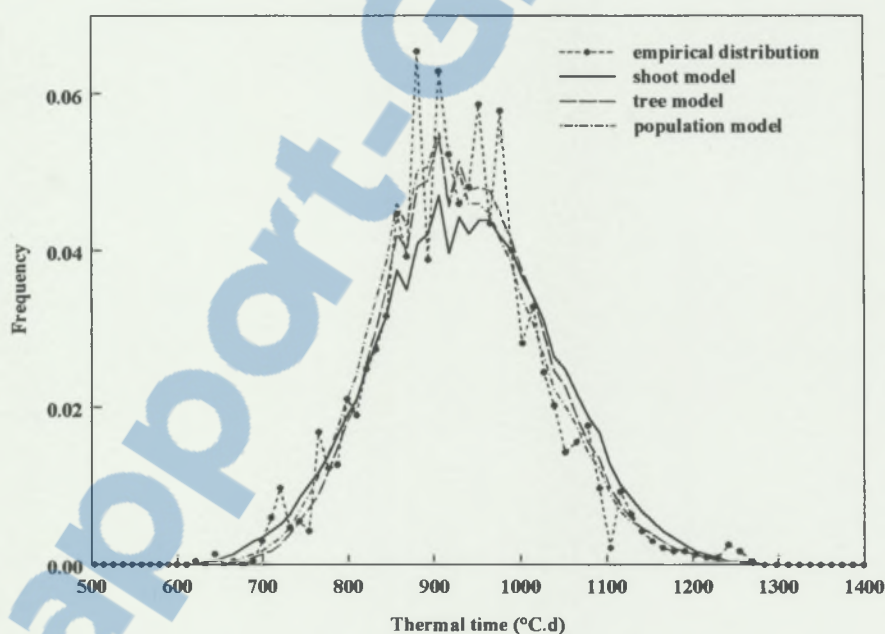


FIG. 2. Empirical and simulated flowering frequency distribution of a group of ten trees of *Psidium cattleianum* (2370 flowers) at 520m of elevation in Réunion Island. Model parameters are estimated from the data.

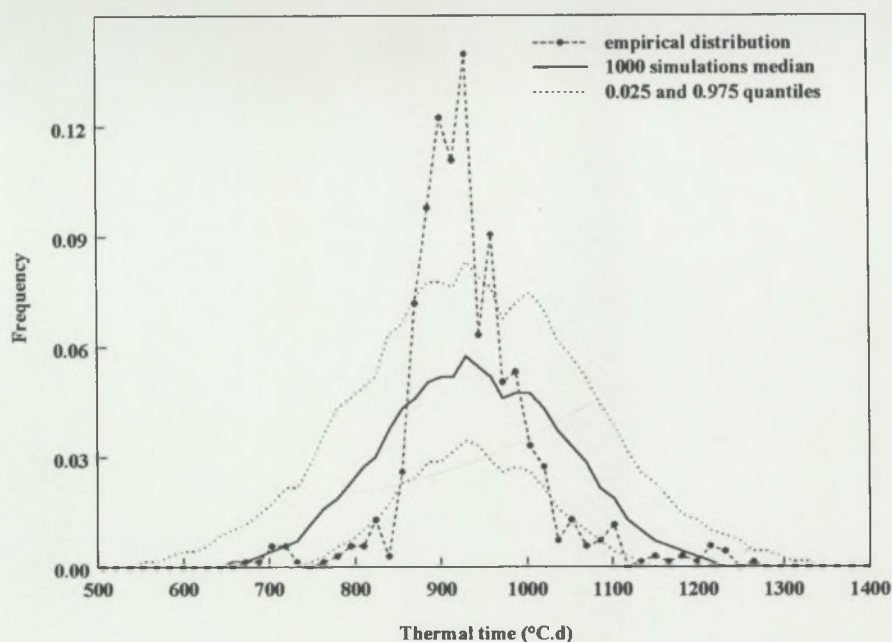


FIG. 3. Validation of the "shoot" model at 200 m of elevation: empirical flowering frequency distribution with the median and the 0.025 and 0.975 quantiles of 1000 simulations run with the "shoot" model fitted at 520 m. The 200 m thermal-time scale, and the empirical numbers of flower buds per shoot were used for the simulations.

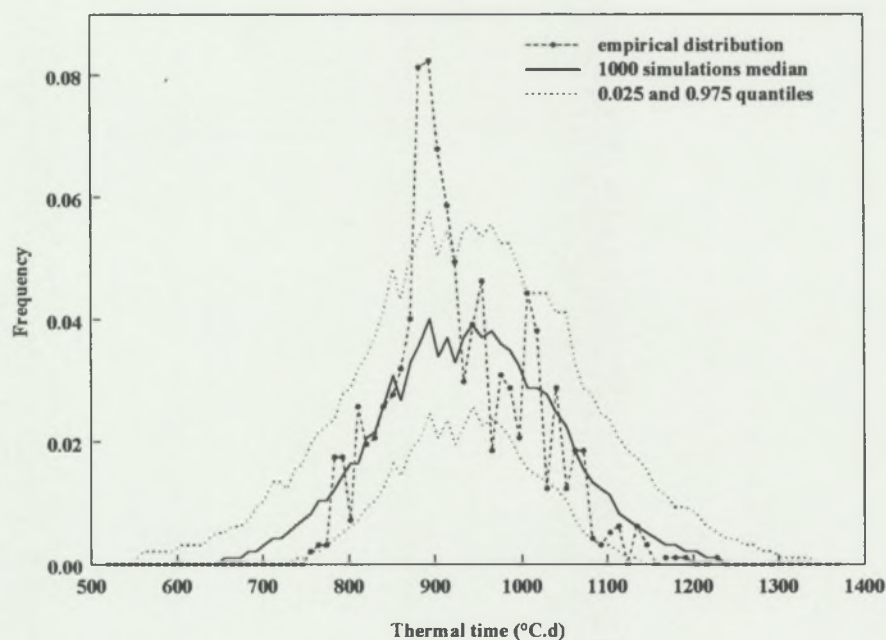


Fig. 4. Validation of the "shoot" model at 890 m of elevation: empirical flowering frequency distribution with the median and the 0.025 and 0.975 quantiles of 1000 simulations run with the "shoot" model fitted at 520 m. The 890 m thermal-time scale, and the empirical numbers of flower buds per shoot were used for the simulations.

Data at 200 and 890 m were used to estimate the validity of the "shoot" model for temperature and population variations. On each site, one thousand simulations were run. The 0.025 and 0.975 quantiles of the simulated flowering frequencies determined for each date t_i are graphically compared to the empirical data (Fig. 3 and 4). This method evaluates the behaviour of the model, but does not indicate if the model is right. Empirical FFDs were satisfactorily included in the range of 95% of the simulated values, except for the flowering peaks observed on these sites. The model gave a good representation of flowering duration and mid-bloom date on both sites. Similar median FFDs were obtained with the "population" model and the "tree" model (data not shown). But the inter-quantile band width was narrower for the simulation with the "population" model. With the "tree" model, it was as wide as with the "shoot" model during the peak flowering, but was narrower at the onset and the end of flowering. The "population" model therefore failed to represent the strong irregularities observed in empirical FFDs. The "tree" model appeared as efficient as the "shoot" model to represent the strong irregularities of empirical FFDs during the peak flowering, but was less efficient for this at the onset and the end of the flowering. This is likely the consequence of the random attribution by the "shoot" model of a flowering threshold to each flowering unit whose weight in the tree flowering is its number of flower bud. So flowering units with numerous flower buds can flower early or late, and increase the simulated flowering frequency at the onset or at the end of the global flowering.

Hypotheses testing

Three hypotheses were made to simplify the general form of the flowering model given by eqn (3) and to derive the "shoot" model [eqn (6)]. These hypotheses have been tested on the 520 m data set to verify their biological relevance for *P. cattleianum*.

Kolmogorov-Smirnov's tests indicated that the differences $d_{j,p}$ followed a central normal distribution on all the trees, but with different standard deviations δ_p (Table 3). The hypothesis of the same normal distribution of the differences $d_{j,p}$ on all the trees was tested by a test of maximum likelihood comparing a model $(a_p', \sigma_p^2, \delta)$ with the same δ for all the trees to a model $(a_p', \sigma_p^2, \delta_p)$ with the estimated δ_p on each tree. Each tree was characterised by the estimated values of its flowering threshold a_p' and its parameter of variance relative to the flowering of its shoots σ_p^2 . The test was not significant at the 5% level ($2 \cdot \Delta LL = 16.6$; $df = 9$;

$\chi^2_{0.05} = 16.9$; $P = 0.06$), indicating that this hypothesis was verified by our sample. However, the test's probability is closed to the significance threshold, suggesting further work to verify this hypothesis in *P. cattleianum*, and eventually modify this assumption for this species.

The parameter of variance relative to flowering at the shoot level $\sigma_{j,p}^2$ [eqn (3)] was hypothesised to have a unique value σ_s^2 independently of the shoot and the tree. We first tested its independence with the shoot by a test of maximum likelihood comparing a model $(a_{j,p}, \sigma_{j,p}^2)$ where each shoot was characterised by the estimated values of its flowering threshold $a_{j,p}$ and its parameter of variance relative to flowering $\sigma_{j,p}^2$, to a model $(a_{j,p}, \sigma_p^2)$ where the shoots of each tree had a common parameter of variance σ_p^2 relative to flowering.

The test was highly significant ($2 \cdot \Delta LL = 1142$; $df = 71$; $\chi^2_{0.05} = 202.5$; $P = 0$). The hypothesis of the same σ_p^2 for all the shoots of a tree was therefore rejected, and *a fortiori* the hypothesis of the same parameter of variance relative to flowering σ_s^2 for all the shoots of the tree population. The $\sigma_{j,p}^2$ values were highly variable on each tree ($\sigma_{j,p}$ coefficient of variation on trees 3 and 6: 72.2% and 74.2% respectively). The $\sigma_{j,p}^2$ traduced the variability in

development-time accumulation among the flower buds of a flowering shoot. Two sources of variation could be put forward. First, slight significant positive correlations were found between $\sigma_{j,p}$ and the number of flowers of the shoot ($r = 0.37$; $n = 89$; $P < 0.001$ for tree 3, and $r = 0.52$; $n = 93$; $P < 0.001$ for tree 6). This suggested a differential assimilate distribution among the flower buds when they are numerous on a shoot, leading to variable development rates, and a larger $\sigma_{j,p}$ on the shoot. Second, variable $\sigma_{j,p}^2$ among the shoots of a tree could be in relation with differences in the effective temperature of their flower buds, linked to the shoot position in the canopy (inside/outside, cardinal orientation).

The flowering thresholds a_p' of the trees at 520 m estimated by the "tree" model followed a normal distribution with mean α_i and variance β_i^2 (Kolmogorov-Smirnov's test: $n = 10$; $ks = 0.2129$; $P = 0.68$).

TABLE 3. Kolmogorov-Smirnov's tests for the central normal distribution of the differences $d_{j,p}$ between the tree flowering threshold and the shoots flowering thresholds of 10 *Psidium cattleianum* trees at 520 m in Réunion Island. The standard deviation (δ_p) of the central normal distributions, estimated by the model or calculated from the data, are presented. The trees 3s and 6s, and the number of flowering shoots per tree are as in Table 1.

Tree	ks	P	Estimated δ_p (°C.d)	Calculated δ_p (°C.d)
1	0.1867	0.50	64.0	66.0
2	0.1300	0.88	55.1	58.4
3	0.0962	0.36	67.8	68.6
4	0.1194	0.92	81.0	83.8
5	0.0991	0.98	91.8	95.5
6	0.0501	0.97	86.1	87.1
7	0.1230	0.89	70.3	72.4
8	0.1385	0.86	70.0	73.9
9	0.1630	0.61	110.3	114.4
10	0.1210	0.98	113.8	118.6
3s	0.0891	0.99	75.1	77.6
6s	0.1488	0.79	95.1	98.6

ks: Kolmogorov-Smirnov's test threshold; P: test's probability.

DISCUSSION

Model structure

A general stochastic model simulating the time course of flowering of an asynchronously flowering population of trees has been presented. The flowering at a particular level (tree or population) is the aggregation of the flowering of the inferior level components (flowering unit or tree respectively). The basic component is the flowering unit, whose flower buds are supposed to follow the same stochastic process of development. The flowering unit must be determined for each species. The model considers the flowering variability both between and within tree. The within tree variability is related to the time course of flowering at the flowering unit level and to the differences in flowering time between flowering units. The

between tree variability is related to different tree flowering thresholds due to genetic and environmental factors. The inputs are a development time scale and the number of flower buds per flowering units or per tree, obtained by observations or as plant growth model outputs. The aggregative structure of the model allows different levels of details in the outputs, from the flowering unit to the population time course of flowering.

The general form of the model [eqn(3)] has a large number of parameters specific to the observed data and is not relevant for prediction purposes. Three hypotheses were then proposed to simplify the model and derive the "shoot" model which requires only four parameters quantifying the components of the flowering variability. These hypotheses have biological meanings and their pertinence should be tested for each species. Moreover, the supposed independence between the differences $d_{j,p}$ and between the tree flowering thresholds implies the independence of flowering of the flowering units of each tree, and of the trees in a population. The influence of environmental and genetic factors on the flowering phenology tends to be under-estimated by the hypotheses as the individual tree has no influence on the distribution of the differences $d_{j,p}$ and on the parameter of variance relative to flowering at the shoot level. Each individual is characterised by its own flowering threshold and by its number of flower buds. Nevertheless, the simulated time course of flowering of n trees is not the superposition at different times of n flowering of the same tree. The model considers for each shoot of a tree its number of flower buds and combines it with a flowering threshold randomly sampled in an estimated distribution. Consequently, each simulation creates a specific time course of flowering for each tree, weighted by its total number of flower buds.

The simulations performed with different levels of flowering synchronisation between and within trees (Fig.1) show that the main flowering patterns are well represented by the "shoot" model. For example, Gentry (1974) found five distinct flowering patterns among the American Bignoniaceae, mainly related to different strategies with regard to pollinators (see also Bawa, 1983; Ims, 1990). They correspond to various levels of synchronisation within and between trees and can be well simulated by the model: Fig. 1B corresponds to the 'cornucopia' flowering pattern and Fig. 1D to the 'big band' flowering pattern (Gentry, 1974).

Two simpler models are derived from the "shoot" model: the "tree" model which neglects the within tree flowering asynchrony, and the "population" model which neglects moreover the between trees flowering asynchrony (e.g. Agostini *et al.*, 1999) and considers that all the flower buds of the population follow the same stochastic process of development time accumulation (i.e. no genetic or environmental influence).

Model adaptability

The model is flexible at two levels. It can be adapted to the biological characteristics of a particular species or environment, and to the objectives and needs of a particular study, with consequences on output accuracy.

The hypotheses made to derive the "shoot" model may not be suited for some species. It is then possible to use different distributions for the differences $d_{j,p}$ and the tree flowering thresholds in eqn (4) and (5), or to fix a tree flowering sequence in the case of a strong environmental (Primack, 1980) or genetic (Primack, 1980; Pors and Werner, 1989; Carthew, 1993; O'Brien and Calder, 1993; Mitchell-Olds, 1996; Hof *et al.*, 1999) influence on flowering phenology. Likewise, a particular distribution of the parameter of variance relative to flowering at the shoot level can be used. The analytic formulation of the "shoot" model

would then be more complex.

The model can be used to simulate the time course of a flowering with a predictive purpose or a descriptive purpose, and the simulation accuracy may not be the same. The most accurate way to reproduce the observed time course of flowering is to estimate the parameters at the tree and/or flowering unit level and fix their value in the simulation process, with variability only induced by the stochastic process of flower bud development. If the objective is to predict the time course of flowering, then simulations are run with the flowering thresholds randomly sampled in distributions whose parameters are estimated besides. Simulations are then less accurate, but the model is of more general use, provided that the population and its environment whose time course of flowering is simulated are included in the model domain of validity defined by the data upon which parameters were fitted. Our study on *P. cattleianum* showed that flowering duration and mid-bloom date are well estimated, even if parameters are fitted in another place. Moreover, the most likely simulated data can be estimated from a large number of simulations (Fig. 3 and 4). A limitation of the model in prediction may arise from a strong lack of balance in the number of flower buds of the components of a flowering level, as illustrated by the model validation results for *P. cattleianum* at 200 and 890 m. Heavy flowering trees may influence the population time course of flowering, particularly if their flowering thresholds are close, or if they are early or late. To estimate these flowering thresholds and fix them in the simulation process is a means to overcome this problem. The general model [eqn(3)] and its different levels of simplification can therefore respond to different objectives and needs for a study, with different levels of simulation accuracy in consequence. The cost of a higher accuracy is expressed by more precise observations, more complex parameters estimation, and simulation limited to the flowering upon which parameters are fitted.

The flowering model can be connected to classical budburst models providing the development time scale (e.g. Chuine *et al.*, 1999). The critical development threshold, or critical state of forcing, used in the budburst models corresponds to the shoot or the tree flowering thresholds of our flowering model where it then follows a specified distribution whose parameters are fitted from the data.

The model calculates the probability that a flower bud bursts between t_{i-1} and t_i . If individual flowers last one day, the model simulates directly the population of open flowers each day. But if individual flowers last n_f days, the model simulates the onset of flowering for each flower. The population of flowers open at t_i is then calculated by adding the number of flowers which have opened the $n_f - 1$ days preceding t_i (and which are still open) to the number of flowers that open between t_{i-1} and t_i . Similar calculation can be made with particular events during the flowering period, such as the pollen release or the stigma receptivity periods (e.g. O'Brien and Calder, 1993), or the flower susceptibility period to pests or pathogens.

The flowering model have been built for and tested on tree species. However, it may be applicable to any flowering species, including annual and herbaceous species. Requisites are the identification of flowering units, as defined in the model construction, and the determination of a development time scale.

Application of the model to P. cattleianum

P. cattleianum flowering shows both between and within tree asynchrony, the latter being more pronounced than the former. The "shoot" model is then the most relevant to represent

the time course of flowering, as indicated by its high likelihood, whereas the "population" model has the lowest likelihood. The integration of asynchrony in a flowering model improves therefore the accuracy of the time course of flowering fit for species concerned by this phenomenon. The flowering duration and the mid-bloom date are satisfactorily estimated.

Data show that flowering is not a smooth process at the tree or at the population level (Fig. 2, 3, 4). The irregularities are due to the flowering of shoots bearing many flower buds, and/or to the simultaneous flowering of several shoots. The aggregative structure of the "shoot" and the "tree" model allow them to simulate such irregularities (Fig. 1B and 1C), whereas the "population" model cannot as it considers the flower buds population as a whole and therefore smoothes the flowering process.

The model parameters δ and β_s estimated at 520 m include the genetic variation of the flowering phenology. The assumption of a similar normal distribution of the differences $d_{j,p}$ on all the trees is acceptable, indicating that the different flowering shoots of a tree are comparable with respect to flowering, and their flowering threshold distribution around the tree flowering threshold is not affected by the tree. The tree flowering thresholds follow a normal distribution at 520 m. This distribution also fits the distributions of the tree flowering thresholds estimated at 200 and 890 m ($n = 8$; $ks = 0.3045$; $P = 0.37$ and $n = 10$; $ks = 0.1873$; $P = 0.81$ respectively). The genetic and environmental variability in tree flowering thresholds included in the normal distribution estimated at 520 m appears therefore representative of this variability on sites with different genetic and environmental conditions.

The hypothesis of a unique parameter of variance relative to flowering at the shoot level is not verified. This parameter is highly variable between the flowering shoots of a tree, suggesting the effect of physiological or micro-environmental factors. Moreover, the distributions of this parameter values are significantly different between individuals, indicating a tree effect likely related to environmental and/or genetic factors. But our data are not relevant to test it. Further work is needed to specify the factors affecting the parameter variability within and between the trees. From a practical point of view, we retain this hypothesis for strawberry guava as the "shoot" model gives a satisfactory fit of the data.

Model applications

The model applications are wide and cover fields of research where the simulation of the time course of flowering is useful: flowering phenology (e.g. Primack, 1980; Augspurger, 1983; Carthew, 1993), interactions between floral biology and reproductive ecology (effective mating population; genes exchange (e.g. Fripp *et al.*, 1987); outcrossing rates (e.g. Murawski and Hamrick, 1991, 1992); pollinators foraging behaviour (e.g. Bawa, 1983); seed and fruit set (e.g. Carthew, 1993)); airborne pollen release distribution; integrated management of flower pest and disease; integration in a production model for crops whose pollination is a key step in seed and fruit set (Lescourret *et al.*, 1999).

The model is also a tool to quantify the within tree variability and the tree flowering thresholds at the individual level or at the population level. These parameters are useful as selection criteria (Hof *et al.*, 1999), to study their variability in a population, or to estimate the contribution of genetic and environmental factors in the flowering variability, provided a relevant experimental design.

LITERATURE CITED

- Agostini D, Habib R, Chadoeuf J. 1999. A stochastic approach for a model of flowering in kiwifruit 'Hayward'. *Journal of Horticultural Science and Biotechnology* 74: 30-38.
- Andersen TB. 1991. A model to predict the beginning of the pollen season. *Grana* 30, 269-275.
- Augspurger CK. 1983. Phenology, flowering synchrony, and fruit set of six neotropical shrubs. *Biotropica* 15, 257-267.
- Bawa KS. 1983. Patterns of flowering in tropical plants. In: Jones CE, Little RJ, eds. *Handbook of experimental pollination*. New York: Scientific and Academic Editions, 394-410.
- Belmonte J, Roure JM. 1991. Characteristics of the aeropollen dynamics at several localities in Spain. *Grana* 30, 364-372.
- Boudry P, Wieber R, Saumitou-Laprade P, Pillen K, Van Dijk H, Jung C. 1994. Identification of RLFP markers closely linked to the bolting gene B and their significance for the study of the annual habit in beets (*Beta vulgaris* L.). *Theoretical Applied Genetics* 88, 852-858.
- Boyer WD. 1973. Air temperature, heat sums, and pollen shedding phenology of longleaf pine. *Ecology* 54, 420-426.
- Cannell MGR, Smith RI. 1983. Thermal time, chill days and prediction of budburst in *Picea sitchensis*. *Journal of Applied Ecology* 20, 951-963.
- Carthew SM. 1993. Patterns of flowering and fruit production in a natural population of *Banksia spinulosa*. *Australian Journal of Botany* 41, 465-480.
- Chuine I, Cour P, Rousseau DD. 1998. Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing. *Plant, Cell and Environment* 21: 455-466.
- Chuine I, Cour P, Rousseau DD. 1999. Selecting models to predict the timing of flowering of temperate trees: implications for tree phenology modelling. *Plant, Cell and Environment* 22, 1-13.
- Citadin I, Raseira MCB, Herter FG, Baptista da Silva J. 2001. Heat requirement for blooming and leafing in peach. *HortScience* 36, 305-307.
- Dacünha-Castelle D, Duflo M. 1982. *Probabilités statistiques. I. Problèmes à temps fixe*. Paris: Masson.
- Dennis B, Kemp KP, Beckwith RC. 1986. Stochastic model of insect phenology: estimation and testing. *Environmental Entomology* 15, 540-546.
- Dodd JC, Estrada A, Jeger MJ. 1992. *Epidemiology of Colletotrichum gloeosporioides in the tropics*. In: Bailey JA, Jeger MJ, eds. *Colletotrichum Biology, pathology and control*. Wallingford: C.A.B. International.
- Frenguelli G, Bricchi E, Romano B, Mincigrucci G, Spieksma FTM. 1989. A predictive study on the beginning of the pollen season for Gramineae and *Olea europea* L. *Aerobiologia* 5, 64-70.
- Fripp YJ, Griffin AR, Moran GF. 1987. Variation in allele frequencies in the outcross pollen pool of *Eucalyptus regnans* F. Muell. throughout a flowering season. *Heredity* 59, 161-171.
- Gentry AH. 1974. Flowering phenology and diversity in tropical Bignoniaceae. *Biotropica* 6: 64-68.
- Hänninen H. 1987. Effects of temperature on dormancy release in woody plants: implications of prevailing models. *Silva Fennica* 21, 279-299.

- Hof L, Keiser LCP, Elberse IAM, Dolstra O. 1999. A model describing the flowering of single plants, and the heritability of flowering traits of *Dimorphotheca pluvialis*. *Euphytica* 110: 35–44.
- Ickovic MR, Boussioud-Corbières F, Sutra JP, Thibaudon M. 1989. Hay fever symptoms compared to atmospheric pollen counts and floral phenology within Paris suburban area in 1987 and 1988. *Aerobiologia* 5: 30–36.
- Ims RA. 1990. The ecology and evolution of reproductive synchrony. *Tree* 5: 135–140.
- Janick J, Moore JN. 1975. *Advances in fruit breeding*. West Lafayette: Purdue University Press.
- Kramer K. 1994. Selecting a model to predict the onset of growth of *Fagus sylvatica*. *Journal of Applied Ecology* 31, 172–181.
- Lehman EL. 1983. *Theory of point estimation*. New York: Wiley and Sons.
- Lescourret F, Blecher N, Habib R, Chadoeuf J, Agostini D, Pailly O, Vaissière B, Poggi I. 1999. Development of a simulation model for studying kiwi fruit orchard management. *Agricultural Systems* 59, 215–239.
- Mathsoft. 1999. *S-Plus 2000 guide to statistics, volume 1*. Seattle: Mathsoft.
- Medan D, Bartoloni N. 1998. Fecundity effects of dichogamy in an asynchronously flowering population: a genetic model. *Annals of Botany* 81: 373–383.
- Mitchell-Olds T. 1996. Genetic constraints on life-history evolution: quantitative-trait loci influencing growth and flowering in *Arabidopsis thaliana*. *Evolution* 50, 140–145.
- Mosseler A, Papadopol CS. 1989. Seasonal isolation as a reproductive barrier among sympatric *Salix* species. *Canadian Journal of Botany* 67, 2563–2570.
- Murawski DA, Hamrick JL. 1991. The effect of the density of flowering individuals on the mating systems of nine tropical tree species. *Heredity* 67, 167–174.
- Murawski DA, Hamrick JL. 1992. Mating system and phenology of *Ceiba pentandra* (Bombacaceae) in Central Panama. *Journal of Heredity* 83, 401–404.
- Murfet IC. 1977. Environmental interaction and the genetics of flowering. *Annual Review of Plant Physiology* 28, 253–278.
- Murray MB, Cannell MGR, Smith RI. 1989. Date of budburst of fifteen tree species in Britain following climatic warming. *Journal of Applied Ecology* 26, 693–700.
- Nienstaedt H. 1974. Genetic variation in some phenological characteristics of forest trees. In: Lieth H, ed. *Phenology and seasonality modelling*. New York: Springer Verlag, 389–400.
- Normand F, Habib R. 2001. Phenology of strawberry guava (*Psidium cattleianum*) in Réunion Island. *Journal of Horticultural Science and Biotechnology* 76, 541–545.
- O'Brien SP, Calder DM. 1993. Reproductive biology and floral phenologies of the sympatric species *Leptospermum myrsinoides* and *L. continentale* (Myrtaceae). *Australian Journal of Botany* 41, 527–539.
- Osawa A, Shoemaker CA, Stedinger JR. 1983. A stochastic model of balsam fir bud phenology utilizing maximum likelihood parameter estimation. *Forest Science* 29: 478–490.
- Pors B, Werner PA. 1989. Individual flowering time in a goldenrod (*Solidago canadensis*): field experiment shows genotype more important than environment. *American Journal of Botany* 76: 1681–1688.
- Primack RB. 1980. Variation in the phenology of natural populations of montane shrubs in New Zealand. *Journal of Ecology* 68: 849–862.
- Sprenst P. 1989. *Applied nonparametric statistical methods*. London: Chapman & Hall.

III- Synthèse et conclusions pour le modèle de floraison du goyavier-fraise

La phénologie de la floraison est un facteur de régulation du rendement chez les espèces dont la fructification ou le poids du fruit dépendent de la pollinisation croisée et/ou de l'intensité de la pollinisation. Le rendement dépend alors de la superposition dans le temps des floraisons du cultivar producteur et du ou des pollinisateurs. Le goyavier-fraise fait partie de ces espèces. Ses caractéristiques biologiques (voir chapitre IV) et l'hétérogénéité génétique des parcelles amènent au besoin de connaître le déroulement de la floraison de chaque plant relativement au déroulement de la floraison des autres plants de la parcelle.

Le modèle de floraison ne doit alors pas se contenter de prévoir la date de début de floraison ou de demi-floraison (50% des fleurs ouvertes), le plus souvent sur la base de sommes de températures, mais doit aussi représenter le déroulement de la floraison. Peu de modèles ont été conçus dans cet objectif. D'autre part, la plupart des modèles de floraison ont été construits pour des vergers monoclonaux, avec l'hypothèse implicite que tous les arbres ont le même déroulement de floraison, qui est, à un facteur d'échelle près, celui du verger. La floraison étant en partie sous contrôle génétique, cette hypothèse est acceptable dans un environnement homogène. Mais on peut s'attendre à ce qu'elle ne soit pas valide dans le cas d'une population comme un verger de goyaviers-fraises de semis.

La floraison individuelle de plants de goyaviers-fraises a été relevée sur des sites naturels à différentes altitudes: 8 plants à 200 m, et 10 plants à 520 et 890 m. On suppose que le sol et l'environnement proche des plants n'affectent pas le déroulement de la floraison. Les caractéristiques variables entre sites sont la température et *a priori* la composition génétique de la population. La floraison individuelle a été suivie sur des pousses florifères de ces plants. Les données enregistrées à 520 m ont été utilisées pour ajuster les modèles, et les données enregistrées à 200 et 890 m ont été utilisées pour la validation des modèles.

Dans un premier temps, nous avons testé sur le goyavier-fraise un modèle stochastique qui a été construit pour la floraison de lianes femelles d'un cultivar de kiwi. Ce modèle considère que le développement d'un bouton floral est un processus stochastique qui consiste en l'accumulation de petits incréments de temps de développement mesurés en degrés.jours. Les boutons floraux d'une parcelle forment une population homogène par rapport au processus d'accumulation de temps de développement. A tout moment t , chaque bouton floral a accumulé une quantité de temps de développement $s(t)$ qui est distribuée normalement avec une moyenne t et une variance $\sigma^2 t$. L'ouverture d'un bouton floral a lieu lorsque la quantité de temps de développement qu'il a accumulé passe un seuil de floraison a . Les paramètres du modèle sont a et σ .

Ce modèle représente bien le déroulement de la floraison à 520 m, mais pas sur les autres sites. Il représente mal la floraison individuelle de chaque plant sur les sites. L'analyse des données de floraison montrent que le déroulement individuel de la floraison varie d'un plant à l'autre, à la fois dans sa période (date de demi floraison) et dans sa durée. Ces différences traduisent un asynchronisme de floraison entre les plants, qui est probablement lié au contrôle génétique de la floraison. Dans ce cas, le poids de chaque plant, représenté par son nombre de fleurs, dans la floraison globale de la parcelle devient important à considérer. La floraison à l'échelle du plant n'apparaît pas comme un phénomène régulier, ce qui pourrait traduire un asynchronisme de floraison au sein du plant. Le déroulement de la floraison à l'échelle du site

doit être considéré comme la superposition des déroulements individuels de floraison des plants, pondérés par leur nombre de fleurs. Le modèle de floraison construit pour le kiwi ne convient donc pas, dans sa forme initiale, comme modèle de floraison du goyavier-fraise.

Un modèle général de floraison a alors été construit en prenant en compte l'asynchronisme de floraison au sein de chaque plant et entre plants. Le cœur du modèle reste le processus stochastique décrit précédemment. On définit au sein du plant des ensembles de boutons floraux, appelés "unités florifères", dont on peut considérer qu'ils ont le même processus d'accumulation de petits incréments de temps de développement, et pour lesquels on applique le modèle précédent, avec des paramètres propres à chaque unité florifère. L'asynchronisme de floraison au sein du plant est caractérisé par la variabilité des seuils de floraison des "unités florifères" d'un même plant. L'espérance des seuils de floraison des "unités florifères" d'un plant représente le seuil de floraison de ce plant, et l'asynchronisme de floraison entre plants est caractérisé par la variabilité des seuils de floraison des plants dans la population. Sous certaines hypothèses, notamment de distribution des valeurs des seuils de floraison des "unités florifères" et des seuils de floraison des plants, ce modèle est simplifié en un modèle 'rameau' à 4 paramètres, dont deux caractérisent l'asynchronisme de floraison au sein du plant et entre les plants. Deux modèles en sont ensuite dérivés: un modèle 'plant' (trois paramètres) qui néglige l'asynchronisme de floraison au sein du plant, et un modèle 'population' (équivalent au modèle "kiwi", deux paramètres) qui néglige de plus l'asynchronisme de floraison entre les plants. Chaque modèle calcule, pour un ensemble de boutons floraux au comportement similaire par rapport à l'accumulation de temps de développement, une distribution de probabilité de floraison qui est ensuite appliquée aux boutons floraux de cet ensemble. Ils prennent donc en compte le poids de chaque ensemble homogène ("unité florifère" ou plant) dans la floraison du niveau supérieur (plant ou population respectivement). Le modèle 'rameau' permet de simuler de nombreux types de floraison observés dans la nature. Une procédure d'estimation des paramètres par maximum de vraisemblance est proposée.

Ces trois modèles ont été ajustés, et leurs qualités d'ajustement comparées, avec les données de floraison obtenues à 520 m d'altitude, l'unité florifère ayant été définie chez le goyavier-fraise comme l'ensemble des boutons floraux portés par les pousses émises par un même rameau. Le modèle 'rameau' est celui qui ajuste le mieux les données à 520 m. L'asynchronisme de floraison est plus prononcé au sein des plants qu'entre les plants. Les deux hypothèses émises pour construire le modèle, sur les distributions normales des seuils de floraison des "unités florifères" au sein du plant et des seuils de floraison des plants au sein de la population, sont vérifiées chez le goyavier-fraise. Par contre l'hypothèse de constance du paramètre de variance associé à la floraison σ^2 entre les unités florifères et entre les plants n'est pas vérifiée. Les déroulements empiriques des floraisons sur les sites à 200 m et 890 m sont représentés de façon satisfaisante par le modèle 'rameau' ou le modèle 'plant'. La durée totale de la floraison et la date de demi-floraison sont bien estimées par le modèle sur chaque site. Les avantages et inconvénients de chaque modèle sont ensuite discutés, et des applications possibles sont présentées.

Dans le cadre d'un modèle global d'élaboration du rendement du goyavier-fraise, le modèle de floraison doit être capable de simuler correctement la floraison à l'échelle d'un plant et à l'échelle de la parcelle. Le modèle descriptif le plus performant est le modèle général appliqué à des données pour lesquelles les paramètres de chaque unité florifère ont été estimés. Le modèle 'rameau', qui intègre des hypothèses simplificatrices, ajuste le mieux les données et est en ce sens un très bon modèle descriptif du déroulement de la floraison des arbres sur laquelle ses paramètres ont été estimés. Par contre, sa capacité à prédire une floraison observée à

l'échelle d'un plant est plus faible que celle du modèle 'plant': sur 10 plants dont le seuil de floraison a été estimé et fixé, des tests de Kolmogorov-Smirnov entre distributions de floraison observées et simulées ont donné des résultats non significatifs (distributions statistiquement non différentes) pour 6 plants sur 10 avec le modèle 'plant', et pour aucun plant avec le modèle 'rameau'. Cette différence est liée à l'aléa introduit en prédiction dans le modèle 'rameau' par le tirage des seuils des unités florifères dans une loi normale. Chaque unité florifère ayant un poids défini par son nombre de boutons floraux, une infinité de couples (seuils, poids) peut être tirée, engendrant une grande variabilité du déroulement de la floraison. Le modèle 'plant' considère les boutons floraux du plant dans leur ensemble, et la variabilité de déroulement de floraison qu'il peut simuler pour un plant donné est plus faible que celle du modèle 'rameau'. L'asynchronisme marqué de floraison entre les unités florifères chez le goyavier-fraise, qui n'est pas explicitement pris en compte par le modèle 'plant', est traduit par la valeur du paramètre de variance associée à la floraison à l'échelle du plant, σ_p . D'un point de vue pratique, les paramètres du modèle 'plant' sont plus aisés à estimer, et les calculs de simulation sont plus rapides. Comme l'échelle d'étude pour la floraison est le plant et que travailler à l'échelle du rameau ne présente pas d'intérêt particulier, le modèle 'plant' est retenu dans le cadre du modèle global d'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise.

Chapitre IV

Etude de la fécondation et de la fructification

I- Introduction

La formation d'un fruit, et le nombre de graines qu'il contient, sont généralement dépendants de la quantité de pollen reçue par la fleur (de Reffye *et al.*, 1978; Snow, 1982; Shore et Barrett, 1984; Falque *et al.*, 1995, 1996) et de la compatibilité du génotype du pollen avec celui de la fleur (Harrison *et al.*, 1994; Dodds *et al.*, 1996; Rodrigo et Herrero, 1996; Cuevas et Polito, 1997). Ces deux facteurs sont déterminés par la pollinisation. Elle correspond au transfert passif (gravité, vent) ou actif (insectes) du pollen, de l'anthère jusqu'au stigmate de la même fleur ou d'une autre fleur. Elle est elle-même conditionnée par plusieurs facteurs pouvant être liés: facteurs climatiques (pluie, vent, température); activité et comportement des insectes pollinisateurs (Vaissière *et al.*, 1996 a,b; Raw, 2000); masse florale (nombre de fleurs ouvertes chaque jour et libérant du pollen) des différents génotypes ainsi que leur niveau de compatibilité et leur éloignement respectifs. La pollinisation est ensuite suivie de deux processus au niveau de la fleur:

- la fécondation, qui comprend plusieurs étapes, de la germination des grains de pollen sur le stigmate à la fusion des gamètes et au développement du zygote. Son résultat, la formation de graines, est sous l'influence de la quantité de pollen déposée sur le stigmate, de la compatibilité des génotypes mâle et femelle, et de la température (Jefferies et Brain, 1984).
- la fructification, qui est le maintien de la fleur sur la plante et le début de développement du jeune fruit. Elle est en général conditionnée, à l'exception des fruits parthénocarpiques, par la réussite de la fécondation et la formation d'une ou de plusieurs graines. Celles-ci libèrent des facteurs hormonaux (Baker *et al.*, 1997) qui empêcheraient l'abscission du jeune fruit et stimuleraient l'activité cellulaire du zygote et de l'ovaire.

L'étape pollinisation-fécondation-fructification du cycle phénologique correspond au passage d'une population de N fleurs à une population de n fruits ($n \leq N$) ayant chacun un nombre variable p de graines ($1 \leq p$ pour les fruits polyspermes). C'est donc une étape clé de l'élaboration du rendement, en particulier chez le goyavier-fraise dont la croissance du fruit est déterminée en partie par son nombre de graines et le nombre de fruits du plant.

Plusieurs modèles relatifs à la pollinisation-fécondation-fructification ont été proposés. Certains ne concernent qu'un processus particulier (Jefferies et Brain, 1984; Falque *et al.*, 1995, 1996) et sont trop restrictifs dans le cadre d'une étude globale de cette étape. D'autres considèrent l'ensemble des trois processus (Brain et Landsberg, 1981; DeGrandi-Hoffman *et al.*, 1989; Lescourret *et al.*, 1998). Ils sont construits à partir de données biologiques afin de représenter de façon fonctionnelle le déroulement des processus chez l'espèce étudiée, et de prendre en compte l'effet de différents facteurs. Les facteurs limitants sont en général liés à la pollinisation croisée, indispensable chez des espèces dioïques comme le kiwi (Lescourret *et*



al., 1998), ou des variétés auto-incompatibles comme chez l'amandier (DeGrandi-Hoffman *et al.*, 1989). Le matériel végétal est composé d'un clone pollinisateur et d'un clone producteur.

Une telle approche suppose une certaine connaissance de la biologie de la reproduction de l'espèce étudiée afin de préciser les processus importants à considérer dans le modèle. Dans le cas du goyavier-fraise, les questions que l'on se pose concernent principalement la fécondation, car de leurs réponses dépendra la façon d'aborder la pollinisation dans le modèle:

- Le goyavier-fraise est-il auto-compatible ou auto-incompatible?
- L'origine du pollen (auto- ou allopollen) a-t-il un effet sur le nombre de graines et la fructification?
- Existe-t-il un effet plant sur la fécondation? En effet, l'incompatibilité est principalement définie par des barrières génétiques (Dodds *et al.*, 1996), et les plants des parcelles de goyavier-fraise sont génétiquement différents.
- Quel est le mode de pollinisation dominant chez le goyavier-fraise?
- Le nombre d'ovules d'une fleur peut-il être limitant pour le nombre de graines formées?

Les éléments de réponse à ces questions sont peu nombreux et très partiels. D'une façon générale, le goyavier-fraise est considéré comme auto-compatible (Hirano et Nakasone, 1969 b; Teatonia *et al.*, 1970; Raseira et Raseira, 1996). Ces deux derniers auteurs notent toutefois que la fructification est plus élevée en pollinisation libre qu'en autopolinisation passive. L'effet du mode de pollinisation sur le nombre de graines n'est pas relevé. Seul Chezhiyan (1988) conclut à l'auto-incompatibilité de *P. cattleianum* var. *lucidum*.

Il apparaît donc indispensable, dans un premier temps, d'acquérir des données sur la biologie de la reproduction du goyavier-fraise pour répondre aux questions précédentes, dans l'objectif de modéliser ensuite l'étape de fécondation-fructification. Les études sur la biologie de la reproduction du goyavier-fraise ont été concentrées sur deux essais: un essai de pollinisation contrôlée, présenté ci-dessous, et la détermination du nombre d'ovules dans les fleurs, présenté en annexe 2.

II- Essai de pollinisation contrôlée

Les objectifs de cet essai étaient de 1) déterminer l'influence de l'auto- et de l'allo-pollen sur la fructification et le nombre de graines formées, 2) identifier un éventuel effet plant sur la fructification et la formation des graines, et 3) déterminer si un vecteur de pollen est nécessaire chez le goyavier-fraise.

1- Matériels et méthodes

Cet essai a eu lieu sur la floraison naturelle de novembre-décembre 1997 d'une parcelle de 4 ans située à 520 m d'altitude. Cinq plants, notés A, B, C, D et E, ont été repérés au hasard sur la parcelle. Sur chaque plant, 8 rameaux portant des pousses florifères, répartis sur l'ensemble de la frondaison, ont été repérés et marqués avant le début de la floraison. Quatre traitements ont été appliqués, chacun sur 2 rameaux par plant:

- allopollinisation saturante: les rameaux de ce traitement sont ensachés dans des sachets en papier de couleur blanche. Chaque jour durant la totalité de la floraison de ces rameaux, les fleurs non émasculées sont pollinisées avec du pollen venant d'un autre plant.
- autopollinisation saturante: ces rameaux sont également ensachés. Chaque jour durant la totalité de la floraison de ces rameaux, les fleurs non émasculées sont pollinisées avec du pollen venant du même plant.
- pollinisation libre: les rameaux ne sont pas ensachés et la pollinisation naturelle a lieu sous l'effet du vent et/ou des insectes.
- auto-pollinisation passive: ces rameaux restent ensachés durant toute la période de floraison. La gravité est le seul vecteur de pollen.

Les pollinisations saturantes ont lieu le matin, vers 9h, peu après l'anthèse. Une fleur est récoltée sur un plant voisin (allopollinisation) ou sur le même plant (autopollinisation), et est retournée et frottée sur une fleur du rameau suivi, de façon à ce que le stigmate de cette dernière, qui dépasse légèrement au-dessus des étamines, soit balayé par les étamines de la fleur donneuse. On suppose que cette technique permet de saturer le stigmate en grains de pollen. Quelques relevés à la loupe ont permis de s'assurer que le stigmate portait de très nombreux grains de pollen, mais aucun comptage de ces derniers n'a été réalisé.

Le déroulement de la floraison de chaque rameau est suivi par le nombre de fleurs ouvertes par jour, à l'exception des rameaux ensachés de l'auto-pollinisation passive. Le nombre de fleurs suivies par traitement et par plant a varié de 21 à 45, avec une moyenne (\pm écart-type) de 32.5 (\pm 6.3). La pluviométrie est enregistrée en continu par un pluviographe. On suppose que les sachets protègent le pollen déposé sur le stigmate d'un effet négatif de la pluie. Les sachets ont été retirés des rameaux 5 jours après la fin de leur floraison.

Le nombre de fruits formés par rameau a été noté un mois après la fin de la floraison. Le taux de fructification est défini comme le rapport entre le nombre de fruits formés et le nombre de fleurs initial. A maturité, les fruits ont été récoltés et leur nombre de graines déterminé. Sur les 256 fruits formés, 235 ont été récoltés et analysés.

L'effet du type de pollinisation et du plant sur le nombre de graines est traité par deux analyses de variance. La pollinisation passive n'a produit que quelques fruits et ce traitement a été retiré des analyses statistiques. Le plant D qui n'a produit des fruits qu'en allopollinisation a été retiré de l'analyse. La première analyse, bifactorielle, a été réalisée sur les données des 3 traitements et des 4 plants restant après ces retraits. Une seconde analyse a ensuite été réalisée sur les données de l'allopollinisation pour y rechercher un éventuel effet plant en incluant le plant D. Les données sont transformées en racine carrée dans les 2 analyses (Dagnelie, 1998b). Les dispositifs des deux analyses étant déséquilibrés, les sommes des carrés de type III sont utilisées. Les moyennes ajustées sont comparées par des tests de Tukey de comparaisons multiples.

L'effet du type de pollinisation et du plant sur le taux de fructification est traité par une analyse de variance bifactorielle équilibrée à 2 répétitions (2 rameaux ayant reçu le même traitement par plant). Une transformation angulaire ($\arcsin \sqrt{x}$) est appliquée aux données avant analyse (Dagnelie, 1998b). Les moyennes sont comparées par des tests de Tukey de comparaisons multiples.

2- Résultats et discussion

Le type de pollinisation et le plant ont un effet très significatif sur le taux de fructification et sur le nombre moyen de graines formées (Tableaux 2 et 3). Ces variables sont plus élevées en allopollinisation saturante qu'en autopolinisation saturante ou en pollinisation libre. Ces deux derniers traitements donnent des résultats proches, avec en moyenne un nombre de graines par fruit un peu plus élevé en pollinisation libre. Le plant a également un effet significatif sur les variables mesurées, avec d'une part les plants A, B et C, qui donnent des résultats semblables, et d'autre part les plants D et E. Le plant D n'a fructifié qu'en allopollinisation saturante, et il ne donne alors pas de résultats significativement différents de ceux des autres plants (Tableaux 2 et 3; $P=0.29$ et $P=0.26$ respectivement pour le taux de fructification et le nombre moyen de graines par fruit). Le plant E présente des taux de fructification et des nombres moyens de graines toujours plus élevés que ceux des autres plants, quel que soit le traitement.

Type de pollinisation ^z	Plants					Moyenne/pol.
	A	B	C	D	E	
Allo-pol. saturante	0.63	0.86	0.96	0.72	0.81	0.80 ^a
Auto-pol. saturante	0.34	0.34	0.43	0.00	0.78	0.38 ^b
Pol. libre	0.22	0.38	0.61	0.00	0.91	0.42 ^b
Moyenne/plant	0.40 ^{bc}	0.53 ^{ab}	0.67 ^{ab}	0.24 ^c	0.83 ^a	0.53

^z: allo-pollinisation saturation; auto-pollinisation saturante; pollinisation libre

Tableau 2: Effet du type de pollinisation et du plant sur le taux de fructification du goyavier-fraise (valeurs non transformées). Les moyennes par plant ou par type de pollinisation suivies de lettres différentes sont significativement différentes (test de Tukey de comparaisons multiples de moyennes). Effet plant: $P<0.001$; effet type de pollinisation: $P<0.001$; interaction: $P=0.07$.

Type de pollinisation ^z	Plants					Moyenne ^x /pol.
	A	B	C	D	E	
Allo-pol. saturante	21.4	22.6	23.2	(17.0)	22.3	22.4 ^a
Auto-pol. saturante	9.2	7.5	9.1	-	16.0	10.4 ^b
Pol. libre	10.9	8.8	6.9	-	27.3	13.5 ^b
Moyenne ^x /plant	13.8 ^b	13.0 ^b	13.1 ^b	-	21.9 ^a	15.4

^z: allo-pollinisation saturation; auto-pollinisation saturante; pollinisation libre

^x: moyennes partielles et moyenne générale ajustées

Tableau 3: Effet du type de pollinisation et du plant sur le nombre moyen de graines par fruit du goyavier-fraise (valeurs non transformées). Le plant D n'est pas pris en compte dans cette analyse bifactorielle. Les moyennes par plant ou par type de pollinisation suivies de lettres différentes sont significativement différentes (test de Tukey de comparaisons multiples de moyennes). Effet plant: $P<0.001$; effet type de pollinisation: $P<0.001$; interaction: $P<0.001$.

En allopollinisation seule, il n'apparaît pas de différence significative entre plants, y compris le plant D ($P=0.26$).

Les distributions du nombre de graines par fruit n'ont pas une forme claire et sont variables entre les plants (Fig. 6, 9 et 10). Quel que soit le type de pollinisation, ces distributions sont en général très étalées et ne présentent pas vraiment de pic. Cela peut être dû en partie aux effectifs réduits de fruits par plant.

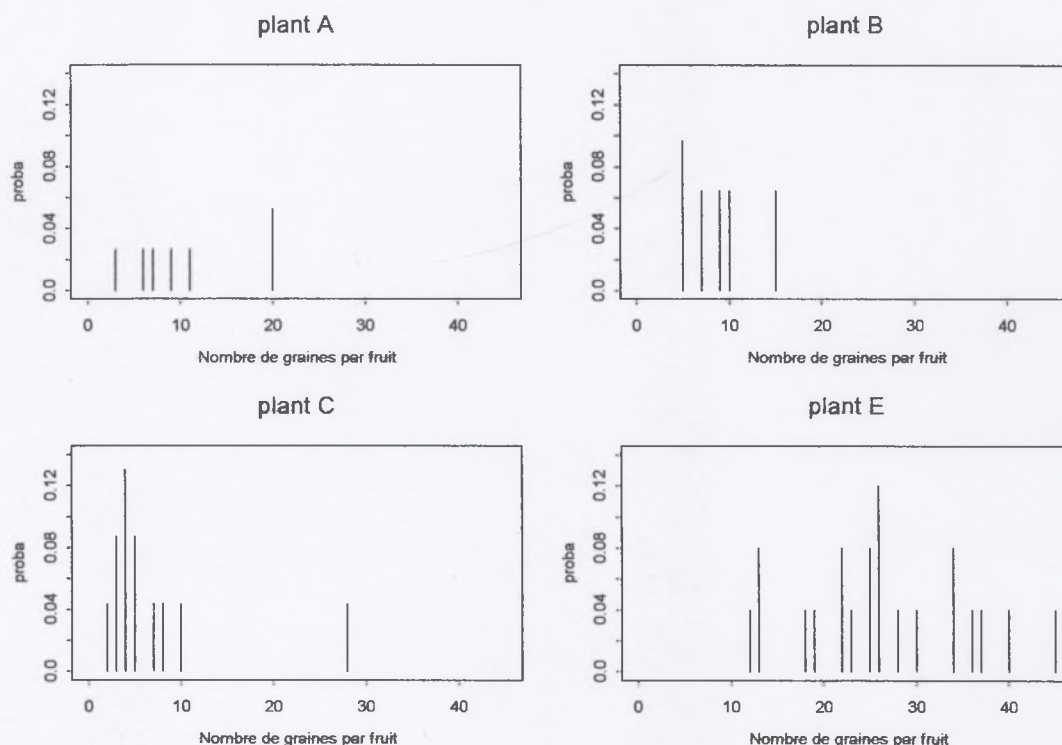


Figure 6: Distributions observées du nombre de graines par fruit après une pollinisation libre sur 4 plants de goyavier-fraise.

Le nombre de graines par fruit a varié de 2 à 46. Comme dans les milliers d'autres fruits dont les graines ont été comptées au cours de diverses études, il n'a jamais été trouvé de fruit sans graine, et ceux contenant une seule graine ont été très rares. Cela suggère que le goyavier-fraise ne présente pas de parthénocarpie spontanée et que la formation des graines est nécessaire à la fructification.

En autopolinisation saturante, une bonne relation linéaire positive est mise en évidence entre le taux de fructification (FRU) à l'échelle du rameau et le nombre moyen de graines (gr) dans les fruits du rameau (Fig. 7A): $FRU = 0.0474 \cdot gr$ ($n=8$; $r^2=0.95$; $P<0.001$). Ce genre de relation n'apparaît pas en allo-pollinisation ou en pollinisation libre (Fig. 7B et 7C).

Ces résultats montrent qu'à l'exception du plant D, les plants suivis sont auto-compatibles, mais le taux de fructification et le nombre moyen de graines par fruit sont plus élevés en allopollinisation saturante, et dans une moindre mesure en pollinisation libre. Nos résultats rejoignent les conclusions de Raseira et Raseira (1996). Un tel comportement est également noté chez d'autres Myrtaceae (Sedgley et Smith, 1989; Proença et Gibbs, 1994), ainsi que chez des espèces horticoles, comme les airelles à corymbes, *Vaccinium corymbosum* L. et *V. corymbosum* L. x *V. angustifolium* Ait. (Harrison *et al.*, 1994; Ehlenfeldt, 2001), l'olivier

(Cuevas et Polito, 1997) ou le pommier (Schneider *et al.*, 2001). Cela suggère une auto-incompatibilité partielle qui est un phénomène assez fréquent chez les Myrtoideae et en particulier la goyave (Nic Lughadha et Proença, 1996). Il apparaît très nettement un effet du plant dans la capacité à fructifier et à former des graines, ce qui suggère une variabilité du niveau d'auto-incompatibilité entre les plants de goyavier-fraise issus de semis. L'architecture des fleurs ne détermine pas *a priori* de barrière physique pour l'auto-incompatibilité. Ces résultats font penser à une auto-incompatibilité génétique de type gamétophytique (Dodds *et al.*, 1996). Les mécanismes de cette auto-incompatibilité partielle restent à être explorés et précisés. L'auto-incompatibilité partielle et variable est connue en horticulture, mais entre cultivars, comme chez le pommier, l'olivier ou l'abricotier. Dans notre étude, et dans la limite du nombre de fleurs suivies, le plant D serait auto-incompatible, et le plant E fortement auto-compatible. Les trois autres plants se situeraient entre ces deux situations.

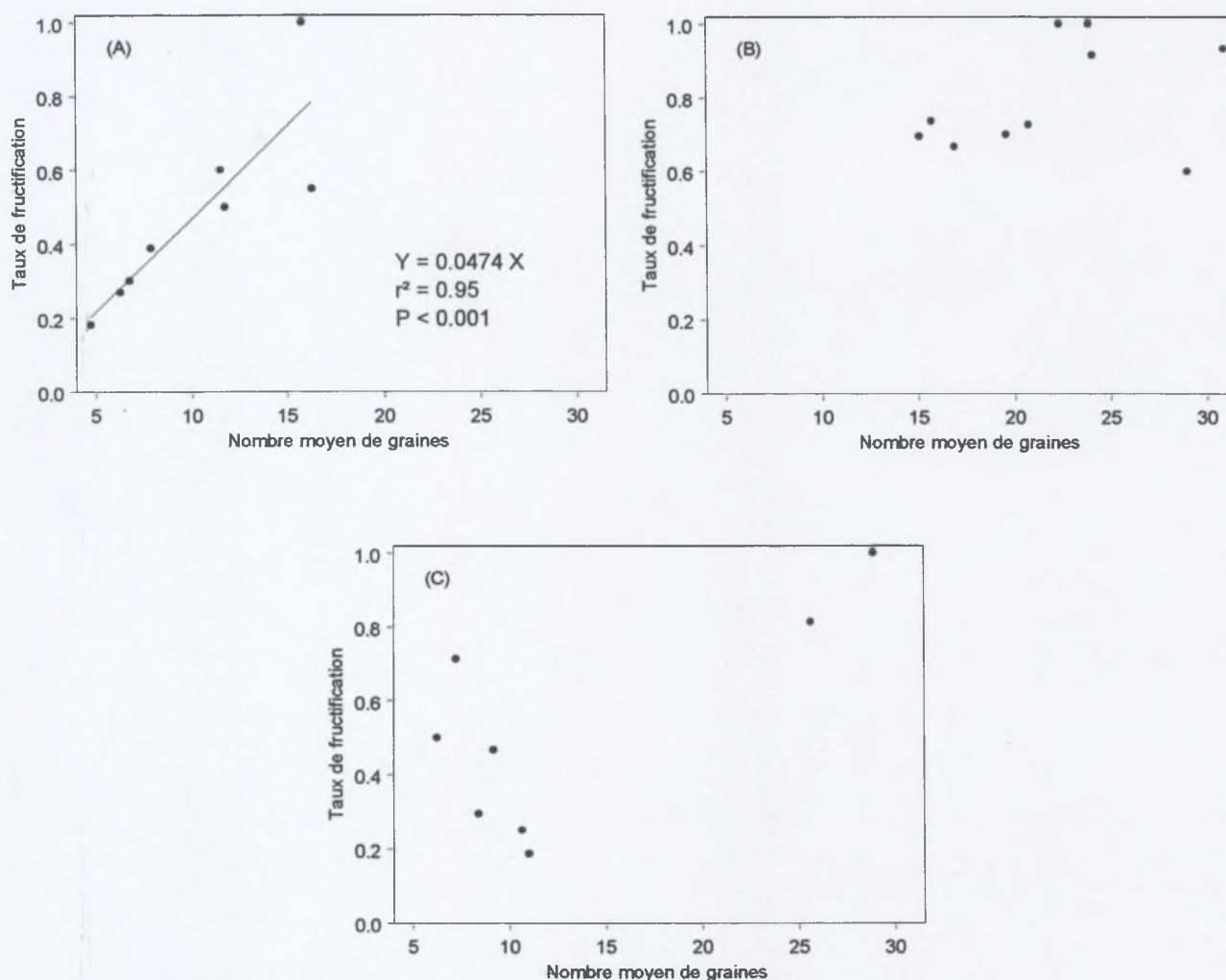


Figure 7: Relation entre le nombre moyen de graines par fruit sur un rameau de goyavier-fraise et le taux de fructification sur ce rameau, A) en autopolinisation saturante, B) en autopolinisation saturante, C) en pollinisation libre.

La fructification en autopolinisation passive a eu lieu sur les plants A, B et C, et a été très faible (3.7% globalement). Les fruits contenaient peu de graines, en moyenne 5.4 ([3; 11]), contre 8.6 pour l'autopolinisation saturante sur ces 3 plants. Ce résultat suggère que la gravité n'est pas un vecteur de pollen efficace, et qu'un transport par le vent ou les insectes est nécessaire. Les abeilles, qui sont nombreuses durant la floraison et récoltent le pollen, ont probablement un rôle important dans la pollinisation du goyavier-fraise, comme en général chez les Myrtoideae (Nic Lughadha et Proença, 1996). Les taux de fructification relevés par Raseira et Raseira (1996) pour ce même traitement sur différentes populations des deux sous-espèces de *P. cattleianum* variaient de 3.9 à 51% et sont en général plus élevés que ceux de notre étude.

En conclusion, cette étude a montré que:

- le goyavier-fraise est une espèce à auto-compatibilité partielle et variable d'un plant à l'autre.
- l'allopollinisation donne des taux de fructification et des nombres moyens de graines par fruit plus élevés que l'autopolinisation, et non significativement différents entre plants.
- il n'apparaît pas de parthénocarpie spontanée chez le goyavier-fraise.
- les distributions de nombres de graines par fruit à l'échelle du plant sont très étalées et n'ont pas une forme bien définie.
- en autopolinisation le taux de fructification est proportionnel au nombre moyen de graines dans les fruits
- un vecteur de pollen (vent, insecte) est nécessaire à la pollinisation.

3- Conséquences pour la modélisation de la pollinisation et de la fécondation-fructification

Les résultats précédents ont apporté des éléments de connaissance sur la biologie de la reproduction du goyavier-fraise. Mais ils font surtout apparaître deux difficultés majeures pour la modélisation des étapes de pollinisation et de fécondation-fructification dans notre démarche de modèle global d'élaboration du nombre de fruits et de leur calibre. Ces difficultés sont liées aux caractéristiques biologiques de la reproduction du goyavier-fraise et à la diversité génétique existant sur les parcelles.

La première difficulté réside dans la variabilité du taux d'auto-incompatibilité qui doit être prise en compte explicitement, pour ses effets sur la fructification et la formation des graines. Cela implique de connaître la distribution, actuellement inconnue, de ce taux au sein d'une population. D'autre part, pour chaque fleur de chaque plant, l'autopollen et l'allopollen doivent être considérés. La pollinisation libre est *a priori* un mélange de ces deux sources de pollen, variable en proportion et en quantité. Nous n'avons pas d'élément sur l'effet de ces sources de variabilité sur la fructification et sur la formation des graines, effet qui dépend probablement du plant. Il a été montré par exemple chez l'airelle à corymbe, que la réponse à une pollinisation avec des mélanges d'auto et d'allopollen dans des proportions variables dépendait du cultivar (Harrison *et al.*, 1994). Dans les travaux de modélisation précédents, il n'est fait référence qu'à du pollen compatible (de Reffye *et al.*, 1978; Brain et Landsberg, 1981; DeGrandi-Hoffman *et al.*, 1989, Lescourret *et al.*, 1998). Cette compatibilité n'apparaît pas explicitement dans les modèles et elle est supposée constante, ce qui est raisonnable dans le cas d'un seul couple de cultivars (producteur, pollinisateur) comme chez le kiwi (Lescourret *et*

al., 1998). Par contre, cette constance est plus discutable dans le cas où plusieurs clones pollinisateurs sont envisagés (DeGrandi-Hoffman *et al.*, 1989) du fait de l'origine génétique de cette compatibilité.

La seconde difficulté est la distribution du nombre de graines par fruit, dont la forme n'est pas claire, et est variable entre plants. Connaître cette forme peut aider à formuler des hypothèses sur les processus impliqués (de Reffye *et al.*, 1978), ou à identifier des pistes pour la modélisation.

Dans le schéma global du modèle d'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise (Fig. 5), la distribution du nombre de graines par fruit à l'échelle du plant est une entrée importante du modèle de croissance du fruit, et sa prédiction doit être aussi précise que possible. Cependant, la modélisation de la pollinisation et de la fécondation-fructification, dont la sortie est la distribution recherchée, ne sera possible que lorsque suffisamment d'éléments de réponse auront été apportés aux problèmes soulevés par nos résultats. Un tel travail demande des études spécifiques et ne peut être envisagé dans le cadre de cette thèse. Nous proposons donc d'exploiter les données dont nous disposons afin de formuler des hypothèses biologiques qui pourront être reprises ultérieurement. L'objectif est de rendre compte de l'effet de différents facteurs sur la distribution du nombre de graines par fruit. Le peu de connaissances sur la biologie de la reproduction du goyavier-fraise nous conduit à adopter des approches descriptives statistiques pour exploiter nos données, et la modélisation apparaît comme un outil pertinent d'aide à l'analyse.

III- Effet de différents facteurs sur la distribution du nombre de graines par fruit à l'échelle du plant

D'après ce qui précède, les facteurs à considérer sont le plant, l'origine génétique du pollen (auto ou allopollen), et la quantité de pollen déposé sur les stigmates, résultat de la pollinisation. La démarche adoptée consiste à construire un modèle général pour représenter la distribution du nombre de graines par fruit à l'échelle du plant, et de tester son comportement en fonction du type de pollinisation et du plant. Les paramètres du modèle sont ensuite explicités en fonction des facteurs étudiés, et des simplifications à partir d'hypothèses biologiques sont recherchées.

1- Le modèle général de distribution du nombre de graines par fruit

Pour une espèce non parthénocarpique et non apomictique, la distribution du nombre de graines par fruit après la fructification est la résultante de deux processus: la fécondation, qui détermine la distribution du nombre d'ovules fécondés par fleur, puis la fructification qui détermine la chute des fleurs fécondées. D'autres chutes peuvent intervenir au cours de la croissance du fruit, liées à des aspects nutritionnels (e.g. Ruiz *et al.*, 2001) ou à des maladies. Ces chutes sont très rares chez le goyavier-fraise et on considère que la distribution du nombre de graines par fruit à la récolte représente bien la distribution du nombre d'ovules fécondés par fleur.

1.1- La distribution du nombre d'ovules fécondés par fleur

Lescourret *et al.* (1998) proposent un modèle de fécondation à l'échelle de la fleur basé sur la description probabiliste de trois étapes: 1) la sélection des grains de pollen fertiles parmi ceux déposés sur le stigmate, c'est à dire ceux capables de produire un tube pollinique atteignant les ovules, 2) la fertilisation des ovules par les tubes polliniques, et 3) la sélection des ovules fertiles. Les hypothèses du modèle sont:

- le dépôt des grains de pollen sur un stigmate suit une loi de Poisson de paramètre μ
- la sélection des graines de pollen fertiles parmi ceux déposés sur le stigmate (étape 1) suit une loi binomiale de paramètre f_p (fertilité pollinique)
- la probabilité qu'un tube pollinique fertile n'atteigne pas un ovule donné parmi les N de l'ovaire est $1 - \frac{1}{N}$ (étape 2), en supposant que les tubes polliniques peuvent atteindre de façon semblable des ovules déjà atteints ou non
- la sélection des ovules fertiles (étape 3) suit une loi binomiale de paramètre F (fertilité ovulaire)
- les ovules d'une fleur sont indépendants les uns par rapport aux autres vis à vis de la fécondation.

Alors la combinaison de ces étapes amène à exprimer la probabilité d'avoir n ovules fécondés parmi les N que comporte la fleur comme une loi binomiale de paramètres N et

$$P = F \cdot \left(1 - e^{-\frac{\mu \cdot f_p}{N}}\right) \quad (\text{Lescourret } et al., 1998)$$

Ces trois étapes sont représentatives du déroulement général de la fécondation, et les hypothèses formulées par les auteurs sont assez générales pour être raisonnablement acceptées dans le cas du goyavier-fraise. Le manque de données sur la reproduction de cette espèce nous amène à considérer que sa fécondation suit ce schéma, et à tester le modèle précédent sur le goyavier-fraise. Cependant, la loi binomiale n'est pas d'utilisation aisée analytiquement. On propose de l'approcher par une loi de Poisson (Dagnelie, 1998a) de paramètre $\theta = F \cdot \lambda$ où F est la fertilité ovulaire et λ est le nombre de tubes polliniques fécondants arrivant au niveau des ovules. En effet, considérons que le nombre d'ovules par fleur, N , est grand par rapport au nombre moyen de tubes polliniques fécondants arrivant au niveau des ovules ($\mu \cdot f$). Cette hypothèse paraît raisonnable vu les distributions respectives des nombres d'ovules par fleur et des nombres de graines par fruit (Annexe 2). Alors $\frac{\mu \cdot f}{N} \rightarrow 0$, et $P \rightarrow 0$. Le paramètre de la

loi de Poisson vers laquelle tend la binomiale vaut $\theta = N \cdot P = N \cdot F \cdot \left(1 - e^{-\frac{\mu \cdot f}{N}}\right)$, qui peut être approché par $\theta \approx N \cdot F \cdot \frac{\mu \cdot f}{N} = F \cdot \mu \cdot f = F \cdot \lambda$. Le paramètre de la loi de Poisson est

indépendant du nombre d'ovules de la fleur, mais dépend du nombre de tubes polliniques fécondants et de la fertilité ovulaire. La probabilité que n ($n > 0$) ovules soient fécondés dans

une fleur est: $p(n) = \frac{e^{-\theta} \cdot \theta^n}{n!} = \frac{e^{-\lambda F} \cdot (\lambda F)^n}{n!}$. Les distributions attendues sont unimodales.

Afin de préciser la forme des distributions empiriques de nombre de graines par fruit avec des échantillons conséquents, nous avons utilisé les données relevées sur 11 plants sur lesquels les graines de 92 à 185 fruits issus de pollinisation libre ont été comptées (voir chapitre V sur la croissance du fruit, essai JBG). Les distributions observées ne présentent en général pas une

forme typique de loi de Poisson (Fig. 8). Leur ajustement par maximum de vraisemblance (Dacūnha-Castelle et Duflo, 1982) à des lois de Poisson, en intégrant différents modèles de fructification (voir ci-dessous § 1.2), ne sont en général pas bons. L'analyse détaillée des écarts entre distributions estimées et observées, et la forme de certaines distributions (Fig. 8, plants 14, 15, 18 et 31) amènent à penser que les distributions sont plutôt bimodales, avec des pics plus ou moins distincts en fonction de leur écart. Ce genre de distribution bimodale a également été observé chez le cacaoyer, et son interprétation est basée sur la distribution des fréquences de grains de pollen sur les stigmates, ajustée par une loi de Pareto, et le fait que soit le nombre de grains de pollen déposé (pic des petits nombres de graines), soit le nombre d'ovules de la fleur (pic des grands nombres de graines) est limitant pour la formation des graines (de Reffye *et al.*, 1978). Dans notre cas, le nombre d'ovules de la fleur n'est pas considéré comme limitant (Annexe 2). De plus, des distributions bimodales sont observées en allopollinisation saturante (exemple Fig. 10 plant A), alors qu'il n'y a probablement pas de limitation du nombre de grains de pollen. Le manque d'interprétation *a priori* de la forme des distributions observées justifie de construire un modèle descriptif statistique.

Le modèle est alors reformulé comme la somme de deux lois de Poisson de paramètres θ_1 et θ_2 ($\theta_1 > 0$ et $\theta_2 > 0$), pondérées par un coefficient de partage p . La probabilité qu'une fleur ait n ($n > 0$) ovules fécondés s'écrit:

$$[1] \quad X(n) = p \cdot \frac{e^{-\theta_1} \cdot \theta_1^n}{n!} + (1-p) \cdot \frac{e^{-\theta_2} \cdot \theta_2^n}{n!}$$

Les ajustements obtenus sur les 11 plants précédents en pollinisation libre sont présentés sur la figure 8.

1.2- La fructification

On considère en général que la probabilité de fructification d'une fleur chez les espèces à fruits polyspermes augmente avec le nombre d'ovules fécondés qu'elle contient (de Reffye *et al.*, 1978; Brain et Landsberg, 1981; Lescourret *et al.*, 1998), ou, ce qui est proche, avec le nombre de grains de pollen compatible reçus (Falque *et al.*, 1995). Cela est basé sur l'hypothèse que les hormones produites par les graines déterminent le potentiel du fruit pour attirer les assimilats nécessaires à sa croissance (Luckwill *et al.*, 1969, cités par Brain et Landsberg, 1981). La rareté des fruits à 1 ou 2 graines dans nos analyses va dans le sens d'une telle hypothèse chez le goyavier-fraise. La probabilité de fructification a alors été exprimée comme une fonction exponentielle décroissante. La probabilité qu'une fleur ayant n ($n > 0$) ovules fécondés donne un fruit (à n graines) est:

$$[2] \quad C(n) = q \cdot (1 - e^{-\tau^n}) \quad n > 0; \quad 0 \leq q \leq 1 \text{ et } \tau > 0$$

où q est le taux de fructification maximal quand $n \rightarrow \infty$, et τ ($\tau > 0$) est un paramètre de forme.

Les estimations des paramètres q et τ sur les données précédentes des 11 plants montrent que 1) le modèle global, combinaison de [1] et [2], est surparamétré, et 2) les valeurs estimées de τ sont élevées ($\tau \approx 8$), ce qui conduit à négliger le terme $e^{-\tau^n}$ quel que soit n .

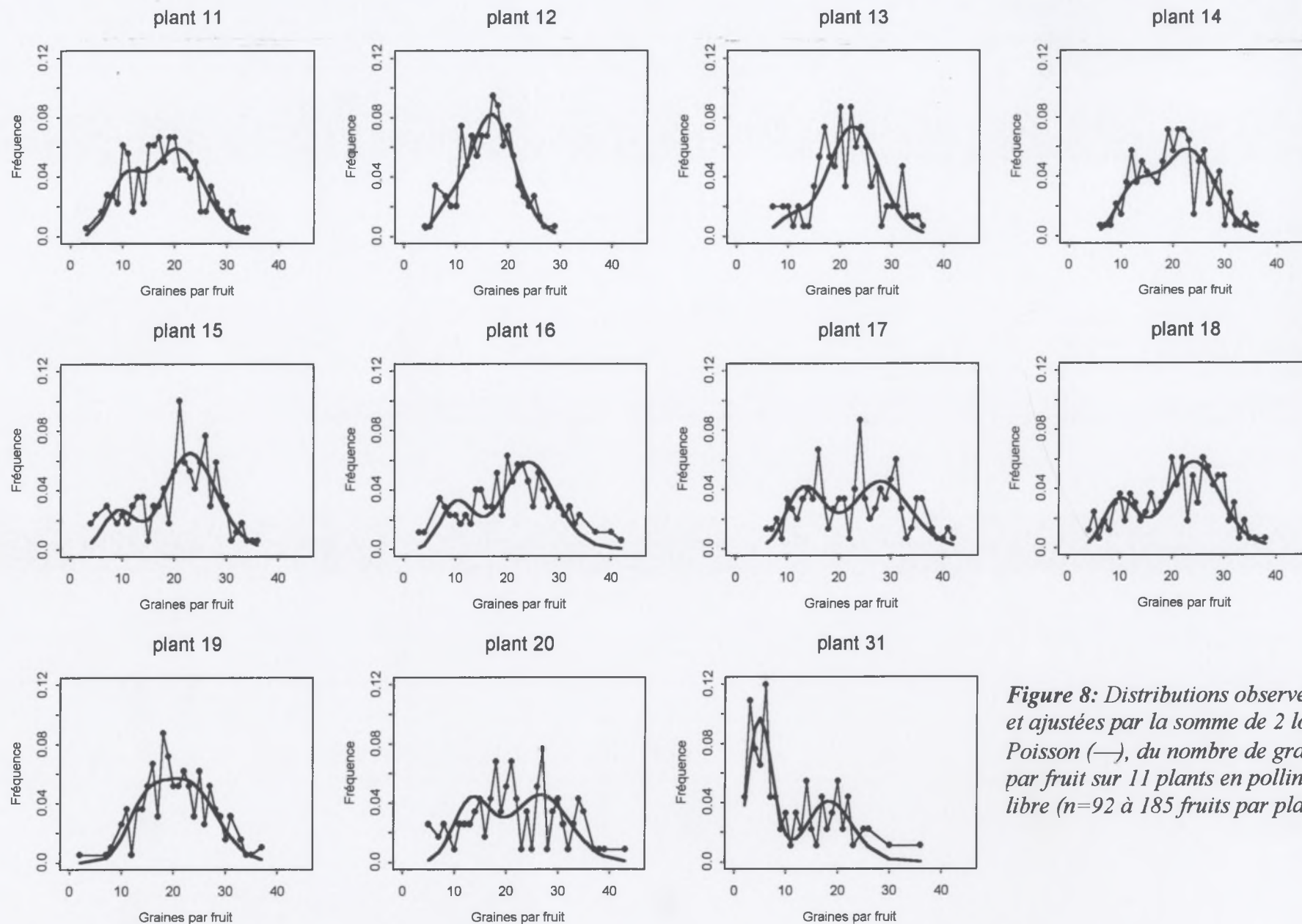


Figure 8: Distributions observées (●) et ajustées par la somme de 2 lois de Poisson (—), du nombre de graines par fruit sur 11 plants en pollinisation libre ($n=92$ à 185 fruits par plant)

La probabilité qu'une fleur avec au moins un ovule fécondé donne un fruit est alors considérée constante, q , et indépendante du nombre d'ovules fécondés.

1.3- Le modèle général de distribution des nombres de graines par fruit

L'événement {une fleur donne un fruit à n graines} est l'intersection des événements {une fleur a n ovules fécondés} et {cette fleur donne un fruit}. D'après ce qui précède, ces événements sont indépendants et la probabilité qu'une fleur donne un fruit à n ($n > 0$) graines est:

$$[3] \quad P(n) = \left(p \cdot \frac{e^{-\theta_1} \cdot \theta_1^n}{n!} + (1-p) \cdot \frac{e^{-\theta_2} \cdot \theta_2^n}{n!} \right) \cdot q$$

La non-fructification, ou chute, d'une fleur correspond aux événements suivants:

- {pas d'ovules fécondés dans la fleur} (pas de fruit parthénocarpique)
- { $n, n \in [1; +\infty)$, ovules fécondés dans la fleur et absence de fructification}

Sa probabilité, notée $P(0)$, s'écrit après développement:

$$[4] \quad P(0) = 1 - q \cdot \left(1 - p \cdot e^{-\theta_1} - (1-p) \cdot e^{-\theta_2} \right)$$

On vérifie que $P(0) + \sum_{n=1}^{n=+\infty} P(n) = 1$

Ce modèle général donne la loi de probabilité d'observer n graines dans un fruit à partir d'une population de fleurs. Cette loi est définie par les équations [3] et [4]. La probabilité $1-P(0)$ correspond à la probabilité de fructification de la population de fleurs, et la probabilité qu'un fruit contienne n graines est $P(n) / [1 - P(0)]$.

2- Effet de l'origine génétique du pollen sur les paramètres du modèle général, et simplification du modèle.

On fait l'hypothèse que l'origine génétique du pollen a une influence sur la loi de probabilité précédente. Les deux facteurs caractérisant l'origine génétique du pollen sont le plant et le type de pollinisation: la notion d'auto et d'allopollen est propre à chaque plant. Leurs effets sur les paramètres du modèle général ont été étudiés.

2.1- Matériels et méthodes

Les données sont celles des 3 traitements allopollinisation, autopollinisation et pollinisation libre de l'essai de pollinisation contrôlée. Les effectifs par plants et type de pollinisation sont relativement limités, en particulier pour l'autopollinisation et la pollinisation libre. Le plant D, dont les fleurs suivies n'ont pas donné de fruit en autopollinisation et en pollinisation libre, n'apparaît qu'en allopollinisation.

Des fruits ont été perdus avant le comptage des graines (champignon en cours de grossissement ou chute spontanée à maturité avant la collecte de l'échantillon), et le modèle a été modifié afin d'englober dans une même estimation des paramètres les fruits observés, les fruits chutés et les fleurs n'ayant pas fructifié. Soit:

a : le nombre de fleurs ayant fructifié

b : le nombre de fleurs n'ayant pas fructifié ($a + b$ = nombre de fleurs suivies)

B : le nombre de fruits perdus après fructification

$K = a - B$: le nombre de fruits observés

En notant Q la probabilité, supposée constante, de perdre un fruit ayant noué, alors la probabilité qu'une fleur donne un fruit observé à n graines est $P(n) \cdot (1 - Q)$, et la probabilité qu'une fleur donne un fruit et que le fruit chute est $(1 - P(0)) \cdot Q$. Le complémentaire est qu'une fleur ne donne pas de fruit, soit $P(0)$. On vérifie que:

$$\sum_{n=1}^{n=+\infty} P(n) \cdot (1 - Q) + (1 - P(0)) \cdot Q + P(0) = 1$$

La vraisemblance de l'échantillon de K fruits ayant chacun n_i graines ($i \in [1; K]$) s'écrit, au coefficient de la multinomiale près (que l'on ne prend pas en compte car il est indépendant des paramètres et n'intervient pas dans leur estimation):

$$V = \prod_{i=1}^{i=K} [P(n_i) \cdot (1 - Q)] \cdot [(1 - P(0)) \cdot Q]^B \cdot P(0)^b$$

Dans un premier temps, les paramètres θ_1 , θ_2 , p et q du modèle général, que l'on appellera alors modèle général complet, sont estimés simultanément pour chaque plant et chaque type de pollinisation. On obtient alors le modèle le plus paramétré possible, qui donne le meilleur ajustement aux données. On teste ensuite des hypothèses simplificatrices visant à réduire le nombre de paramètres du modèle.

L'estimation des paramètres est réalisée par maximum de vraisemblance (Dacünha-Castelle et Duflo, 1982) en minimisant $LV = -\log(V)$ grâce à un programme de minimisation de fonctions nonlinéaires sur le logiciel Splus 2000 (Mathsoft, 1999).

En développant $\frac{\partial LV}{\partial Q} = 0$, on vérifie que $\hat{Q} = \frac{B}{a}$.

Les hypothèses simplificatrices sont testées en réestimant les paramètres du modèle sous chaque hypothèse, puis en comparant la qualité d'ajustement du modèle ainsi obtenu à celle du modèle général complet par un test de maximum de vraisemblance (Lehman, 1983). Ce test évalue le gain de maximum de vraisemblance d'un modèle par rapport à un autre en tenant compte de la différence du nombre de paramètres entre ces modèles.

2.2- Résultats

2.2.1- Estimation des paramètres du modèle général complet

La log-vraisemblance du modèle général complet vaut 1079.1 pour 52 paramètres (Tableaux 4a, 4b et 4c). En autopollinisation, la distribution du nombre de graines par fruit du plant B ne montre qu'un seul pic correspondant à θ_1 . La fonction d'estimation des paramètres ajuste p à 1, et θ_2 est mal estimé. En pollinisation libre, pour des raisons semblables, les θ_2 des plants B et

C sont mal estimés, ainsi que le θ_1 et le p du plant E (1 seul pic dans la distribution de graines, correspondant à θ_2).

plant	effectif	θ_1	θ_2	p	q	LV
A	16	13.5112	31.3265	0.5551	0.6400	70.89
B	39	11.1491	28.1880	0.3271	0.8667	167.98
C	21	11.2335	29.6560	0.3510	0.9565	84.76
D	21	11.6194	28.3888	0.6791	0.7241	100.88
E	24	11.4206	28.1222	0.3466	0.7692	123.97
Total:						548.48

Tableau 4a: Estimation des paramètres du modèle général complet pour l'allopollinisation

plant	effectif	θ_1	θ_2	p	q	LV
A	11	6.2188	16.4370	0.7113	0.3060	53.41
B	11	7.4502	17.8707	1.0000	0.3931	44.68
C	15	3.1142	16.9517	0.5869	0.4007	75.82
E	20	9.1359	21.2474	0.4374	0.7353	100.21
Total:						274.11

Tableau 4b: Estimation des paramètres du modèle général complet pour l'autopollinisation

plant	effectif	θ_1	θ_2	p	q	LV
A	7	7.1076	19.4312	0.6962	0.2052	43.85
B	11	7.5724	12.4884	0.7473	0.3751	53.07
C	12	4.9650	27.9945	0.9172	0.6111	59.11
E	22	12.0000	30.0549	0.1730	0.8889	100.53
Total:						256.56

Log-vraisemblance totale du modèle pour les 3 modes de pollinisation: 1079.14

Nombre total de paramètres du modèle pour les 3 modes de pollinisation: 52

Tableau 4c: Estimation des paramètres du modèle général complet pour la pollinisation libre

Ces résultats montrent que:

- les distributions de nombres de graines par fruit des plants suivis peuvent être ajustées en général par une distribution bimodale, y compris en auto et en allopollinisation. De ce fait, les deux pics, qui avaient été identifiés sur des distributions provenant de fruits issus de pollinisation libre, ne correspondent pas à des fleurs pollinisées soit par de l'autopollen (pic des fruits contenant peu de graines), soit par de l'allopollen (pic des fruits contenant beaucoup de graines)
- en allopollinisation: θ_1 et θ_2 sont élevés et varient peu entre plants. Le rapport $\alpha = \frac{\theta_2}{\theta_1}$ est stable et vaut 2.44 en moyenne. Le coefficient de partage p est variable. Le taux de fructification q est peu variable et est élevé en général.

- En autopolinisation: θ_1 et θ_2 sont variables entre les plants et ont des valeurs plus faibles qu'en allopollinisation. Le rapport α est variable et est en moyenne plus élevé (3.23) qu'en allopollinisation. Les paramètres p et q sont variables; q a des valeurs plus faibles qu'en allopollinisation, sauf pour le plant E.
- En pollinisation libre: θ_1 et θ_2 sont variables entre les plants et ont des valeurs intermédiaires par rapport à celles obtenues en allo- et en autopolinisation. Le rapport α est variable et intermédiaire entre ceux de l'allo- et de l'autopolinisation (2.88 en moyenne). Les paramètres p et q sont variables, avec des valeurs très dispersées pour q .

Ces résultats confortent l'hypothèse d'une influence de l'origine génétique du pollen sur les paramètres. Les valeurs intermédiaires des paramètres estimés en pollinisation libre suggèrent que ce type de pollinisation est une combinaison d'auto- et d'allopollinisation. La variabilité des paramètres met en évidence un effet plant en autopolinisation et en pollinisation libre.

2.2.2- Simplification du modèle général

Des hypothèses simplificatrices sont testées sur le modèle dans le cas de l'allo- et de l'autopolinisation pour lesquelles l'origine du pollen est identifiée, alors que ce n'est pas le cas pour la pollinisation libre.

a- Simplification du modèle pour l'allopollinisation

Du fait de la faible variabilité des paramètres estimés du modèle général complet en allopollinisation, le modèle peut être simplifié au maximum pour l'allopollinisation, en considérant θ_1 , θ_2 (ou $\alpha = \frac{\theta_2}{\theta_1}$), p et q constants entre les plants. Le modèle à 20 paramètres

devient alors un modèle à 4 paramètres (Tableau 5). Le test de maximum de vraisemblance n'est pas significatif ($P = 0.15$), indiquant l'équivalence entre ces deux modèles.

plant	θ_1	α	p	q	LV	LV gen	2·ΔLV
A	11.8974	2.4384	0.4372	0.7950	73.99	70.89	6.22
B	11.8974	2.4384	0.4372	0.7950	169.48	167.98	3.00
C	11.8974	2.4384	0.4372	0.7950	87.74	84.76	5.96
D	11.8974	2.4384	0.4372	0.7950	103.73	100.88	5.71
E	11.8974	2.4384	0.4372	0.7950	124.42	123.97	0.89
Totaux:					559.37	548.48	21.78

LV: log-vraisemblance; LV gen: log-vraisemblance du modèle général; $2 \cdot \Delta LV = 2 \cdot (LV - LV \text{ gen})$, à comparer au seuil à 5% du test de maximum de vraisemblance

Différence de paramètres entre ce modèle et le général: $20 - 4 = 16$

Test de maximum de vraisemblance: $\text{seuil}_{0.05} = 26.3$ ($P_{2 \cdot \Delta LV} = 0.15$)

Tableau 5: Estimation des paramètres pour le modèle simplifié à θ_1 , α , p et q constants en allopollinisation

b- Simplification du modèle pour l'autopollinisation

La plupart des hypothèses simplificatrices qui ont été testées dans ce cas (α constant; θ_1 et α constants; p constant; q constant; p et q constants) ont été rejetées. Cela est probablement à relier à la variabilité élevée des paramètres dans le modèle général complet pour l'autopollinisation. L'hypothèse de θ_2 constant est acceptée d'un point de vue statistique ($P = 0.32$; tableau 6a). Cependant, le plant B n'a pas de pic correspondant à θ_2 . Il s'ajuste avec $p=1$ et n'a pas de poids dans l'analyse. L'hypothèse acceptée est donc que θ_2 peut être considéré constant pour les plants qui présentent un pic correspondant à ce paramètre.

Le rejet de l'hypothèse de p constant ($P = 0.02$; tableau 6b) est lié au plant B qui a un poids important dans la différence de maximum de vraisemblance entre les modèles ($2 \cdot \Delta LV$) dû à la mauvaise estimation de son paramètre p . De ce fait, il paraît raisonnable d'accepter l'hypothèse de p constant en autopollinisation, malgré le test significatif (il ne l'est pas sur les 3 plants A, C, E: $P=0.72$).

Plant	θ_1	θ_2	p	q	LV	LV gen	$2 \cdot \Delta LV$
A	6.2986	18.8563	0.7284	0.3060	53.80	53.41	0.78
B	7.4502	18.8563	1.0000	0.3931	44.68	44.68	0.00
C	3.2568	18.8563	0.6014	0.3995	76.30	75.82	0.97
E	7.9694	18.8563	0.3379	0.7354	101.07	100.21	1.74
Totaux :					275.85	274.11	3.49

LV: log-vraisemblance; LV gen:log-vraisemblance du modèle général; $2 \cdot \Delta LV = 2 \cdot (LV - LV \text{ gen})$, à comparer au seuil à 5% du test de maximum de vraisemblance

Différence de paramètres entre ce modèle et le général: $16 - 13 = 3$

Test de maximum de vraisemblance: $\text{seuil}_{0.05} = 7.81$ ($P_{2 \cdot \Delta LV} = 0.32$)

Tableau 6a: Estimation des paramètres pour le modèle simplifié à θ_2 constant en autopollinisation.

plant	θ_1	θ_2	p	q	LV	LV gen	$2 \cdot \Delta LV$
A	6.1734	16.1908	0.6513	0.3060	53.49	53.41	0.15
B	7.0303	17.8708	0.6513	0.3931	48.24	44.68	7.14
C	3.1534	17.0845	0.6513	0.4014	75.94	75.82	0.24
E	10.1275	22.4387	0.6513	0.7353	101.28	100.21	2.15
Totaux:					278.95	274.11	9.68

LV: log-vraisemblance; LV gen:log-vraisemblance du modèle général; $2 \cdot \Delta LV = 2 \cdot (LV - LV \text{ gen})$, à comparer au seuil à 5% du test de maximum de vraisemblance

Différence de paramètres entre ce modèle et le général: $16 - 13 = 3$

Test de maximum de vraisemblance: $\text{seuil}_{0.05} = 7.81$ ($P_{2 \cdot \Delta LV} = 0.02$)

Tableau 6b: Estimation des paramètres pour le modèle simplifié à p constant en autopollinisation.

2.3- Conclusions

Les paramètres θ_1 , θ_2 , p et q du modèle de fécondation-fructification varient en fonction du type de pollinisation et du plant. La pollinisation libre apparaît comme un mélange d'autopollinisation et d'allopollinisation, ce qui est réaliste.

En allopollinisation, il n'y a pas d'effet du plant sur les paramètres, et un paramétrage unique pour tous les plants peut être retenu.

En autopollinisation, les paramètres sont variables entre les plants, ce qui est à rapprocher de la variabilité du nombre moyen de graines par fruit et du taux de fructification constatées dans l'essai de pollinisation contrôlée. Toutefois, l'hypothèse de θ_2 constant est acceptée pour les plants qui ont un pic correspondant à ce paramètre. L'hypothèse de p constant est acceptable, ce qui signifie que le partage de la distribution du nombre de graines par fruit entre les deux lois de Poisson est constant.

Ces résultats sont toutefois à relativiser en raison des faibles effectifs des échantillons, tant le nombre de plants (effet plant sur les paramètres) que le nombre de fruits par traitement et par plant (distribution des nombres de graines par fruit).

3- Explicitation des paramètres du modèle général: le modèle en λ

3.1- Construction du modèle en λ

Afin de relier la variabilité des paramètres du modèle général à l'effet des facteurs étudiés sur la fécondation, les paramètres θ_1 et θ_2 ont été explicités. D'après la construction des lois de Poisson (voir § 1.1), leur paramètre vaut:

$$\theta = F \cdot \lambda$$

où: F est la fertilité ovulaire, soit la probabilité qu'un ovule fécondé donne une graine,
 λ est le nombre de tubes polliniques fécondant arrivant au niveau des ovules.

Les résultats de l'essai de pollinisation contrôlée, corroborés par les résultats précédents sur les paramètres du modèle général, montrent que l'allopollen est fortement compatible, quelle que soit son origine, alors que la compatibilité de l'autopollen est partielle et variable entre plants. Il faut donc distinguer les tubes polliniques d'allo et d'autopollen arrivant au niveau des ovules. Alors λ s'écrit:

$$\lambda = V \cdot (\lambda_1 \cdot f_1 + \lambda_2 \cdot f_2)$$

où: V est la viabilité pollinique, c'est à dire le taux de germination du pollen sur le stigmate, que l'on considère constante et indépendante de l'origine génétique du pollen.
 λ_1 est la quantité d'autopollen déposée sur le stigmate
 λ_2 est la quantité d'allopollen déposée sur le stigmate
 f_1 est la compatibilité pour l'autopollen, ou autocompatibilité ($f_1 \in [0;1]$)
 f_2 est la compatibilité pour l'allopollen.

L'expression de λ est surparamétrée (V, f_1, f_2), et l'on choisit de fixer la compatibilité pour l'allopollen en supposant qu'elle est totale et indépendante des génotypes male et femelle, soit $f_2 = 1$. Cette hypothèse est en contradiction avec les observations d'incompatibilité croisée variable entre cultivars chez la goyave (Seth, 1960) ou l'olive (Cuevas et Polito, 1997). Elle est cependant confortée d'une part par le fait qu'il n'y a pas de différence significative du nombre moyen de graines par fruit entre les plants en allopollinisation, et d'autre part par l'acceptation, sur le modèle général, de l'hypothèse que les paramètres θ_1 et θ_2 sont constants en allopollinisation.

Alors $\theta = F \cdot V \cdot [\lambda_1 \cdot f_1 + \lambda_2] = R \cdot [\lambda_1 \cdot f_1 + \lambda_2]$ où R est un coefficient de réduction. La fertilité ovulaire et la viabilité pollinique interviennent en produit et ne peuvent être estimés séparément.

Les distributions de nombres de graines par fruit obtenues en auto- et en allopollinisation saturantes peuvent être ajustées en général par une distribution bimodale. De ce fait, les hypothèses que les deux pics sont dû à des fleurs auto ou allopollinisées, ou à des fleurs mal (pluie, absence d'insecte) ou bien pollinisées sont écartées. L'hypothèse d'un phénomène progressif, par exemple la baisse de la fertilité au fur et à mesure de l'avancement de la floraison liée à des ressources limitées (Delph, 1986), est écartée car dans ce cas, une distribution large et aplatie est attendue. On fait l'hypothèse que l'on a deux populations distinctes de fleurs par plant, l'une donnant des fruits avec peu de graines, et l'autre des fruits avec beaucoup de graines. Chacune est caractérisée par une autocompatibilité f_i et par une fertilité ovulaire F . Cette hypothèse s'écrit:

$$\theta_1 = R_1 \cdot [\lambda_{11} \cdot f_{11} + \lambda_{12}]$$

$$\theta_2 = R_2 \cdot [\lambda_{21} \cdot f_{21} + \lambda_{22}]$$

où: $R_1 = F_1 \cdot V$ (F_1 : fertilité ovulaire des fleurs correspondant à θ_1)

$R_2 = F_2 \cdot V$ (F_2 : fertilité ovulaire des fleurs correspondant à θ_2)

λ_{11} : quantité d'autopollen déposée sur le stigmate des fleurs correspondant à θ_1

λ_{12} : quantité d'allopollen déposée sur le stigmate des fleurs correspondant à θ_1

λ_{21} : quantité d'autopollen déposée sur le stigmate des fleurs correspondant à θ_2

λ_{22} : quantité d'allopollen déposée sur le stigmate des fleurs correspondant à θ_2

f_{11} : autocompatibilité des fleurs correspondant à θ_1

f_{21} : autocompatibilité des fleurs correspondant à θ_2

Il n'y a pas eu émasculatation des fleurs avant la pollinisation manuelle. Compte tenu du mode de pollinisation par frottement, il est probable que des grains d'autopollen venant de la fleur suivie se soient déposés sur le stigmate lors de l'opération. Notons λ_m cette quantité d'autopollen déposée sur le stigmate, et λ_M la quantité de pollen déposée par la fleur donneuse. On suppose, de part l'homogénéité du mode opératoire au cours de l'essai, que les quantités de pollen λ_m et λ_M sont constantes quels que soient la fleur et le type de pollinisation. Alors:

$$\text{Pour l'allopollinisation, } \theta_{1allo} = R_1 \cdot [\lambda_m \cdot f_{11} + \lambda_M]$$

$$\theta_{2allo} = R_2 \cdot [\lambda_m \cdot f_{21} + \lambda_M]$$

$$\text{Pour l'autopollinisation, } \theta_{1auto} = R_1 \cdot [\lambda_m + \lambda_M] \cdot f_{11}$$

$$\theta_{2auto} = R_2 \cdot [\lambda_m + \lambda_M] \cdot f_{21}$$

Notons $u_1 = R_1 \cdot \lambda_m$ et $u_2 = R_2 \cdot \lambda_m$. Les expressions précédentes deviennent:

$$\text{Pour l'allopollinisation: } \theta_{1\text{allo}} = u_1 \cdot [f_{11} + \lambda_M / \lambda_m] \quad \theta_{2\text{allo}} = u_2 \cdot [f_{21} + \lambda_M / \lambda_m]$$

$$\text{Pour l'autopollinisation: } \theta_{1\text{auto}} = u_1 \cdot [1 + \lambda_M / \lambda_m] \cdot f_{11} \quad \theta_{2\text{auto}} = u_2 \cdot [1 + \lambda_M / \lambda_m] \cdot f_{21}$$

En autopollinisation, on attribue la valeur 0 à f_{11} (resp. f_{12}) si le pic correspondant à $\theta_{1\text{auto}}$ (resp. $\theta_{2\text{auto}}$) n'existe pas. Alors $\theta_{1\text{auto}}$ (resp. $\theta_{2\text{auto}}$) = 0, et la loi de Poisson correspondante n'est pas calculée.

Les paramètres θ_1 et θ_2 du modèle général (équations [3] et [4]) sont remplacés par ces expressions. Les paramètres par plant sont alors u_1 , u_2 , λ_M / λ_m , f_{11} , f_{21} , p et q . Les trois premiers sont par hypothèse égaux pour tous les plants et types de pollinisation. Comme pour le modèle général, ces paramètres sont estimés simultanément sur chaque plant en allo- et en autopollinisation. Le plant D, qui n'a pas de données en autopollinisation et pour lequel on considère $f_{11} = f_{21} = 0$ (plant auto-incompatible) a été retiré du jeu de données. Le modèle ainsi paramétré est comparé au modèle général par un test de maximum de vraisemblance. Des hypothèses simplificatrices visant à réduire le nombre de paramètres sont ensuite testées.

3.2- Estimation des paramètres du modèle en λ

La log-vraisemblance du modèle en λ vaut 723.48 pour 27 paramètres (Tableau 7). Le test de maximum de vraisemblance comparant ce modèle avec le modèle général n'est pas significatif ($P = 0.62$; tableau 7). Ce résultat est cependant à relativiser car seule l'allopollinisation contribue à la différence de log-vraisemblance entre les modèles. Cela est lié à l'expression des θ_{auto} comme un produit de 3 facteurs qui s'équilibrent pour ne pas modifier la log-vraisemblance de chaque plant. Ce test indique qu'il y a équivalence entre ces deux modèles, et son résultat conforte l'approche utilisée pour expliciter les paramètres θ_1 et θ_2 .

modefec	plant	u_1	u_2	λ_M / λ_m	f_{11}	f_{21}	p	q	LV	LV gen	2·ΔLV
Allo	A	0.3713	0.9024	31.911	0.5113	0.5563	0.5319	0.6400	71.93	70.89	2.09
Allo	B	0.3713	0.9024	31.911	0.6088	0.0000	0.3543	0.8667	168.23	167.98	0.50
Allo	C	0.3713	0.9024	31.911	0.2547	0.5715	0.3606	0.9565	84.91	84.76	0.29
Allo	E	0.3713	0.9024	31.911	0.7452	0.7135	0.3764	0.7692	124.30	123.97	0.67
Auto	A	0.3713	0.9024	31.911	0.5113	0.5563	0.7125	0.3060	53.41	53.41	0.00
Auto	B	0.3713	0.9024	31.911	0.6088	0.0000	0.9859	0.3987	44.68	44.68	0.00
Auto	C	0.3713	0.9024	31.911	0.2547	0.5715	0.5871	0.4007	75.82	75.82	0.00
Auto	E	0.3713	0.9024	31.911	0.7452	0.7135	0.4350	0.7353	100.21	100.21	0.00
Totaux:									723.48	721.71	3.55

LV: log-vraisemblance; LV gen: log-vraisemblance du modèle général; 2·ΔLV = 2·(LV – LV gen), à comparer au seuil à 5% du test de maximum de vraisemblance

Différence de paramètres entre ce modèle et le général: 32 – 27 = 5

Test de vraisemblance: seuil_{0.05} = 11.1 ($P_{2 \cdot \Delta LV} = 0.62$)

Tableau 7: Estimation des paramètres pour le modèle en λ .

Le modèle en λ estime pour chaque plant une valeur de f_{11} et f_{21} , quantifiant ainsi la notion d'autocompatibilité partielle et variable du goyavier-fraise par deux composantes qui interviennent sur chaque loi de Poisson de la distribution des nombres de graines par fruit. Nos résultats suggèrent que ces deux composantes sont indépendantes ($r = 0.49$, $P = 0.41$), mais cela demande à être vérifié sur un échantillon plus grand de plants.

λ_M/λ_m est très supérieur à 1, ce qui est cohérent avec la signification de ce paramètre et conforte la bonne réalisation des pollinisations.

Les valeurs estimées des paramètres u_1 , u_2 et λ_M/λ_m peuvent être testées sur le plant D qui n'a pas été utilisé pour l'estimation. Pour ce plant, $f_{11} = f_{21} = 0$, soit $\theta_{1allo} = u_1 \cdot \lambda_M/\lambda_m$ et $\theta_{2allo} = u_2 \cdot \lambda_M/\lambda_m$. D'après les valeurs estimées de u_1 , u_2 et λ_M/λ_m , (Tableau 7), on obtient les estimations $\theta_{1D} = 11.85$ et $\theta_{2D} = 28.80$, qui sont des valeurs très proches des estimations de θ_1 ($=11.62$) et θ_2 ($=28.39$) pour ce plant par le modèle général en allopollinisation (Tableau 4a).

3.3- Simplification du modèle en λ

Les hypothèses simplificatrices testées sur le modèle en λ ont eu pour objectif d'explicitier les différences de comportement entre plants et entre types de pollinisation. Le modèle simplifié avec p constant en auto- et en allopollinisation, mais avec des valeurs différentes, et q constant en allopollinisation, est équivalent au modèle général ($P = 0.30$; tableau 8). La log-vraisemblance du plant B en autopollinisation ne change pas car si $f_{21} = 0$ (soit $\theta_{2auto} = 0$), la log-vraisemblance est fonction du produit pq , et non de p et de q . Alors q_{auto} reste libre et s'ajuste pour conserver pq constant.

modefec	plant	u_1	u_2	λ_M/λ_m	f_{11}	f_{21}	p	q	LV	LV gen	2·ΔLV
Allo	A	0.2574	0.6241	46.286	0.5018	0.5336	0.3892	0.8106	74.54	70.89	7.30
Allo	B	0.2574	0.6241	46.286	0.6117	0.0000	0.3892	0.8106	168.83	167.98	1.70
Allo	C	0.2574	0.6241	46.286	0.2547	0.5736	0.3892	0.8106	87.13	84.76	4.75
Allo	E	0.2574	0.6241	46.286	0.8011	0.7447	0.3892	0.8106	124.48	123.97	1.02
Auto	A	0.2574	0.6241	46.286	0.5018	0.5336	0.5683	0.3059	53.80	53.41	0.77
Auto	B	0.2574	0.6241	46.286	0.6117	0.0000	0.5683	0.6916	44.68	44.68	0.00
Auto	C	0.2574	0.6241	46.286	0.2547	0.5736	0.5683	0.4005	75.83	75.82	0.02
Auto	E	0.2574	0.6241	46.286	0.8011	0.7447	0.5683	0.7353	100.58	100.21	0.76
Totaux:									729.86	721.71	16.31

LV: log-vraisemblance; LV gen: log-vraisemblance du modèle général; 2·ΔLV = 2·(LV – LV gen), à comparer au seuil à 5% du test de maximum de vraisemblance

Différence de paramètres entre ce modèle et le général: $32 - 18 = 14$

Test de vraisemblance: $\text{seuil}_{0.05} = 23.7$ ($P_{2 \cdot \Delta LV} = 0.30$)

Tableau 8: Estimation des paramètres pour le modèle en λ simplifié avec p constant en auto- et en allopollinisation, et q constant en allopollinisation

Il n'y a par contre pas équivalence entre le modèle en λ avec p et q constants en auto- et en allopollinisation et le modèle général ($P = 0.01$). Par rapport au modèle précédent, on fait l'hypothèse supplémentaire que la probabilité de fructification pour une fleur ayant au moins un ovule fécondé en autopollinisation est constante, alors que cette valeur est très variable ([0.31; 0.74]). Ce sont d'ailleurs les plants A et E, qui ont des valeurs estimées de q extrêmes,

qui ont un poids élevé dans la différence de log-vraisemblance des modèles. D'un point de vue biologique, accepter cette hypothèse revient à ne pas tenir compte de la variabilité du taux de fructification en autopolinisation qui a été constatée dans l'essai de pollinisation contrôlée, ni de sa relation avec le nombre moyen de graines des fruits. Nous ne retenons donc pas cette hypothèse supplémentaire.

Les simplifications du modèle général en allopollinisation se retrouvent dans celles du modèle en λ . Les paramètres p et q sont considérés constants dans les deux modèles. Les paramètres θ_1 et θ_2 sont constants dans le modèle général. Dans le modèle en λ , le paramètre λ_M/λ_m est très supérieur à f_{11} et f_{21} , qui peuvent alors être négligés, d'où $\theta_{1allo} \approx u_1 \cdot \lambda_M/\lambda_m$ et $\theta_{2allo} \approx u_2 \cdot \lambda_M/\lambda_m$ qui sont constants.

En autopolinisation, l'hypothèse de θ_2 constant avait été retenue dans le modèle général pour les plants dont le pic correspondant existe. Dans le modèle en λ , chaque plant est défini par sa valeur de f_{21} , et on ne peut retenir l'hypothèse de f_{21} constant car ce paramètre est nul pour les plants B et D qui n'ont pas de pic correspondant à θ_{2auto} . Le modèle en λ permet ainsi d'explicitier par les autocompatibilités la variabilité de comportement entre les plants.

Les distributions de probabilité obtenues avec le modèle général complet et le modèle simplifié en λ donnent des abscisses proches pour les deux pics (Fig. 9 et 10). Le poids de chaque pic, défini par p , et la position relative des courbes en hauteur, définie par q , varient entre les modèles. Les deux modèles collent assez bien aux données qui sont cependant peu nombreuses et donnent donc des distributions très irrégulières.

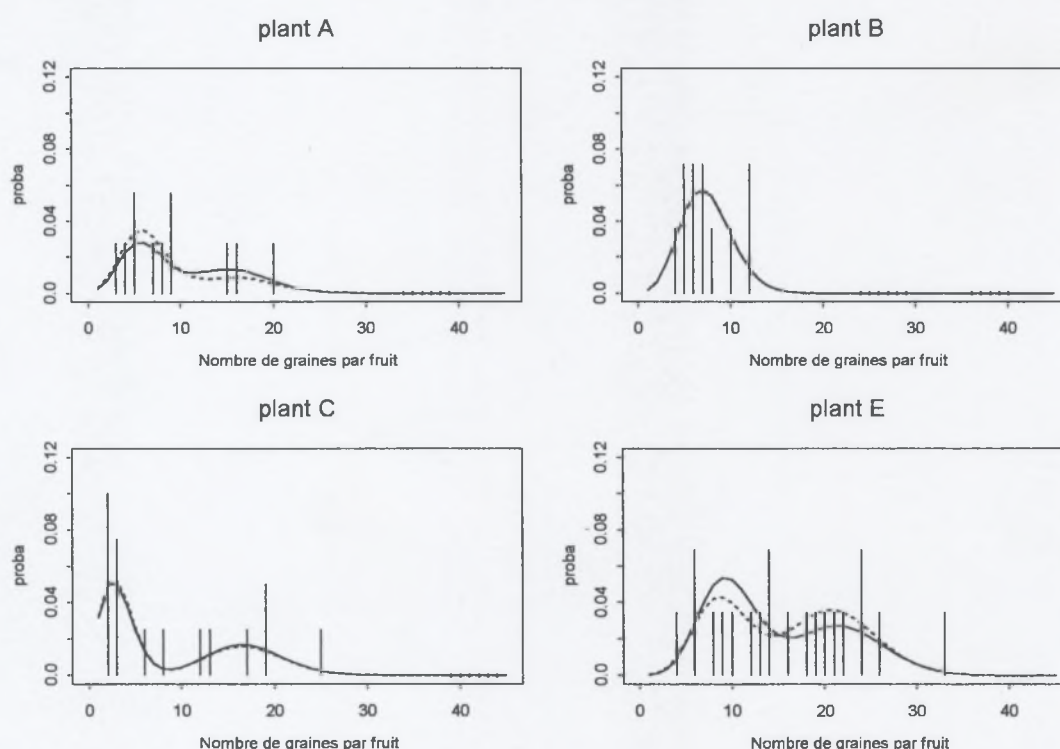


Figure 9: Distributions observées (barres verticales), ajustées avec le modèle général (pointillés), et avec le modèle simplifié en λ (trait plein) du nombre de graines par fruit après une autopolinisation saturante sur 4 plants de goyavier-fraise.

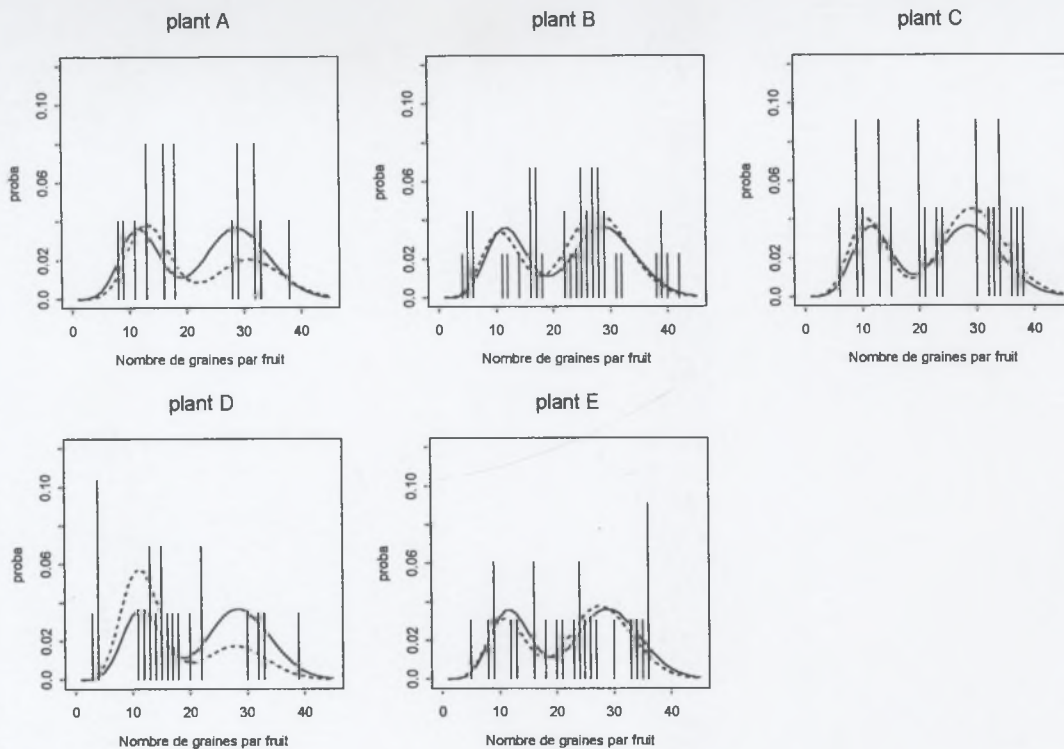


Figure 10: Distributions observées (barres verticales), ajustées avec le modèle général (pointillés) et avec le modèle simplifié en λ (trait plein) du nombre de graines par fruit après une allopollinisation saturante sur 5 plants de goyavier-fraise.

4- Discussion - Conclusions

L'effet du plant et de l'origine du pollen sur la distribution des nombres de graines par fruit chez le goyavier-fraise a été étudié à travers la construction d'un modèle statistique. Il décrit cette distribution comme la somme de deux lois de Poisson et intègre les effets de facteurs affectant la pollinisation et la fécondation-fructification. Les intensités des lois de Poisson dépendent des conditions de pollinisation ($u_1, u_2, \lambda_M/\lambda_m$) et des autocompatibilités f_{11} et f_{21} , propres à chaque plant, de chacune des deux populations de fleurs sous-jacentes à chaque loi de Poisson. Le coefficient de partage entre les deux lois de Poisson, p , est indépendant du plant, mais dépend de l'origine génétique du pollen. En allopollinisation, un poids plus fort est donné à la loi de Poisson correspondant aux nombres élevés de graines par fruit. Le taux de fructification q est indépendant du plant en allopollinisation, et est variable en autopolinisation, notamment en fonction du nombre de graines dans les fruits (Fig. 7A). Ce modèle permet d'estimer le niveau d'autocompatibilité d'un plant à partir des résultats de pollinisations saturantes contrôlées.

Le modèle d'analyse des résultats a été construit à partir des données de l'essai de pollinisation contrôlée qui constituent des échantillons restreints de plants et de fruits par plant. Les résultats doivent donc être relativisés par rapport à ces effectifs. Cependant, le petit nombre de plants est en partie compensé par leur diversité de comportement en autopolinisation qui a permis d'émettre certaines hypothèses sur le système d'auto-incompatibilité du goyavier-fraise. Mais il ne permet pas d'avoir une idée de la distribution des valeurs de f_{11} et f_{21} au sein

d'une population. De même, les valeurs des taux de fructification ne reflètent que ceux des plants étudiés et ne peuvent être généralisées. Ainsi, le taux constant de fructification en allopollinisation saturante (0.81) correspond à un maximum qui peut être observé en moyenne sur une parcelle, ces conditions de pollinisation étant optimales. Or les taux de fructification de 95 cycles de production relevés sur des parcelles à différentes altitudes et à différentes saisons ont varié de 0.05 à 0.95, avec une moyenne de 0.54. Quatorze valeurs (14.7% du total) étaient au-dessus de la valeur de 0.81 et ont varié de 0.82 à 0.95.

Les faibles échantillons de fruits par type de pollinisation et par plant donnent des distributions de graines par fruit probablement peu représentatives, mais les grandes tendances sont montrées. Ainsi, le plant B n'a pas donné de fruit contenant un nombre élevé de graines en autopollinisation (Fig. 9), ce qui nous a amené à considérer $f_{2I} = 0$ pour ce plant à partir d'un échantillon de 11 fruits. En pollinisation libre, où l'autopollinisation peut prédominer, ce plant n'a pas non plus donné de fruit avec un nombre élevé de graines (Fig. 6), ce qui conforte l'idée que $f_{2I} = 0$ pour ce plant.

Le modèle proposé peut apporter des éléments de réponse aux questions soulevées par les résultats, à condition d'utiliser des jeux de données, et donc des dispositifs expérimentaux, adaptés aux objectifs recherchés. Quel que soit le dispositif, des auto- et des allopollinisations saturantes doivent être réalisées sur chaque plant pour permettre l'estimation des paramètres du modèle. Pour avoir en moyenne des effectifs par plant semblables de fruits issus d'auto- et d'allopollinisation, nos résultats sur les taux de fructification indiquent qu'il faut réaliser deux fois plus d'autopollinisations que d'allopollinisations par plant. Si l'objectif recherché est l'obtention de distributions des nombres de graines par fruits plus représentatives, il est nécessaire de polliniser plus de fleurs par plant. Le temps passé à polliniser les fleurs peut être réduit en pollinisant un grand nombre de fleurs par plant chaque jour, par exemple au moment du pic de floraison, au lieu de polliniser toutes les fleurs de quelques rameaux, ce qui oblige à des passages quotidiens pour seulement quelques fleurs ouvertes. Si l'objectif recherché est de déterminer la distribution des valeurs de f_{1I} et f_{2I} au sein d'une population, le dispositif expérimental sera orienté vers un nombre élevé de plants sur lesquels seront faites les auto- et allopollinisations, avec un nombre suffisant de fleurs par traitement (au moins 40-50 fleurs par plant et par traitement d'après notre essai) afin de valider la présence de deux pics sur les distributions des nombres de graines par fruit et de pouvoir estimer correctement les paramètres du modèle.

Une hypothèse forte du modèle est qu'il y a deux populations de fleurs par plant, l'une ayant une capacité à donner des fruits avec peu de graines, et l'autre une capacité à donner des fruits avec de nombreuses graines. La répartition des fleurs entre ces deux populations est influencée par le type de pollinisation, qui détermine la valeur du coefficient de partage p , et non par le plant lui-même. Plusieurs hypothèses sur l'origine de ces populations ont déjà été écartées. L'effet de l'origine du pollen sur le paramètre p suggère que des facteurs génétiques sont impliqués dans le déterminisme de ces deux populations. L'auto-incompatibilité est en général déterminée par un locus multi-allélique. Le goyavier-fraise étant fortement polyploïde, on peut supposer en autopollinisation des effets différenciés sur les processus de l'autocompatibilité en fonction du génotype du gamète mâle par rapport au génotype du parent et/ou du gamète femelle. D'autre part, les résultats de Raseira et Raseira (1996) conduisent ces auteurs à supposer une apomixie facultative, qui n'a toutefois pas été vérifiée. Celle-ci pourrait s'exprimer de façon différenciée en fonction du génotype des gamètes mâles. Si l'apomixie est vérifiée, un modèle de fécondation basé sur l'hypothèse que les graines proviennent de la fécondation des ovules par des tubes polliniques est probablement inadapté.

Une connaissance plus fine de la biologie de la reproduction du goyavier-fraise est donc nécessaire pour envisager la construction d'un modèle de fécondation-fructification sur des bases biologiques solides.

D'autres hypothèses peuvent être avancées pour l'existence de deux populations de fleurs. Sedgley et Smith (1989) émettent l'hypothèse d'une variabilité de la fertilité chez les fleurs de genres à floraison groupée tels que *Eucalyptus* (Myrtaceae). Sanzol et Herrero (2001), dans une synthèse sur la période effective de pollinisation, rappellent que la qualité reproductrice de la fleur individuelle peut être affectée par divers facteurs (nutritionnels, orientation de la branche, températures avant floraison).

La relation positive entre le nombre de graines et le taux de fructification q en autopolinisation (Fig. 7A) remet en cause la constance de q qui est un résultat obtenu sur des fruits issus de pollinisation libre, type de pollinisation où cette relation n'apparaît pas (Fig. 7C). Il serait intéressant, sur un jeu de données plus conséquent que le notre, de tester un modèle de fructification dépendant du nombre d'ovules fécondés en autopolinisation.

Chapitre V

Etude et modélisation de la croissance du fruit du goyavier-fraise

La croissance du fruit est l'étape finale du cycle de production. Elle comporte deux composantes. La première, temporelle, définit la date de maturité ou de récolte du fruit. La seconde, quantitative, détermine le poids, ou calibre, du fruit à maturité ou à la récolte. Pour le goyavier-fraise, la date de récolte que l'on considère, et qui est celle utilisée en pratique, correspond à la date de pleine maturité quand le fruit chute spontanément du plant. Dans l'objectif du modèle global d'élaboration du nombre de fruits et de leur calibre, les sorties attendues du modèle de croissance du fruit sont la distribution des poids de fruit et la distribution des dates de récolte (Fig. 5). Ces données sont des éléments de décision pour l'agriculteur pour préparer le chantier de récolte et la commercialisation, et pour décider d'éventuels actes techniques sur la croissance du fruit.

Compte tenu du caractère polysperme du fruit, et par analogie avec les facteurs affectant en général la croissance des fruits, nous faisons l'hypothèse que la croissance du fruit du goyavier-fraise est affectée par le nombre de graines du fruit et par le rapport feuilles:fruits du plant. Ces facteurs sont déterminés (Fig. 5) par la pollinisation-fécondation-fructification (nombre de fruits et de graines par fruit) et en partie par l'élaboration du nombre de fleurs (nombre de feuilles issues des nouvelles pousses).

Dans une première partie, une démarche générale pour modéliser l'effet de facteurs sur la croissance individuelle du fruit est présentée (Normand, F., Chadœuf, J. et Habib, R. Coupling growth factors and statistical errors with a fruit growth model to predict fruit weight distribution. Manuscript soumis à *Journal of Horticultural Science & Biotechnology*). Un modèle d'erreur est intégré pour évaluer, par simulation, la précision des sorties du modèle. Cette approche est appliquée à la croissance du fruit du goyavier-fraise. La forme exponentielle de la courbe de croissance du fruit au moment de la maturité, et la faible précision sur la prédiction de la date individuelle de récolte conduisent à modéliser le poids final du fruit et sa date de récolte de façon indépendante. Un modèle de prédiction du poids final du fruit intégrant l'effet des facteurs de croissance cités ci-dessus est présenté dans cette première partie.

La seconde partie a pour objet la prédiction des dates de récolte. La date de récolte individuelle étant mal expliquée par les facteurs de croissance précédents, un modèle de prédiction de la distribution des dates de récolte à l'échelle du plant en fonction du rapport feuilles:fruits du plant et du nombre moyen de graines dans ses fruits est construit. Les démarches de modélisation et de simulation sont semblables à celles utilisées pour la modélisation de la croissance du fruit.

I- Couplage des facteurs de croissance et des erreurs statistiques à un modèle de croissance pour prévoir la distribution des poids de fruit

Coupling growth factors and statistical errors with a fruit growth model to predict fruit weight distribution

By F. NORMAND¹, J. CHADŒUF² and R. HABIB³

¹ Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Département des Productions Fruitières et Horticoles, BP 180, 97455 Saint Pierre cedex, Ile de la Réunion, France

² Institut National de la Recherche Agronomique, Biométrie, Domaine St Paul, Site Agroparc, F-84914 Avignon cedex 9, France

³ Institut National de la Recherche Agronomique, Plante et Systèmes de Culture Horticoles, Bât. A, Domaine St Paul, Site Agroparc, F-84914 Avignon cedex 9, France

Soumis à *Journal of Horticultural Science & Biotechnology*

SUMMARY

The fruit size distribution at harvest is of critical importance in horticulture. Several models based on fruit growth curve have been proposed to predict this distribution. Some models considered the effects of growth factors, but they rarely accounted for statistical errors which define, with the model pertinence, the model accuracy. We proposed to express fruit growth model parameters as functions of growth factors, leading to a nested growth model. Statistical errors were integrated in the model and a method based on simulation was used to estimate the prediction limits and the contribution of each source of error on the prediction variability. This approach was then applied to built a model to predict fruit weight distribution at maturity for strawberry guava (*Psidium cattleianum*), with the seed number and the tree leaf:fruit ratio as growth factors. Strawberry guava fruit has a biphasic growth curve without final plateau: the fruit drops spontaneously at full maturity while it is in a rapid growth phase. Growth curve was modelled as the sum of a logistic and an exponential function. The model gave information on the growth factors effects on each growth period. But the prediction accuracy was limited, due to a rough statistical relationship between the ratio of fruit weight at maturity on fruit weight after the first growth phase and growth factors. A second model was derived and directly fitted on the data. It was simpler, more accurate and robust, but it did not provide information on growth factors effects on fruit growth. The importance of the experimental range where the model was fitted was pointed out for a general use of the model. Results also suggested that strawberry guava fruit growth was affected by factors not considered in this study, such as local inter-fruit competition.

The distribution of the fruit size, or of the fruit weight, at harvest is of critical importance in horticulture to optimise orchard profitability. Its prediction, as early as possible after bloom, allows to prepare appropriately fruit packaging and marketing, and to intervene if necessary on fruit growth with specific techniques such as fruit thinning. The classical approach to determine the fruit size distribution at harvest is to measure a fruit size sample early in the season and to project forward in time using appropriately scaled fruit growth curves (Welte, 1990; Lakso *et al.*, 1995; Gandar *et al.*, 1996; Bramardi *et al.*, 1998). The prediction accuracy, given by the prediction limits, is also important for the user and may be support of decision. However, it is rarely provided by growth models.

Fruit growth curves are affected by external factors such as temperature or radiation, in particular for temperate fruit species, so that parameters estimation values vary with season and location (Lakso *et al.*, 1995; Gallego *et al.*, 1997; Bramardi *et al.*, 1998; Stanley *et al.*, 2000). Likewise, fruit growth curves are affected by tree factors such as cultivar (Lakso *et al.*, 1995; Fanizza and Colonna, 1996) or fruit load (Burge *et al.*, 1987; Inglese *et al.*, 1995). They also vary greatly between individual fruit, even on a same tree (Génard and Bruchou, 1993).

The identification of the main growth factors and their integration in fruit growth models would lead to models of more general use to predict fruit size distribution. One approach is the construction of process-based models (Buwalda, 1991; Grossman and DeJong, 1994; Fishman and Génard, 1998). They are deterministic and contain many parameters to be estimated. They need climatic data collected at a small time-step as input. They are generally more suitable for research purposes than for the prediction of fruit weight distribution at harvest as a decision support. A probably simpler and more applied approach is to work on the fruit growth curve, but very few work has been done in this way. Gandar *et al.* (1996) integrated a form of time-varying response of fruit growth to temperature at the level of the differential equations of the model. Lescourret *et al.* (1999) built individual kiwifruit growth curves with growth factors acting on a reference relative growth rate or on the potential growth rate. A different way, that we propose to explore, is to specify relationships between the model parameters and growth factors, and to express the former as function of the latter. Indeed, changes in fruit growth curves are related to variability in parameters values (Génard and Bruchou, 1993; Lakso *et al.*, 1995; Fanizza and Colonna, 1996; Stanley *et al.*, 2000) rather than to a change of the growth curve model. The model relying growth factors to fruit growth would then be formulated as nested functions.

It is particularly important for prediction models to give the prediction limits of an estimated value. These limits are related to the model pertinence and to the statistical errors which are however rarely considered in fruit growth studies. Nevertheless, to quantify the different levels of statistical error allows to determine their contribution in model prediction and eventually reduce them by adapting sampling and measurement efforts on variables with respect to the influence of their uncertainty on prediction accuracy. Wallach and Génard (1998) showed in particular that variability in input or parameter values affects the model variance contribution to the mean square error of prediction. The integration of statistical errors in a prediction model allows to account for the over-all variability, partly related to the growth factors not explicitly considered in the model. The above-mentioned nested models contain specific sources of error representing the cost of the growth factors integration into the model. Their modelling is thus important.

The purpose of this paper is to present a fruit growth model coupled with the effect of growth factors on the model parameters, and with the related statistical errors. This approach

is then applied to build a strawberry guava (*Psidium cattleianum* Sabine) fruit growth model and use it to predict the fruit weight distribution at harvest. A simulation method is proposed to estimate the contribution of the different sources of statistical errors and the prediction limits of the individual final fruit weight and of the final fruit weight distribution.

MATERIALS AND METHODS

Data collection

Strawberry guava is a subtropical shrub producing small berries (1-4 cm in diameter) containing several small hard seeds. It is a promising new fruit species in Brazil (Raseira and Raseira, 1996) or in Réunion Island (Normand, 1994). The growth factors are unknown for this species, but two classical factors may be put forward. The polyspermous fruit suggests that the seed number (*SN*) may have an effect on growth as in kiwifruit (Pyke and Alspach, 1986), custard apple (George *et al.*, 1986) or saskatoon (McGarry *et al.*, 1998). Leaf:fruit ratio (*LFR*) is an indicator of potential assimilates production per fruit which affects fruit growth in various species (Hansen, 1989; Reddy and Singh, 1991). We propose to study the influence of these two factors on strawberry guava growth.

Fruit growth was studied in a 4 year-old strawberry guava orchard located at 480 m of elevation on the East side of Réunion Island (lat. 21°06' S, long. 55°32' E). Plants were from seedling, and spacing was 3 m x 1 m. Just after the flowering period in October 1997, 27 trees of similar size, foliage density and stem section at the soil level were selected. Nine *LFR* were obtained at the tree level by combining three levels of defoliation (removal of 0, 50 or 75 % of the leaves) with three levels of fruit thinning (low, medium and high fruit load). A complete randomised block design with three replicates was used. Leaves and fruit were then counted on each tree and additional leaves or fruit removal was eventually performed so that the three replicates of a same treatment have about the same *LFR* (range: 0.8 - 13.0). Treatments were set up at fruit set, 11 to 17 days after mid-bloom. Newly emerged shoots were removed during fruit growth in order to avoid a leaf number increase. The drop of immature fruit is rare after fruit set in strawberry guava.

Five fruit per tree were randomly sampled and labelled after treatment set up. Their length (excluding persistent sepals) and diameter were measured weekly from 24 days after mid-bloom, and every 2-3 days when the skin colour began to turn red. Studied fruit were picked up at full maturity (i.e. natural drop), measured, weighted, and their seeds were counted. Because of fruit loss at harvest, data were available for 132 fruit. The other fruit of the 27 studied trees were harvested every 2-3 days at full maturity and were counted and weighted.

Fruit growth was expressed as the change in cumulated fruit fresh weight, latter called fruit weight. For each date, fruit weight (*W*) was estimated from fruit length (*l*) and diameter (*d*) with the relation: $W = 1.7 \cdot 10^{-4} \cdot \pi \cdot d^2 \cdot l$ ($r^2 = 0.997$; $P = 0$). This relation was obtained with data measured in two other experiments on fruits from fruit set to harvest. It was stable across year, orchard and *LFR* in the range 0.7-6 (unpublished data). We verified that fruit weight at maturity was well estimated by this relation, even for *LFR* > 6. The actual fruit weight at maturity was used for the final point of fruit growth curves.

The time scale was expressed as accumulated degree.days since mid-bloom. Data collected in orchards over several years and elevations showed that the delay between mid-bloom and mid-harvest was 3,134°C.d with a base temperature of -4°C. This base temperature

was statistically estimated by different methods (Arnold, 1959; Yang *et al.*, 1995) giving concordant results, and has no physiological significance. Negative statistical base temperatures have already been reported (Wielgolaski, 1999).

The fruit growth model was validated at two levels. First, on the same production cycle and the same trees than previously, but on different fruit: 35 mature fruit were chosen in the total harvest of each tree in order to represent a large range of fruit size. Their individual length, diameter, weight and seed number were determined. The second level of validation used data collected during the 98-99 production cycle on a 6 year-old orchard located at 520 m of elevation. Climatic conditions were similar during the first (480 m) and the second (520 m) production cycles used for validation: from mid-bloom to mid-harvest on the studied trees, mean daily air temperature was 21.6 and 21.0°C, and total rainfall was 3,215 and 3,691 mm respectively. Eleven trees of similar size were sampled, and their fruit and leaves were counted at fruit set to determine their individual *LFR*. All fruit were harvested in March-June 1999 as in the previous experiment. During the harvest peak, 5 to 10 mature fruit per tree were randomly sampled among the fruit harvested each date. Their individual weight and seed number were determined. Over-all, 1,683 fruit were analysed, with 92 to 185 fruit per tree.

General approach of the nested fruit growth models and related statistical errors

The model: We first present the relation model coupling a fruit growth model with the growth factors. Let consider a fruit growth model where a variable y related to growth is a function of time t expressed in time units or as physiological time:

$$[I] \quad y(t) = f(t, \Theta) + \varepsilon$$

where Θ is a vector of parameters θ_i , $i \in [1, p]$, and ε is a random term with expectation zero.

We suppose that some factors, constant during fruit growth, affect the fruit growth curve. They can be intrinsic factors, such as seed number or fruit position in the canopy, or factors related to the shoot (e.g. shoot leaf:fruit ratio, orientation) or to the tree (e.g. crop load, cultivar). They generate variability on fruit growth among fruit, shoots or trees respectively. We suppose that the other growing conditions are not limiting. We denote U the vector of the growth factors values for a particular fruit. We assume that each growth model parameter θ_i is a linear combination of predictors $X_i(U)$ functions of U :

$$[II] \quad \theta_i(U) = X_i(U)\Omega_i + \varepsilon_a$$

where Ω_i is the vector of p_i parameters of the linear combination, and ε_a is a random term associated to the linear regression with expectation zero.

The fruit growth model coupled with growth factors is then defined as:

$$[III] \quad y(t, U) = f(t, \Theta(U)) + \varepsilon$$

where $\Theta(U)$ is the vector of parameters $\theta_i(U)$ (equation [II]). This nested model integrates the growth factors effects through their influence on the growth model parameters, which allows to simulate individual fruit growth curve.

Parameters estimation: Direct estimation of such models is quite impossible due to the number of parameters involved and to the high nonlinearity of the growth model f (France and Thornley, 1984). Therefore, a two steps method is proposed. Individual growth is measured on N fruit: for a fruit j ($j \in [1, N]$), a growth variable y_j is measured at Q_j different times t_j . We have then Q_j pairs (t_{jq}, y_{jq}) , $q \in [1, Q_j]$. Moreover, the value U_j of the growth factors affecting the fruit j growth are known.

First, the growth curve parameters θ_{ij} are estimated for each fruit j (equation [I]). We assume that the ratio of the variance of the estimated parameter $\hat{\theta}_{ij}$ for a fruit j on the variance of the true parameters θ_{ij} within the population of fruit, i.e. $\text{var}(\varepsilon_a)$, is low, so that we will neglect in the following the estimation errors on parameters $\hat{\theta}_{ij}$ and suppose them fixed and named θ_{ij} .

Second, the parameters Ω_i of the linear regression between growth model parameters θ_i and predictors functions of growth factors $X_i(U)$ are estimated from the data $(\theta_{ij}, X_i(U_j))$, $j \in [1, N]$. The generic formulation of the linear regression is $\theta_i(U) = X_i(U)\hat{\Omega}_i + \varepsilon_{\theta_i}$, where $\hat{\Omega}_i$ is the estimator of the linear regression vector of parameters, and ε_{θ_i} is the residual error of the linear regression, assumed to be normally distributed with expectation zero and variance $\text{var}(\varepsilon_{\theta_i})$.

Prediction: Our objective is to predict at time t the value of the growth variable, $\hat{y}(t, U)$, for any one fruit characterised by the growth factors value U , as well as the variability of this prediction, in particular for fruit different than those upon which the models were fitted. The predicted value is $\hat{y}(t, U) = f(t, \hat{\Theta}(U))$, where $\hat{\Theta}(U)$ is the vector of estimated parameters $\hat{\theta}_i(U) = X_i(U)\hat{\Omega}_i^*$ ($\hat{\Omega}_i^*$ is a particular estimation of the estimator $\hat{\Omega}_i$).

A confidence interval can be calculated for each particular value $\hat{\theta}_i(U)$ estimated through the linear regression (Dagnelie, 1998). But this confidence interval cannot be used through the nonlinear growth model f to calculate a confidence interval for $\hat{y}(t, U)$. We propose to supply this problem with a numerical method. It consists in simulating the variability of $\hat{\theta}_i(U)$ by numerous random sampling in $\hat{\theta}_i(U) = X_i(U)\hat{\Omega}_i + \varepsilon_{\theta_i}$. For each sampled $\hat{\theta}_i(U)$, a predicted $\hat{y}(t, U)$ is sampled in $\hat{y}(t, U) = f(t, \hat{\Theta}(U)) + \varepsilon_g$, where ε_g is the residual error of the growth curve fit, and $\hat{\Theta}(U)$ is the vector of previously sampled parameters $\hat{\theta}_i(U)$, in order to estimate $\hat{y}(t, U)$ variability for a given t and U . As stated by Dagnelie (1975, p 292) for the confidence interval calculation of a particular value estimated through a linear regression, two independent sources of error intervene and add up. The first one is the estimation error of the vector of parameters $\hat{\Omega}_i$. The value $\hat{\Omega}_i^*$ estimated from the data follows in \mathbb{R}^p a multinormal distribution with mean Ω_i , the true parameter values, and with variance-covariance matrix C_i . If the measurements were repeated on the same fruit, different values θ'_{ij} would have been obtained for the growth curve parameters. Alike, a new vector of estimated parameters $\hat{\Omega}'_i$ would have been obtained, which follows a multinormal distribution with mean Ω_i and with variance-covariance matrix C'_i . Thus a source of variability is linked to the data used to fit the

linear regression. However, if only one estimation $\hat{\Omega}_i^*$ is done, which is generally the case, and if one wants to take into account the statistical error on $\hat{\Omega}_i$ estimation, we propose to consider that $\hat{\Omega}_i^*$ is the true value of parameters Ω_i , and that $\hat{\Omega}_i$ variability follows a multinormal distribution Δ_i with mean 0 and variance-covariance matrix C_i estimated from the data, so that $\hat{\Omega}_i = \hat{\Omega}_i^* + \Delta_i$. These two approximations cannot be quantified, but they are reasonably acceptable for a large sample size N . Thus one simulates a fitting repetition of the linear regression when sampling an error in Δ_i to calculate an estimated vector of parameters $\hat{\Omega}_i$.

The second source of error is the particular variability of the individual value $\hat{\theta}_i(U)$, which is assumed to follow a central normal distribution with variance the residual variance of the regression $\text{var}(\varepsilon_{\theta_i})$. This variability is related to the data dispersion around each model. This dispersion is due to the measurement errors, but also to the effect of growth factors, acting at the fruit level for some of them, that are not explicitly taken into account in the model. So, the prediction variability must account for the statistical errors at the individual fruit level.

In the same way, the particular variability of the individual value $\hat{y}(t, U)$ must be considered. This variability is assumed to follow a central normal distribution with variance the residual variance of the growth model, $\text{var}(\varepsilon_g)$.

Then the prediction of a particular value $\hat{\theta}_i(U)$ is a random sampling in the variable

$$[\text{IV}] \quad \hat{\theta}_i(U) = X_i(U)\hat{\Omega}_i + \varepsilon_{\theta_i} = X_i(U)\hat{\Omega}_i^* + X_i(U)\Delta_i + \varepsilon_{\theta_i}$$

And the prediction of an individual value $\hat{y}(t, U)$ is a random sampling in the variable

$$[\text{V}] \quad \hat{y}(t, U) = f(t, \hat{\Theta}(U)) + \varepsilon_g$$

where $\hat{\Theta}(U)$ is the vector of sampled parameters $\hat{\theta}_i(U)$ (equation [IV]).

A large number of such simulations for the same t and U give an estimated distribution for $\hat{y}(t, U)$, so that its variability can be estimated, for example as the variance of the simulated values $\text{var}[\hat{y}(t, U)]$. A confidence interval for a predicted value at the 5% level can be estimated by the 0.025 and 0.975 quantiles of the estimated distribution.

The variance $\text{var}[\hat{y}(t, U)]$ depends on the statistical errors Δ_i , ε_{θ_i} and ε_g . It could be useful to calculate the contribution of each source of error to $\text{var}[\hat{y}(t, U)]$. The analytical calculation of these contributions depends on the growth model formulation which may be complex. As previously, simulation is an interesting method to estimate the contribution of each source of error on a model prediction. This is now illustrated with the strawberry guava fruit growth model.

Application to the strawberry guava fruit growth model

We anticipate the results on strawberry guava fruit growth pattern in order to present the model. The fruit weight at maturity, W_f , is the output of interest to determine the fruit weight distribution at harvest. Growth is exponential before maturity and a small error on the maturity date prediction may lead to a large error on W_f prediction through the fruit growth curve. Then, W_f and maturity date must be predicted separately. We focus our study on W_f prediction.

From the shape of the individual growth curve, W_f can be written, independently of time, as $W_f = K \cdot P + \varepsilon_f$ where P is the cumulated growth during the first growth phase, K is a proportion parameter between W_f and P (see Results section), and ε_f is a random term. The parameters P and K are linear combinations of predictors function of the growth factors SN and LFR : $P = h(LFR, SN)$, and $K = g(\log(LFR))$ (see Results section). We propose to integrate the effect of SN and LFR in $W_f = K \cdot P + \varepsilon_f$ and to estimate, using simulation, the variability of the prediction $\hat{W}_f(U)$ and the contribution of each source of error to this variability. For a fruit characterised by $U = (LFR, SN)$, the prediction $\hat{W}_f(U)$ is a random variable (equation [V]):

$$[VI] \quad \hat{W}_f(U) = (X_K(U) \hat{\Omega}_K^* + X_K(U) \Delta_K + \varepsilon_K) \cdot (X_P(U) \hat{\Omega}_P^* + X_P(U) \Delta_P + \varepsilon_P) + \varepsilon_r$$

where $X_K(U) = (1, \log(LFR))$; $X_P(U) = (LFR, SN)$; $\hat{\Omega}_K^*$ and $\hat{\Omega}_P^*$ are the linear regressions parameters values estimated from the data; Δ_K and Δ_P are the parameters estimation errors, and ε_K and ε_P the residual errors, of the linear regressions of K and P on $X_K(U)$ and $X_P(U)$ respectively; ε_r is the residual error of the growth curve fit. The variance-covariance matrix of Δ_K and Δ_P , and the variances of ε_K , ε_P and ε_r are estimated from the data.

Although K is a scale factor between W_f and P , we assume that the estimators \hat{K} and \hat{P} are independent (we verified the independence of ε_K and ε_P : $\text{cov}(\varepsilon_K, \varepsilon_P) = -0.007 \sim 0$). $\hat{W}_f(U)$ is not normally distributed as products of normal distributions appear in its expression. Its expectation $E[\hat{W}_f(U)]$ and variance $\text{var}[\hat{W}_f(U)]$ are given by (see Appendix):

$$[VII] \quad E[\hat{W}_f(U)] = (X_K(U) \hat{\Omega}_K^*) \cdot (X_P(U) \hat{\Omega}_P^*)$$

$$[VIII] \quad \text{var}[\hat{W}_f(U)] = \text{var}(B_1) + \text{var}(B_2) + \text{var}(B_3) + \text{var}(\varepsilon_r)$$

$E[\hat{W}_f(U)]$ is the deterministic part of $\hat{W}_f(U)$ in equation [VI]. The variance $\text{var}[\hat{W}_f(U)]$ is the sum of four variances related to the different sources of error: B_1 depends only on the estimation errors Δ_K and Δ_P of the linear regression parameters; B_2 depends only on the residual errors ε_K and ε_P of the linear regressions, and B_3 depends on the products of the two previous sources of error.

The weight of each component of $\text{var}[\hat{W}_f(U)]$ is evaluated using simulation. The variance $\text{var}(B_1)$ is estimated from 50,000 values $\hat{W}_f^{B_1}(U)$ generated from equation [VI] by a random sampling of errors Δ_K and Δ_P in two central multinormal distributions calculated with the variance-covariance matrix of Δ_K and Δ_P respectively, and $\varepsilon_K = \varepsilon_P = \varepsilon_r = 0$. Then $\text{var}(B_1) = \text{var}[\hat{W}_f^{B_1}(U)]$. Likewise, $\text{var}(B_2)$ is estimated from 50,000 values of $\hat{W}_f^{B_2}(U)$ generated from equation [VI] by a random sampling of errors ε_K and ε_P in central normal distributions with variance $\text{var}(\varepsilon_K)$ and $\text{var}(\varepsilon_P)$ respectively, and $\Delta_K = \Delta_P = \varepsilon_r = 0$. Then $\text{var}[\hat{W}_f(U)]$ is estimated from 50,000 values of $\hat{W}_f(U)$ generated from equation [VI] with the errors ε_K , ε_P , Δ_K and Δ_P . $\text{var}(B_3)$ is obtained by difference. $\text{var}(\varepsilon_r)$ is calculated from the fitting of the $n=132$

individual growth curves: $\text{var}(\varepsilon_r) = \frac{\sum_{i=1}^{i=n} SCE_i}{\sum_{i=1}^{i=n} ddl_i}$ where SCE_i and ddl_i are respectively the residual

sum of square and the degrees of freedom of the fruit i growth curve fit.

Statistical analysis

Analysis of variance was performed to compare the *LFR* effects on the mean W_f per tree. Differences between means were assessed using Tukey's multiple comparison test. Growth curves were fitted for each individual fruit with a minimisation program for nonlinear function of Splus 2000 statistical package using the least-square criterion (Mathsoft, 1999). Relationships between variable and predictors were investigated by linear regression.

The purpose of the model is to predict final fruit weight distribution based on individual fruit weight. The model was validated on two data sets described in the data collection section. For each data set, individual fruit weights were distributed into 1g-classes in order to build the observed distribution. Such distributions were calculated at the tree level and at the level of the set of studied trees, later called orchard for simplicity. Let F be the number of fruit of the observed distribution. A simulation was performed as a random sampling of F individual fruit weights $\hat{W}_f(U)$ in the random variable $\hat{W}_f(U)$ (equation [VI]). These values were distributed into 1 g-classes in order to build a simulated distribution. One thousand simulated distributions, each of F fruit, were calculated in the same way. The mean simulated distribution and the prediction limits at the 5% level were constructed at the class level: for each 1g-class, the mean and the 0.025 and 0.975 quantiles of the 1000 simulated absolute frequencies were calculated. Thus we determined the most likely absolute frequencies simulated, or predicted, by the model in each class. The prediction region belted by the quantiles of each class was then compared to the actual distribution.

RESULTS

Actual W_f of the 132 studied fruit ranged from 1.0 to 22.9 g, with $\text{mean} \pm \text{sd} = 7.4 \pm 4.3$ g, and SN ranged from 2 to 27 with $\text{mean} \pm \text{sd} = 12.1 \pm 5.6$. Actual tree *LFR* varied from 0.8 to 13.4, and had a highly significant effect on the mean W_f per tree (Table I). A logarithmic

relationship linked the mean W_f per tree and LFR (Figure 1). For a given LFR , W_f increased linearly with SN (Figure 2).

TABLE I

Effect of the mean tree leaf:fruit ratio, obtained by various levels of fruit thinning and leaves removal, on the mean strawberry guava fruit weight at maturity (n=3; $P<0.001$)

Fruit load	Number of leaves	Mean leaf:fruit ratio	Mean fruit weight ^z (g)
High	Low	0.8	3.6 ^h
High	Medium	1.5	4.3 ^{fgh}
High	High	2.3	5.6 ^{efgh}
Medium	Low	1.5	5.3 ^{gh}
Medium	Medium	2.7	7.0 ^{deg}
Medium	High	4.0	8.1 ^{cd}
Low	Low	4.3	8.4 ^{bcd}
Low	Medium	6.9	10.2 ^b
Low	High	13.0	12.4 ^a

^z: Means followed by different superscript letters are significantly different (Tukey's multiple comparison test).

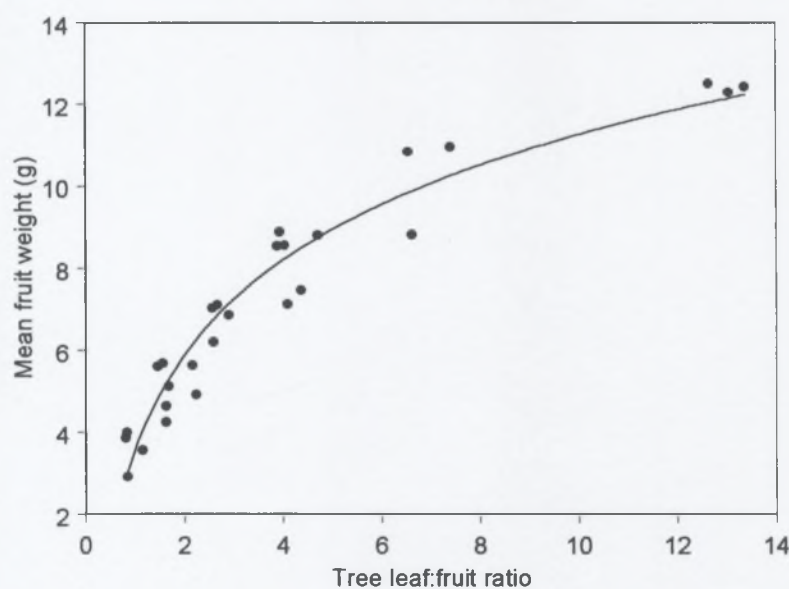


FIG. 1

Relationship between the tree leaf:fruit ratio and the mean strawberry guava fruit weight at maturity. Equation of the curve fit: $Y = 3.34 \cdot \log(x) + 3.56$; $n=27$; $r^2=0.94$; $P<0.001$.

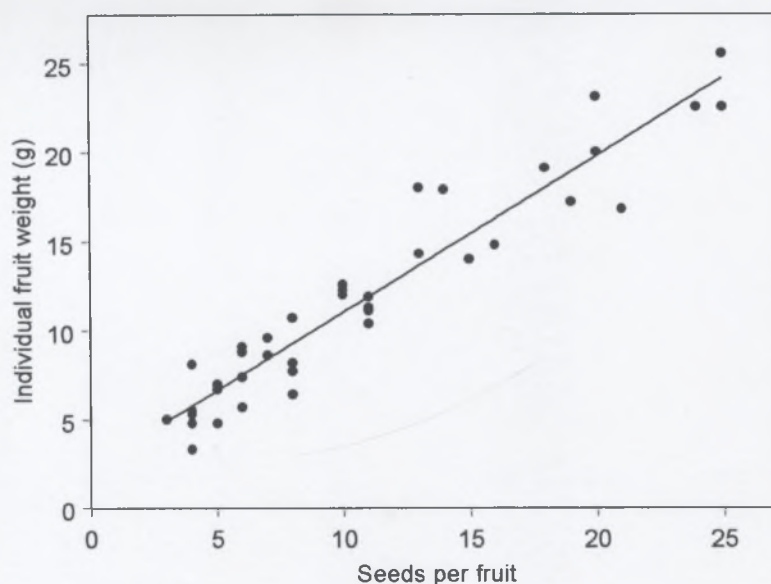


FIG. 2

Relationship between the seed number and the individual fruit weight at maturity on a strawberry guava tree with a 12.6 leaf:fruit ratio. Equation of the linear fit: $Y = 0.87 \cdot x + 2.27$; $n=40$; $r^2=0.91$; $P<0.001$.

Individual strawberry guava fruit growth curve may be divided into three periods: one period of rapid growth, another of much slower growth, and a last period of exponential growth (Figure 3). It is noteworthy that the fruit fell spontaneously at maturity while they were in a rapid growth period. No noticeable growth reduction was observed as in classical double-sigmoid growth pattern (Coombe, 1976). Like Génard and Bruchou (1993) for peach, or Fanizza and Colonna (1996) for grape berry, we assumed the cumulated growth of strawberry guava fruit to be according a two phases growth characterised by a two growth curves model. The first model, which describes the first growth phase, is represented by a logistic function, and the second model, which describes the final growth phase, is represented by an exponential function. The period of slow growth corresponds to the overlapping of the two growth curves. The fruit growth model was defined as:

$$[\text{IX}] \quad \begin{cases} W(t) = \frac{P}{1 + e^{-\alpha(t-t_l)}} + e^{\beta(t-t_E)} + \varepsilon_w \\ W_f = K \cdot P + \varepsilon_f \end{cases}$$

Where t is development-time expressed in degree-days, and ε_w and ε_f are random terms. The parameters P , α , t_l , β , t_E , and K have a biological meaning: P is the cumulated fruit growth during the first growth phase; t_l is the date of the maximal absolute growth rate during the first growth phase; α is proportional to the relative growth rate at t_l ; β is the constant relative growth rate during the second growth phase; and t_E is the date when the growth rate of the second growth phase is β . K is a proportion parameter between W_f and P . The

parameters P , α , t_1 , β and t_E were estimated for each fruit. K was calculated from the actual W_f and the estimated P value for each fruit.

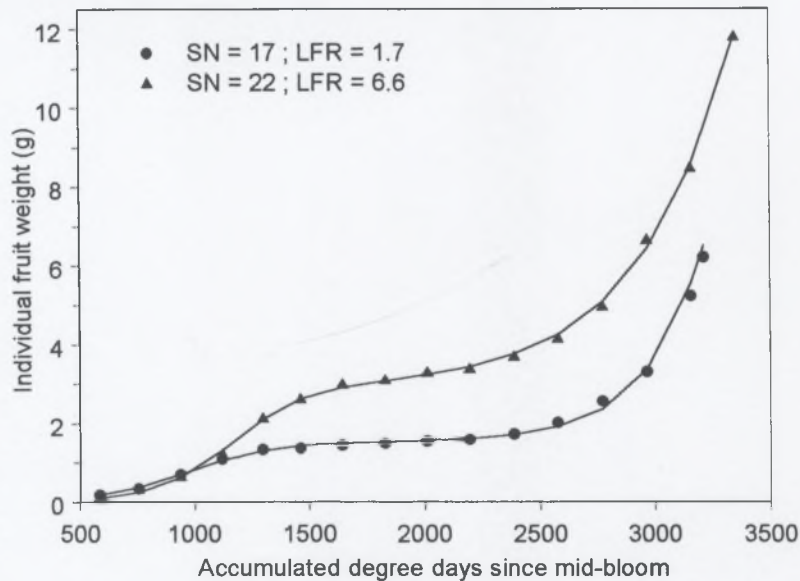


FIG. 3

Empirical (points) and modelled (line) cumulated growth of two strawberry guava fruits with different seed number (SN) and tree leaf:fruit ratio (LFR).

The model fitted well the 132 individual growth curves (Figure 3). Apart for one fruit, the residual standard error was lower than 0.4 g. Thus the splitting of individual growth curves into two parts was satisfactory. Estimated parameter values were variable within the population of fruit. The ratios of the variance of P estimated for each fruit by the variance of P within the population of fruit were low ($2 \cdot 10^{-4}$ to 0.21), then verifying the condition to neglect the statistical error on P estimation. P was a linear combination of LFR and SN : $P = 0.081 \cdot SN + 0.133 \cdot LFR$ ($R^2=0.96$, $n=132$, $P<0.001$). This is concordant with the relative position of individual growth curves which diverged as early as the first growth phase according to their respective LFR and SN (Figure 3). K was poorly related to $\log(LFR)$: $K = 3.745 + 0.804 \cdot \log(LFR)$ ($r^2=0.45$; $n=132$; $P<0.001$).

For pairs $U=(LFR, SN)$ included in the experimental range ($0.8 < LFR < 13.4$ and $2 < SN < 27$), $E[\hat{W}_f(U)]$, later called predicted $W_f(U)$, were calculated from equation [VII] and variances $\text{var}[\hat{W}_f(U)]$ were estimated by simulation (Figure 4). The variance $\text{var}[\hat{W}_f(U)]$ was generally high and increased with predicted $W_f(U)$. However, the coefficient of variation was high for the small predicted $W_f(U)$ and low for the large ones. The contribution of the different sources of error to $\text{var}[\hat{W}_f(U)]$ is shown on Table II for four contrasting (LFR, SN) pairs. The variance $\text{var}(B3)$ was very small ($\text{mean}[\text{var}(B3)]=0.002 \text{ g}^2$) and we neglected it. The main contribution was due to $B2$ which depended on the residual error of the regressions between the growth model parameters and the growth factors. This was likely due to the low accuracy of the regression of K on $\log(LFR)$. Both $\text{var}(B1)$ and $\text{var}(\varepsilon_r)$ were low.

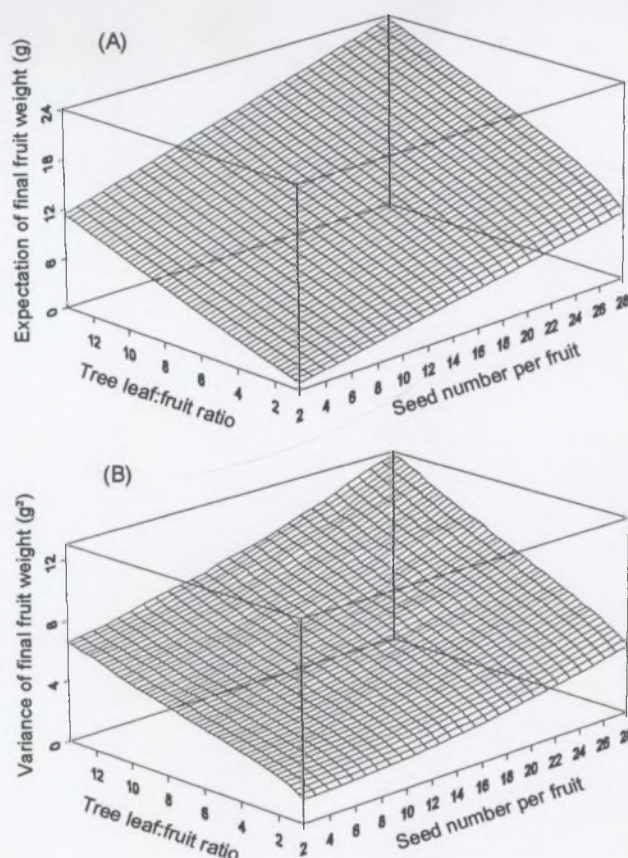


FIG. 4

Expectation (A) and variance (B) of strawberry guava fruit weight at maturity predicted with a nested growth model, with the seed number and the tree leaf:fruit ratio as growth factors.

TABLE II

Predicted strawberry guava fruit weight at maturity ($E[\hat{W}_f(U)]$) and contribution of the different sources of statistical errors (B1, B2, resid) to the variance of the prediction ($\text{var}[\hat{W}_f(U)]$) by a nested growth model (model 1) for four contrasting pairs of seed number and tree leaf:fruit ratio. B1: errors related to the parameters estimation in the regression between growth model parameters and growth factors; B2: residual errors of the regression between growth model parameters and growth factors; resid: residual error of the strawberry guava growth curve fit

Fruit seed number	Tree leaf:fruit ratio	$E[\hat{W}_f(U)]$ (g)	var(B1) (g ²)	var(B2) (g ²)	var(resid) (g ²)	$\text{var}[\hat{W}_f(U)]$ (g ²)
5	2	2.9	0.01	2.56	0.03	2.56
25	2	9.9	0.12	4.99	0.03	5.14
5	10	9.7	0.18	5.40	0.03	5.61
25	10	18.8	0.26	9.41	0.03	9.70

In order to evaluate the cost in accuracy on prediction of the nested model (model 1), we derived a new model from model 1. The development of the product $W_f = K \cdot P$ with the relationships between K , P and the growth factors, and adding a constant term, expressed $W_f(U)$ as a linear combination of SN , LFR , $SN \log(LFR)$ and $LFR \log(LFR)$. After fitting with the data, three parameters were not significantly different from zero: $W_f(U) = 3.180 + 0.314 \cdot SN \log(LFR)$ ($r^2=0.81$, $n=132$, $P<0.001$). This model (model 2) belonged to the same family than model 1, but had not a nested structure. Its parameters were directly fitted on raw data. Although the variability of a particular estimation $\hat{W}_f(U)$ by the linear model 2 and its components can be exactly calculated (Dagnelie, 1998), the previous approach using simulation was used to estimate them (Table III) in order to compare the results with those of model 1 (Table II). We verified that the variability of each of the four predicted values with model 2 (Table III) were similar by simulation or by exact calculation. $\text{var}[\hat{W}_f(U)]$ was rather constant in the range considered and lower than in model 1, except for low SN and LFR . Model 2 had a bad behaviour for $LFR \leq 1$ as then $\log(LFR) \leq 0$. Consequently, predicted $\hat{W}_f(U)$ was constant or decreased as SN increased, which was not consistent with our experimental results, and negative $\hat{W}_f(U)$ could be predicted. Therefore, the validity of model 2 is limited to $LFR > 1$.

TABLE III

Predicted strawberry guava fruit weight at maturity ($E[\hat{W}_f(U)]$) and contribution of the different sources of statistical errors (PE, resid) to the variance of the prediction ($\text{var}[\hat{W}_f(U)]$) by a direct model (model 2) for four contrasting pairs of seed number and tree leaf:fruit ratio.

PE: error related to the model parameters estimation; resid: residual error of the model

Fruit seed number	Tree leaf:fruit ratio	$E[\hat{W}_f(U)]$ (g)	$\text{var}(\text{PE})$ (g ²)	$\text{var}(\text{resid})$ (g ²)	$\text{var}[\hat{W}_f(U)]$ (g ²)
5	2	4.3	0.05	3.58	3.63
25	2	8.6	0.03	3.58	3.61
5	10	6.8	0.03	3.58	3.61
25	10	21.3	0.37	3.58	3.95

For the two levels of validation, SN range was larger than the experimental range where models were fitted. For the first level of validation, on different fruit from the same trees and production cycle, SN ranged from 1 to 42, with $\text{mean} \pm \text{sd} = 10.7 \pm 6.6$, and 2.2% of the fruit had SN outside the experimental range. $\hat{W}_f(U)$ prediction with model 1 was consistent with actual data at the tree (Figure 5A) and at the orchard level (Figure 5C). But low fruit weights were over-estimated in trees with $LFR > 12$ (Figure 5B). This was due to an over-estimation of P when $SN < 12$ and $LFR > 12$. As model was based on statistical relationships, this was likely a consequence of the weak presence of such fruit in the experimental range. Model 2 reproduced satisfactorily the actual $W_f(U)$ distributions at the tree level, whatever the tree LFR , and at the orchard level (data not shown).

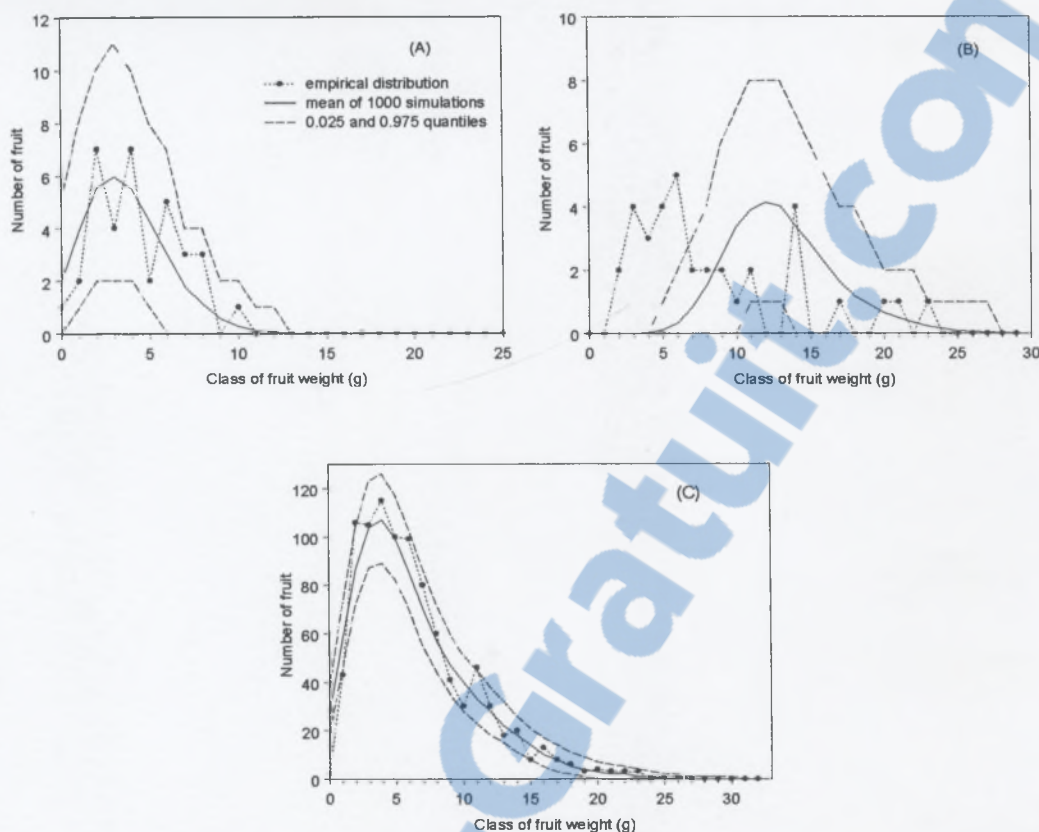


FIG. 5

Validation of a nested strawberry guava fruit growth model on the same trees and production cycle than model fitting, but on different fruit. Empirical and predicted, with 95% prediction region, fruit weight distributions at harvest (1g class), on trees with leaf:fruit ratio of 1.6 (A) and 13.0 (B), and on a set of 27 trees (C) which leaf:fruit ratio varied from 0.8 to 13.4.

For the second level of validation on a different orchard and year, tree LFR ranged from 1.3 to 11.1 (8 trees with $LFR < 2$) and SN ranged from 2 to 43 (mean \pm sd = 19.8 ± 7.7 ; 16.6% of the fruit had SN outside the experimental range). Model 1 over-estimated fruit weights, both at the tree (Figures 6A and B) and at the orchard (Figure 6C) levels, for the same reason than previously and because model 1 tended to over-estimate actual $W_f(U)$ for $LFR < 2$ and $SN > 17$, i.e. 52.4% of the fruit. The range of model 2 synthetic predictor $SN_{log}(LFR)$ was larger than in the experimental range, but weight of fruit outside this range were satisfactorily predicted. Model 2 predictions were more consistent with the actual fruit weights for certain trees (Figure 7A), but still over-estimated fruit weights for most trees (Figure 7B) and at the orchard level (Figure 7C).

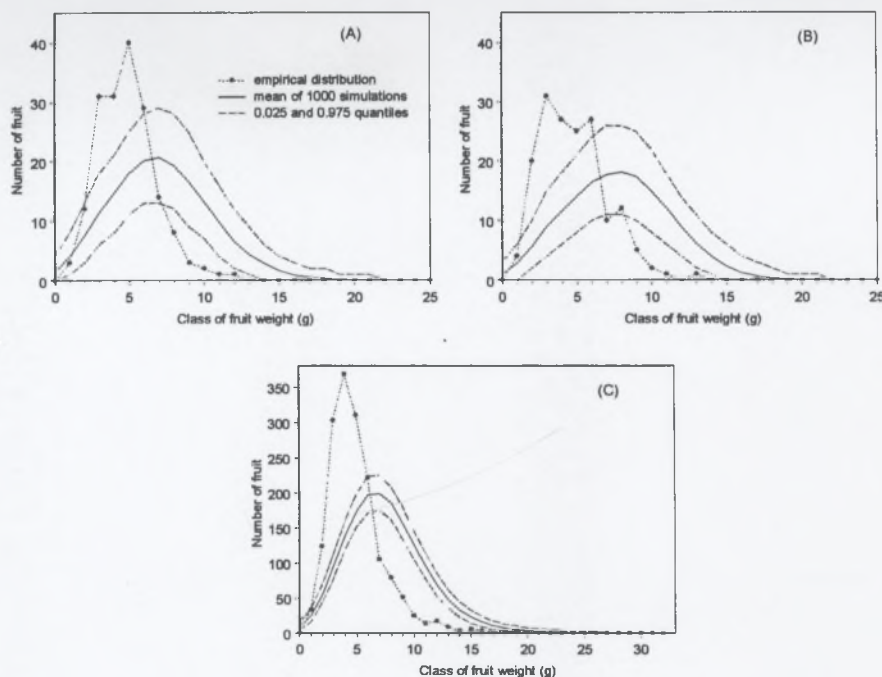


FIG. 6

Validation of a nested strawberry guava fruit growth model on a different orchard and year than model fitting. Empirical and predicted, with 95% prediction region, fruit weight distributions at harvest (1g class), on trees with leaf:fruit ratio of 1.4 (A) and 1.9 (B), and on a set of 11 trees (C) which leaf:fruit ratio varied from 1.3 to 11.1.

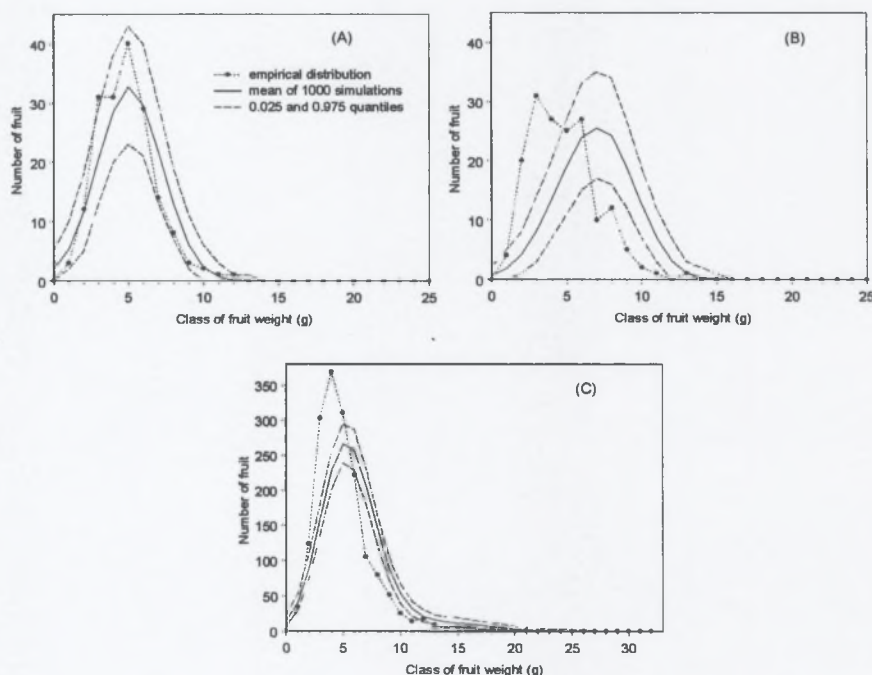


FIG. 7

Validation of a direct strawberry guava fruit growth model on a different orchard and year than model fitting. Empirical and predicted, with 95% prediction region, fruit weight distributions at harvest (1g class), on trees with leaf:fruit ratio of 1.4 (A) and 1.9 (B), and on a set of 11 trees (C) which leaf:fruit ratio varied from 1.3 to 11.1.

DISCUSSION

The strawberry guava fruit has a biphasic growth curve with a final exponential growth before full maturity. Similar growth pattern is reported on highbush blueberry (*Vaccinium corymbosum* L.) (Huang *et al.*, 1997) and on saskatoon (*Amelanchier alnifolia* Nutt.) (McGarry *et al.*, 1998). Silva Galho *et al.* (2000) also observed a biphasic growth curve for strawberry guava fruit in Brasil, but with a final plateau before maturity likely due to the fact that data were not individual fruit weight but mean weight of fruit harvested at each date. The fruit growth curves and final weights were highly variable. The seed number and the tree leaf:fruit ratio explained to a large extent this variability, justifying to study growth at the fruit level within each tree.

The aim of the model is not the prediction of individual fruit weight *per se*, but the prediction of fruit weight distribution at harvest, which is built from individual fruit weight in order to take into consideration the variability between fruit and the prediction variability of each fruit. The same approach could be applied at the growth curve equation in order to predict fruit weight distribution and its variability at any date. This permitted to account for growth factors effects on model predictions, and to precise how each parameter was affected by growth factors. It was assumed that each growth model parameter was related to growth factors. Nevertheless, if no relationship could be evidenced between a parameter and growth factors, either this parameter could be set to a constant value if it was not very variable, either it could be sampled in an empirical distribution of values in order to account for its variability unexplained by the considered growth factors. These relationships accounted for the specific effect of growth factors on the different growth phases, which could be of interest to study biphasic growth pattern as that of strawberry guava. This approach is however limited by the hypothesis that the growth factors are constant during growth. This constraint have few consequences from a practical point of view for strawberry guava. Indeed, reduced crop load effect is larger on fruit weight when trees are thinned at fruit set, which is done by growers. Crop load could then be considered constant. Growing conditions such as water availability must be constant during fruit growth. This approach differs from the widespread multiplicative method used in crop modelling (Whisler *et al.*, 1986). The growth factors effects are then applied on the growth rate or relative growth rate of a reference curve, and the way each factor acts on growth curve parameters is not explicited. However, growth factors effects may then vary during fruit growth. (e.g. late thinning: Lescourret *et al.*, 1999).

The results suggest underlying biological processes and direct further research. The relation $W_f = K \cdot P$ indicated that a potential final fruit weight was set up during the first growth period and was expressed during the second growth period, as observed for peach growing in non limiting conditions (Batjer and Westwood, 1958). This suggests that the first growth period was primarily related to cell division and expansion and the second growth period to cell expansion, as in saskatoon which has a growth pattern similar to that of strawberry guava (McGarry *et al.*, 2001). The growth factors had specific effects on the two growth periods through their relationships with P and K . As observed for several fruit (Ho, 1992), SN and LFR had a direct effect on the first growth period. The second growth period was only dependent of LFR , suggesting a role of carbohydrate availability. As a large variability in K was evidenced between fruit within a tree, the local competition between fruit might affect the second growth period. Further studies are required to precise the physiology of strawberry guava fruit growth.

LFR range on unthinned trees in the second level of validation was within the experimental range, suggesting this latter was representative of most tree *LFR* in orchards. But *SN* distribution differed from experimental range, likely in relation with pollination conditions. The over-estimation of fruit weights by model 1 when $LFR > 12$ in the first level of validation was partly due to fruit which pairs (*LFR*, *SN*) were outside the experimental range where the statistical relation between *P* and growth factors was fitted. A wider and more even distribution of pairs (*LFR*, *SN*) would be needed for a more robust fit of model parameters. This implies to study growth with more fruit per tree, or to control the pollination, in order to have a larger *SN* distribution for each *LFR* to fit model 1. Model 2 validation, which was less sensitive to validation range, showed however that W_f was over-estimated on some trees, suggesting that other(s) factor(s) affected fruit growth, probably at the tree or fruit level, such as for example local competition between fruit as stated above. These factors might have been likely stronger on the trees of the second level of validation, explaining thus the over-estimation of the final weight distribution by both models (Figures 6C and 7C). Further experiments are required to test this hypothesis and improve the models robustness.

The variability of a prediction was estimated using simulation, and two approximations concerning the estimation error on the linear regression parameters were done. Our large sample (132 fruit) makes reasonable these approximations, all the more as the estimation errors on the linear regression parameters have a small contribution on the total variability of the prediction (Tables II and III). We verified moreover with the simple linear regression of model 2 that the variability of a predicted value was not affected by the calculation method (simulation or exact calculation). We supposed that our data set is representative of the actual seed numbers and tree leaf:fruit ratios encountered in orchards, so that the estimated parameters values and the statistical variability of our data set are used for prediction on other data sets. As stated above, this is likely not the case, and weight of fruit outside the experimental range are not well predicted. We assumed that the error on the estimated parameters $\hat{\Omega}_i$ follows a multinormal distribution with a variance-covariance matrix C_i estimated from the data. C_i depends on the residual variance and therefore, for a given data set, on the chosen estimator (here the least square estimator). So the statistical variability which is simulated by the prediction model depends on the estimator and is not the same than the natural variability of the observed phenomenon.

The modelling of statistical errors showed that the model accuracy was limited, as illustrated by a large variability of each prediction. Provided an appropriate growth model and a sufficient number of observations to ensure accuracy on parameters estimation, the main contribution to the prediction variability was due to the residual errors of the regressions between the growth curve model parameters and growth factors. The weight of these residual errors on output variability is the cost, in accuracy, of the possibility to predict fruit weight, by applying the approach at the growth curve equation, at any time during growth and for any factor value included in the validity range of the regressions. Then the highest part of growth model parameters variance explained by the growth factors, the highest the model accuracy. Hence the necessity to identify the main factors affecting growth and to include them in the model. The cost of the nested model on prediction accuracy is illustrated by $\hat{W}_f(U)$ variance with model 2, which is 1.4 to 2.5 times lower than that of model 1, except when both *SN* and *LFR* are low (Table III). Model 2 expressed a direct relationship between final fruit weight and the growth factors, without taking into account the growth curve model. The lower variability on predicted $W_f(U)$ was related to the direct fit of model 2 parameters on data. The residual error of the fit was the major contribution to the prediction variability. Model 2 was

useful from a practical point of view for its simplicity, accuracy and robustness to predict final fruit weight, but it swept away the information on the effect of growth factors on the growth curve given by model 1.

In conclusion, a nested model has proven useful to express the strawberry guava growth as a function of two growth factors, *SN* and *LFR*, and predict the final fruit weight distribution. The variability of each prediction and of the predicted $W_f(U)$ distribution was estimated using simulation. The model accuracy was limited, in relation with a rough statistical relationship between one model parameter and growth factors. The model gave information on the underlying biological processes during fruit growth. A more accurate and robust model was derived from the nested model with a direct parameter estimation, but it did not provide information on the effects of growth factors on the growth pattern.

APPENDIX

The random variable associated to a prediction of the final weight of a strawberry guava fruit characterised by $U = (LFR, SN)$ is given by equation [VI]:

$$\hat{W}_f(U) = (X_K(U)\hat{\Omega}_K^* + X_K(U)\Delta_K + \varepsilon_K) \cdot (X_P(U)\hat{\Omega}_P^* + X_P(U)\Delta_P + \varepsilon_P) + \varepsilon_r$$

where $X_K(U) = (1, \log(LFR))$ and $X_P(U) = (LFR, SN)$ are the vectors of predictors for K and P respectively, $\hat{\Omega}_K^*$ and $\hat{\Omega}_P^*$ are the vectors of parameters estimated from the data. Δ_K and Δ_P are the binormal distributions of parameters estimation errors with mean zero and whose variance-covariance matrix are those of the estimated parameters of the linear regressions of K and P on $X_K(U)$ and $X_P(U)$ respectively; ε_K and ε_P are central normal distributions with variance equal to the residual variance of these regressions ($\text{var}(\varepsilon_K)$ and $\text{var}(\varepsilon_P)$ respectively); ε_r is a central normal distribution with variance equal to the residual variance of the growth curve fit, $\text{var}(\varepsilon_r)$.

$\hat{W}_f(U)$ is not normally distributed as products of normal distributions appear in its expression. We calculate its expectation $E[\hat{W}_f(U)]$ and variance $\text{var}[\hat{W}_f(U)]$. Although K is a scale factor between W_f and P , we assume that the estimators \hat{K} and \hat{P} are independent. Then:

$$[X] \quad E[\hat{W}_f(U)] = (X_K(U)\hat{\Omega}_K^*) \cdot (X_P(U)\hat{\Omega}_P^*)$$

To calculate $\text{var}[\hat{W}_f(U)]$, we write $\hat{W}_f(U) = D + B_1 + B_2 + B_3 + \varepsilon_r$ with:

$$D = (X_K(U)\hat{\Omega}_K^*) \cdot (X_P(U)\hat{\Omega}_P^*)$$

$$B_1 = (X_K(U)\hat{\Omega}_K^*)(X_P(U)\Delta_P) + (X_P(U)\hat{\Omega}_P^*)(X_K(U)\Delta_K) + (X_K(U)\Delta_K)(X_P(U)\Delta_P)$$

$$B_2 = X_K(U)\hat{\Omega}_K^* \varepsilon_P + X_P(U)\hat{\Omega}_P^* \varepsilon_K + \varepsilon_K \varepsilon_P$$

$$B_3 = X_K(U)\Delta_K \varepsilon_P + X_P(U)\Delta_P \varepsilon_K$$

D is the deterministic part of $\hat{W}_f(U)$; B_1 depends only on the estimation errors Δ_K and Δ_P of the linear regression parameters; B_2 depends only on the residual errors ε_K and ε_P ; B_3 depends on the products of the two previous sources of error.

Therefore: $\text{var}[\hat{W}_f(U)] = \text{var}(B_1 + B_2 + B_3) + \text{var}(\varepsilon_r)$, or:

$$\text{var}[\hat{W}_f(U)] = \text{var}(B_1) + \text{var}(B_2) + \text{var}(B_3) + 2 \cdot \text{cov}(B_1, B_2) + 2 \cdot \text{cov}(B_1, B_3) + 2 \cdot \text{cov}(B_2, B_3) + \text{var}(\varepsilon_r)$$

$\text{cov}(B_1, B_2) = \text{cov}(B_1, B_3) = \text{cov}(B_2, B_3) = 0$ as we assumed that the estimators \hat{K} and \hat{P} are independent. Then:

$$[XI] \quad \text{var}[\hat{W}_f(U)] = \text{var}(B_1) + \text{var}(B_2) + \text{var}(B_3) + \text{var}(\varepsilon_r)$$

REFERENCES

- ARNOLD, C. Y. (1959). The determination and significance of the base temperature in a linear heat unit system. *Proceedings of the American Society for Horticultural Science*, **74**, 430–45.
- BATJER, L. P. and WESTWOOD, M. N. (1958). Size of Elberta and J.H. Hale peaches during the thinning period as related to size at harvest. *Proceedings of the American Society for Horticultural Science*, **72**, 102–5.
- BRAMARDI, S. J., CASTRO, H. R. and ZANELLI, M. L. (1998). Fruit growth pattern of pear cv Bartlett and Packham's Triumph to improve hand thinning. *Acta Horticulturae*, **475**, 283–93.
- BURGE, G. K., SPENCE, C. B. and MARSHALL, R. R. (1987). Kiwifruit: effects of thinning on fruit size, vegetative growth, and return bloom. *New Zealand Journal of Experimental Agriculture*, **15**, 317–24.
- BUWALDA, J. G. (1991). A mathematical model of carbon acquisition and utilisation by kiwifruit vines. *Ecological Modelling*, **57**, 43–64.
- COOMBE, B. G. (1976). The development of fleshy fruits. *Annual Review of Plant Physiology*, **27**, 507–28.
- DAGNELIE, P. (1975). *Théorie et méthodes statistiques. Applications agronomiques. Volume 2*. Les Presses Agronomiques de Gembloux, Gembloux, Belgique.
- DAGNELIE, P. (1998). *Statistique théorique et appliquée. Tome 2, inférence statistique à une et deux dimensions*. De Boeck & Larcier s.a., Bruxelles, Belgique.
- FANIZZA, G. and COLONNA, G. (1996). A mathematical function to study berry growth variability in grapes (*Vitis vinifera*). *Journal of Wine Research*, **7**, 117–23.
- FISHMAN, S. and GENARD, M. (1998). A biophysical model of fruit growth: simulation of seasonal and diurnal dynamics of mass. *Plant, Cell and Environment*, **21**, 739–52.
- FRANCE, J. and THORNLEY, J. H. M. (1984). *Mathematical models in agriculture*. Butterworths, London, United Kingdom.
- GALLEGO, P. P., MARTINEZ, A. and ZARRA, I. (1997). Analysis of the growth kinetic of fruits of *Actinidia deliciosa*. *Biologia Plantarum*, **39**, 615–22.
- GANDAR, P. W., HALL, A. J. and DE SILVA, H. N. (1996). Deterministic models for fruit growth. *Acta Horticulturae*, **416**, 103–12.
- GENARD, M. and BRUCHOU, C. (1993). A functional and exploratory approach to studying growth: the example of the peach fruit. *Journal of the American Society for Horticultural Science*, **118**, 317–23.
- GEORGE, A. P., NISSEN, R. J. and KO, H. L. (1986). Productivity of custard apple (*Annona atemoya* Hort.): factors affecting yield and fruit size. *Acta Horticulturae*, **175**, 37–41.
- GROSSMAN, Y. L. and DEJONG, T. (1994). PEACH: a simulation model of reproductive and vegetative growth in peach trees. *Tree Physiology*, **14**, 329–45.
- HANSEN, P. (1989). Source/sink effects in fruits: an evaluation of various elements. In: *Manipulation of fruiting*. (Wright, C.J., Ed). Butterworths, London, United Kingdom, 29–37.
- HO, L. C. (1992). Fruit growth and sink strength. In: *Fruit and seed production. Aspects of development, environmental physiology and ecology*. (Marshall, C. and Grace, J., Eds). Cambridge Press University, Cambridge, United Kingdom, 101–24.
- HUANG, Y. H., JOHNSON, C. E., LANG, G. A. and SUNDBERG, M. D. (1997). Pollen sources influence early fruit growth of southern highbush blueberry. *Journal of the American Society for Horticultural Science*, **122**, 625–9.

- INGLESE, P., BARBERA, G., LA MANTIA, T. and PORTOLANO, S. (1995). Crop production, growth, and ultimate size of cactus pear fruit following fruit thinning. *HortScience*, **30**, 227–30.
- LAKSO, A. N., CORELLI GRAPPADELLI, L., BARNARD, J. and GOFFINET, M. C. (1995). An expolinear model of the growth pattern of the apple fruit. *Journal of Horticultural Science*, **70**, 389–94.
- LESCOURRET, F., BLECHER, N., HABIB, R., CHADŒUF, J., AGOSTINI, D., PAILLY, O., VAISSIERE, B. and POGGI, I. (1999). Development of a simulation model for studying kiwi fruit orchard management. *Agricultural Systems*, **59**, 215–39.
- MATHSOFT (1999). *S-Plus 2000 guide to statistics, Volume 1*. Data Analysis Products Division, Mathsoft, Seattle, USA.
- MCGARRY, R., OZGA, J. A. and REINECKE, D. M. (1998). Patterns of saskatoon (*Amelanchier alnifolia* Nutt.) fruit and seed growth. *Journal of the American Society for Horticultural Science*, **123**, 26–9.
- MCGARRY, R., OZGA, J. A., REINECKE, D. M. (2001). Differences in fruit development among large- and small-fruited cultivars of saskatoon (*Amelanchier alnifolia*). *Journal of the American Society for Horticultural Science*, **126**, 381–5.
- NORMAND, F. (1994). Le goyavier-fraise, son intérêt pour l'île de la Réunion. *Fruits*, **49**, 217–27.
- PYKE, N. B. and ALSPACH, P. A. (1986). Inter-relationships of fruit weight, seed number and seed weight in kiwifruit. *New Zealand Agricultural Science*, **20**, 153–6.
- RASEIRA, M. C. B. and RASEIRA, A. (1996). *Contribuição ao estudo do araçazeiro*, *Psidium cattleianum*. EMBRAPA-CPACT, Pelotas, Brazil.
- REDDY, Y. T. N. and SINGH, G. (1991). Further studies on the relationship between leaf number and area and fruit development in mango (*Mangifera indica* L.). *Journal of Horticultural Science*, **66**, 471–8.
- SILVA GALHO, A., FERNANDES LOPES, N., RASEIRA, A. and BACARIN, M. A. (2000). Crescimento do fruto de araçá (*Psidium cattleianum* Sabine). *Revista Brasileira de Fruticultura (Jaboticabal)*, **22**, 223–5.
- STANLEY, C. J., TUSTIN, D. S., LUPTON, G. B., MCARTNEY, S., CASHMORE, W. M. and DE SILVA, H. N. (2000). Towards understanding the role of temperature in apple fruit growth responses in three geographical regions within New Zealand. *Journal of Horticultural Science & Biotechnology*, **75**, 413–22.
- WALLACH, D. and GENARD, M. (1998). Effect of uncertainty in input and parameter values on model prediction error. *Ecological Modelling*, **105**, 337–45.
- WELTE, H. F. (1990). Forecasting harvest fruit size during the growing season. *Acta Horticulturae*, **276**, 275–82.
- WIELGOLASKI, F. E. (1999). Starting dates and basic temperatures in phenological observations of plants. *International Journal of Biometeorology*, **42**, 158–68.
- WHISLER, F. D., ACOCK, B., BAKER, D. N., FYE, R. E., HODGES, H. F., LAMBERT, J. R., LEMMON, H. E., MCKINION, J. M. and REDDY, V. R. (1986). Crop simulation models in agronomic systems. *Advances in Agronomy*, **40**, 141–208.
- YANG, S., LOGAN, J. and COFFEY, D. L. (1995). Mathematical formulae for calculating the base temperature for growing degree days. *Agricultural and Forest Meteorology*, **74**, 61–74.

II- Synthèse de la partie précédente et conclusion pour le modèle de prédiction de la distribution des poids de fruit à maturité.

La prédiction précoce de la distribution des poids de fruit à maturité est un indicateur décisif pour préparer en conséquence la récolte et la commercialisation, et décider d'opérations techniques telles que l'éclaircissage. Le poids du fruit à maturité dépend de nombreux facteurs liés au fruit, à la plante, ou environnementaux, qui conditionnent par leurs effets sur la croissance du fruit la distribution des calibres à maturité sur une parcelle. L'approche proposée pour prendre en compte ces effets est d'exprimer les paramètres de la courbe de croissance du fruit en fonction de ces facteurs. Le modèle de croissance s'exprime alors sous forme de modèles emboîtés. Une procédure d'estimation des paramètres est proposée. Afin d'évaluer la précision sur les sorties du modèle, un modèle d'erreurs est intégré. Il peut permettre de plus d'estimer la contribution de chaque source d'erreur à l'erreur totale. La variabilité sur les sorties du modèle est estimée par simulation.

Cette approche est appliquée à la construction d'un modèle de croissance du fruit du goyavier-fraise, avec comme facteurs de croissance considérés le nombre de graines du fruit (SN) et le rapport feuilles:fruits du plant (LFR). La courbe de croissance du goyavier-fraise est biphasique et est modélisée comme la somme d'une fonction logistique et d'une fonction exponentielle. Du fait de la croissance exponentielle finale du fruit et de la faible précision sur la prédiction de la date de récolte d_{rec} , cette dernière et le poids du fruit à maturité W_f doivent être modélisés séparément, indépendamment de la courbe de croissance. La démarche précédente est alors appliquée pour construire un modèle de prédiction de W_f . Cette variable peut être écrite comme le produit du poids du fruit P à la fin de la première phase de croissance par le rapport K entre W_f et P : $W_f = K \cdot P$. Les paramètres K et P s'expriment en fonction des facteurs de croissance étudiés. La première phase de croissance dépend du nombre de graines du fruit (SN) et de la disponibilité moyenne en assimilats (LFR). La seconde phase de croissance, avant la maturité, ne dépend que de LFR .

Ces relations donnent des informations sur l'effet de ces facteurs sur chaque phase de la courbe de croissance. La précision sur les sorties du modèle est moyenne, à cause de l'erreur résiduelle élevée de la relation entre le paramètre K et sa variable explicative. La validation du modèle sur d'autres fruits des mêmes plants et du même cycle de production que les données ayant servi à ajuster le modèle est bonne. Il surestime par contre la distribution des poids à maturité pour des fruits d'une autre parcelle.

Un second modèle est dérivé en développant analytiquement le rapport $W_f = K \cdot P$ et en ajustant ce modèle directement aux données. Le poids W_f s'exprime alors comme une fonction linéaire d'un prédicteur synthétique $SN \log(LFR)$. La précision sur les sorties est meilleure que celle du modèle précédent. Mais ce modèle ne spécifie pas les effets des facteurs sur les phases de croissance du goyavier-fraise. Il se comporte bien en validation sur d'autres fruits des mêmes plants et du même cycle de production que les données ayant servi à l'ajustement. Il prédit correctement les distributions de poids de fruits de certains plants d'une autre parcelle, mais tend à surestimer globalement les poids de fruit sur cette autre parcelle. Ce résultat laisse supposer qu'un facteur non considéré dans cette étude, agissant à l'échelle du fruit ou du plant, affecte la croissance du fruit. Ce pourrait être la compétition locale entre les fruits. Cette étude met en évidence l'importance du domaine expérimental pour déterminer le domaine de validité de modèles statistiques.

Dans le cadre de notre étude sur l'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise, le modèle le plus simple, qui est aussi le plus précis et le plus robuste, est retenu. Il relie le poids du fruit à maturité à un prédicteur synthétique $SN\log(LFR)$. Ce modèle linéaire peut être amélioré par l'ajustement de ses deux paramètres sur un domaine expérimental plus large et plus équilibré, et par la prise en compte d'autres facteurs de croissance évoqués ci-dessus. Un modèle de prédiction de la distribution des dates de récolte reste à construire.

III- Modèle de prédiction de la distribution des dates de récolte

1- Introduction

La sortie pratique d'un modèle de croissance du fruit, d'un point de vue temporel, est la distribution des dates individuelles de récolte des fruits à l'échelle de la parcelle, qui permet d'organiser le chantier de récolte et la commercialisation. La date de récolte d'un fruit est avant tout déterminée par la température au cours de sa croissance, qui est le plus souvent prise en considération par une échelle de temps thermique accumulé à partir d'une certaine date (Munoz *et al.*, 1986; Fleisch and Bartholomew, 1987; Bittenbender and Kobayashi, 1990). Le fort effet de la température sur la durée de développement du fruit de goyavier-fraise a déjà été montré (Chapitre II, Annexe 1).

Quelques travaux montrent que la date de récolte d'un fruit peut aussi être affectée par d'autres facteurs, dont le principal est la charge en fruits ou le rapport feuilles:fruits. Jones (1931) note chez la pêche qu'une augmentation du rapport feuilles:fruits sur des rameaux annelés tend à accélérer le développement du fruit et sa maturité. Inglese *et al.* (1995) montrent chez le figuier de Barbarie (*Opuntia ficus-indica* Mill.) que les fruits sur les cladodes les moins chargés sont plus précoces que ceux sur les cladodes très chargés. Barone *et al.* (1994) montrent que la récolte sur des oliviers dont la charge en fruits est réduite de moitié est plus précoce d'un mois par rapport à celle de plants non éclaircis. D'autres facteurs sont également cités: Volz *et al.* (1995) observent que les pommes var. Royal Gala situées sur des rameaux vigoureux ou bien éclairés, ou qui ont un nombre de graines élevé, ont des dates de récolte plus précoces. Les fruits d'airelles à corymbe (*Vaccinium corymbosum* L. et *V. corymbosum* L. x *V. angustifolium* Ait.) issus de pollinisation croisée ont plus de graines et arrivent à maturité plus tôt que ceux issus d'autopollinisation (Lang et Danka, 1991; Harrison *et al.*, 1994). Dans ces études, la contribution de l'origine du pollen et du nombre de graines dans la précocité n'est cependant pas définie. Ehlenfeldt (2001) montre qu'au sein de chaque type de pollinisation, les fruits ayant plus de graines ont tendance à mûrir plus tôt.

Ces résultats nous amènent à faire l'hypothèse que le rapport feuilles:fruits du plant (LFR) et le nombre de graines du fruit (SN) peuvent affecter la date de récolte du goyavier-fraise, qui correspond à la chute spontanée du fruit à maturité. Elle doit alors être estimée à l'échelle du fruit, et la distribution des dates individuelles de récolte à l'échelle de la parcelle est ensuite obtenue par agrégation. Une relation linéaire de qualité moyenne a été établie entre les dates de récolte (d_{rec}) de fruits dont la croissance a été suivie (voir Matériels et Méthodes ci-dessous) et ces prédicteurs: $d_{rec} = 2933.8 + 526.5/LFR + 4.9 \cdot SN$; $R^2=0.63$; $n=132$; $P<0.001$. L'effet de SN est significatif ($P=0.02$), mais faible, et cette variable n'apporte pas une forte contribution. La retirer de la régression affecte peu la qualité de cette dernière, et $1/LFR$ seul explique alors 61% de la variabilité de d_{rec} . Le faible effet de SN est illustré en observant un

large échantillon de fruits par plant. Une très grande variabilité de d_{rec} est alors observée pour un même SN (Fig. 11).

Il apparaît donc que si LFR rend bien compte de l'effet du plant sur d_{rec} , SN n'est pas un prédicteur pertinent pour expliquer et estimer de façon satisfaisante la date de récolte individuelle du fruit. Ces deux prédicteurs sont insuffisants pour une prévision satisfaisante de la date individuelle de récolte du fruit, ce qui justifie la modélisation de la date de récolte et du poids du fruit à maturité indépendamment de la courbe de croissance. La recherche de facteurs explicatifs de la variabilité de la date de récolte individuelle du fruit implique des expérimentations complémentaires. Ces résultats apporteraient sans doute une meilleure connaissance des mécanismes en jeu dans la croissance du fruit. En terme de modélisation, des facteurs de nature architecturale ou physiologique devront probablement être considérés, et ils sortent du cadre de notre démarche. Nous proposons donc de déterminer la distribution des dates individuelles de récolte des fruits directement à l'échelle du plant, en considérant comme facteurs LFR et le nombre moyen de graines dans les fruits du plant (SN_{moy}).

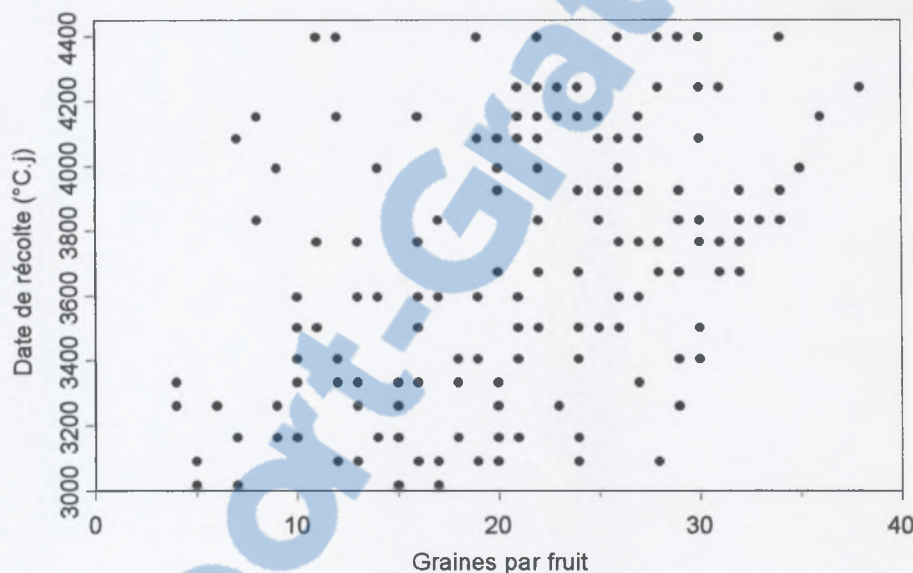


Figure 11: Distribution des dates individuelles de récolte de 165 fruits d'un plant ($LFR = 1.9$) en fonction de leur nombre de graines. Les dates sont exprimées en degrés.jours accumulés depuis la demi-floraison de la parcelle.

2- Matériels et méthodes

Le modèle de prédiction de la distribution des dates de récolte à l'échelle du plant a été construit à partir des observations réalisées sur les essais présentés dans la partie I de ce chapitre, et qui sont rappelées ici brièvement pour ce qui concerne les dates de récolte. Sur un premier essai, noté PR, (480 m d'altitude, récolte naturelle 1998), une large gamme de LFR ([0.8; 13.4]) a été mise en place sur 27 plants à la nouaison (9 traitements x 3 répétitions). La croissance individuelle de 5 fruits par plant a été suivie de la nouaison jusqu'à la pleine maturité (données utilisées pour établir la relation précédente entre d_{rec} et $1/LFR$ et SN). Pour chaque plant, on dispose de la distribution des dates de récolte de l'ensemble des fruits et d'une estimation de la distribution du nombre de graines par fruit (40 fruits/plant). Sur un

second essai, noté JBG et réalisé sur une autre parcelle et une autre année (520 m d'altitude, récolte naturelle 1999), on dispose, pour 11 plants dont le *LFR* a été mesuré à la nouaison ([1.3; 11.1]), de la distribution des dates de récolte de tous les fruits et d'une distribution estimée du nombre de graines dans les fruits (92 à 185 fruits/plant). Pour 10 de ces plants, le déroulement individuel de la floraison a été estimé à partir du suivi d'un échantillon de fleurs (plants ayant servi au paramétrage du modèle de floraison, Chapitre III, partie II). Sur les 2 parcelles, les températures moyennes et les pluviométries entre la date de demi-floraison et la date de demi-récolte des plants suivis ont été similaires: 21.6 et 21.0°C, 3215 et 3691 mm pour PR et JBG respectivement.

Les distributions des dates de récolte par plant sont souvent exagérément étalées par quelques fruits très précoces ou très tardifs dont l'importance dans la récolte totale est faible, mais dont le poids dans la distribution peut être élevé. On a donc travaillé sur des distributions réduites par plant contenant 98% des observations, obtenues par retrait de 1% des dates de récolte en début et en fin de distribution. Cela a permis d'éliminer ces fruits très précoces ou très tardifs.

Le facteur température est pris en compte par l'utilisation d'une échelle de temps thermique (Annexe 1). Toutes les durées et les dates sont exprimées en degrés.jours (°C.j) accumulés à partir de la demi-floraison de la parcelle (essai PR) ou de l'ensemble des 11 plants suivis (essai JBG), sauf indication contraire dans le texte.

Les distributions des dates de récolte par plant sont en général symétriques et de forme à peu près gaussienne. On choisit de les représenter par une loi normale, ce qui revient à les caractériser par la moyenne et l'écart-type de cette loi, correspondant à la date de demi-récolte du plant et à un facteur d'étalement de la distribution autour de cette date. Pour chaque plant, on choisit comme moyenne la médiane de la distribution observée ($d_{0.5rec}$), et comme écart-type, celui de la distribution observée (σ_{rec}). La démarche consiste alors à rechercher, comme pour le modèle de croissance du fruit (Chapitre V-I), des relations linéaires entre $d_{0.5rec}$ d'une part, et σ_{rec} d'autre part, et des prédicteurs fonctions de *LFR* et SN_{moy} . Comme pour le modèle de croissance du fruit, les erreurs statistiques sont considérées dans le modèle de prédiction afin d'estimer la variabilité des prédictions. Pour tester les régressions linéaires sur les données ayant servi à les paramétrer, les limites de prédiction individuelles sont calculées par simulation (100 000 tirages), au lieu d'un calcul exact de l'intervalle de confiance (Dagnélie, 1998b).

La gamme expérimentale de *LFR* des deux essais est large. Dans l'essai PR, la gamme expérimentale de SN_{moy} varie de 8.2 à 14.4. Dans l'essai JBG, elle varie de 11.3 à 22.8. Construire le modèle avec les données de l'essai PR et le valider avec les données de l'essai JBG, comme cela a été fait pour le modèle de croissance du fruit, revient à travailler avec une gamme expérimentale de SN_{moy} réduite et peu représentative de la gamme relevée dans des suivis au cours de plusieurs années ([9.1; 23.4]). Une relation statistique estimée sur ces données aura un domaine de validité et un intérêt limités.

Les données relatives aux plants des deux essais PR et JBG, soit 38 plants, ont donc été rassemblées afin i) de travailler avec une variabilité de SN_{moy} plus grande pour construire le modèle, et ii) d'établir des relations dans un domaine expérimental plus large et représentatif. Cette démarche implique l'hypothèse qu'il n'y a pas d'autre facteur, en particulier relatif à la parcelle et aux conditions de croissance, qui affectent de façon majeure les variables étudiées. L'utilisation des deux jeux de données pour établir le modèle nous empêche de réaliser une validation externe par manque de données indépendantes.

Les sorties du modèle ont été confrontées aux distributions de dates individuelles de récolte observées sur l'ensemble des plants suivis sur chaque parcelle. Pour la validation, les dates de récolte pour les prédictions sont les mêmes que celles observées. A chaque simulation, les paramètres de la distribution normale des dates individuelles de récolte sont tirés dans une variable dépendant des caractéristiques LFR et SN_{moy} du plant. Puis les dates de récolte des fruits du plant sont tirées aléatoirement dans cette loi estimée. Enfin, les dates de récolte ainsi estimées pour tous les fruits de tous les plants sont agrégées en distribution de dates à l'échelle de la parcelle. Mille simulations sont ainsi réalisées, donnant à chaque fois une distribution simulée. Comme pour la validation du modèle de croissance du fruit, la moyenne et les quantiles 0.025 et 0.975 des 1000 fréquences absolues simulées pour chaque date de récolte sont déterminés. Ces quantiles définissent les limites de prédiction à 95% du modèle.

3- Résultats et discussion

La récolte de tous les fruits du goyavier-fraise est très étalée sur chaque plant: 669 à 2151 °C.j sur l'essai PR, soit 25 à 84 j, et 1018 à 2378 °C.j sur l'essai JBG, soit 41 à 101 j. A l'échelle du groupe de plants, la récolte est étalée sur 2651 °C.j (102 j) sur l'essai PR, et 2404 °C.j (102 j) sur l'essai JBG. La récolte est significativement plus tardive et plus étalée sur les plants dont LFR est faible (Tableau 9, Fig. 12). Les différences sont particulièrement importantes pour les petites valeurs de LFR . A l'échelle du fruit, ce retard dans la récolte est lié à un allongement de la période de croissance faible ou nulle précédant la croissance exponentielle finale (Fig. 13).

Charge en fruits	Nombre de feuilles	LFR moyen	$d_{0.5rec}$ (°C.j)	σ_{rec} (°C.j)
Forte	Faible	0.8	3551 c	307.1 c
Forte	Moyen	1.5	3150 ab	222.1 b
Forte	Elevé	2.3	3053 ab	168.5 ab
Moyenne	Faible	1.5	3261 b	187.2 ab
Moyenne	Moyen	2.7	3096 ab	166.2 ab
Moyenne	Elevé	4.0	3060 ab	150.8 ab
Faible	Faible	4.3	3170 ab	160.5 ab
Faible	Moyen	6.9	3070 ab	164.0 ab
Faible	Elevé	13.0	2933 a	139.5 a

Tableau 9: Effet du rapport feuilles:fruits du plant (LFR), obtenu par différents niveaux de la charge en fruits et du nombre de feuilles, sur les paramètres de la loi normale $N(d_{0.5rec}, \sigma_{rec})$ représentant la distribution des dates de récolte des fruits du plant. $d_{0.5rec}$ et σ_{rec} sont exprimés en degrés.jours accumulés depuis la demi-floraison de la parcelle. Pour chaque variable, les moyennes suivies de lettres différentes sont significativement différentes ($n=3$, $P<0.001$)

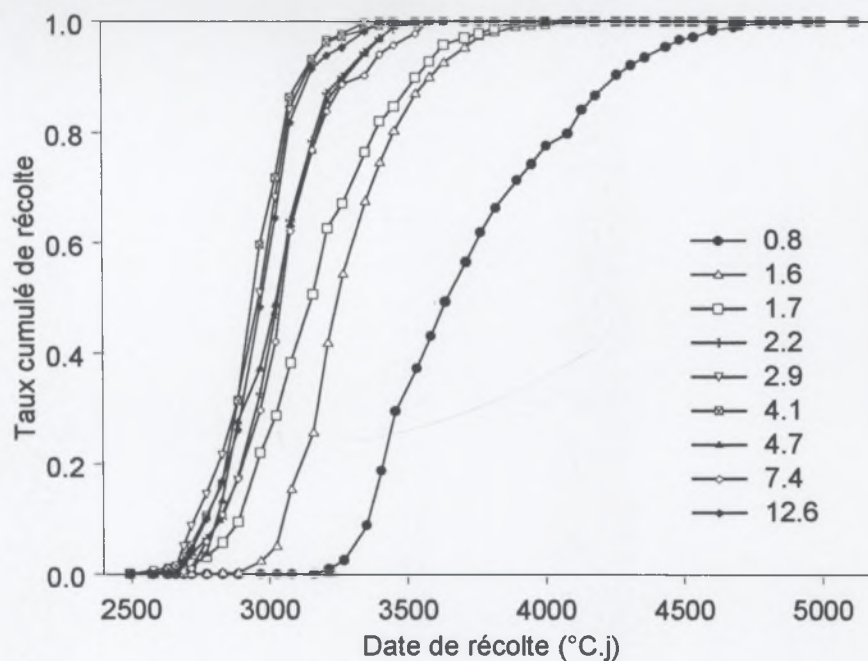


Figure 12: Evolution cumulée de la totalité de la récolte de 9 plants dont le rapport feuilles:fruits varie de 0.8 à 12.6 (essai PR, 9 plants du bloc 1).

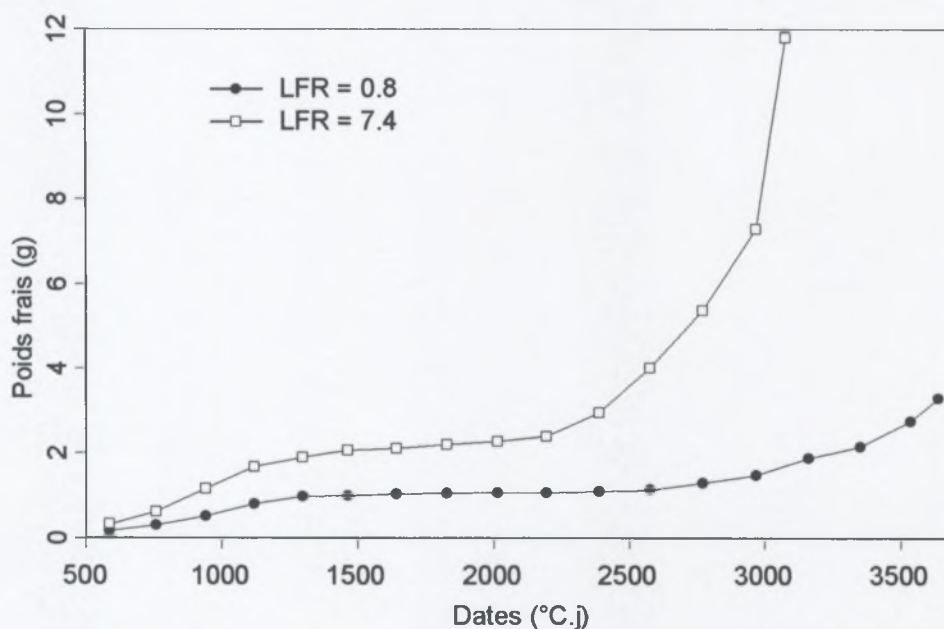


Figure 13: Croissance cumulée en poids:fruits de deux fruits contenant chacun 17 graines et provenant de plants ayant des rapports feuilles:fruits (LFR) différents.

De bonnes relations linéaires existent entre $d_{0.5rec}$ et σ_{rec} , et LFR et SN_{moy} :

$$\begin{aligned} d_{0.5rec} &= 2551.5 + 389.7/LFR + 39.3 \cdot SN_{moy} & R^2=0.81; n=38; P<0.001 \\ \sigma_{rec} &= 134.4/LFR + 11.6 \cdot SN_{moy} & R^2=0.97; n=38; P<0.001 \end{aligned}$$

Ces relations, établies avec des données obtenues dans des conditions différentes, montrent le rôle majeur de LFR et SN_{moy} comme facteurs affectant $d_{0.5rec}$ et σ_{rec} à l'échelle du plant. Cependant, SN_{moy} n'a un effet significatif sur σ_{rec} que lorsque les valeurs de LFR sont inférieures à 2.2 (Fig. 14). L'effet de SN_{moy} sur $d_{0.5rec}$ en fonction de LFR , bien que moins net, présente la même tendance (Fig. 15). La validation des relations précédentes sur les données ayant servi à les établir est satisfaisante (Fig. 16 et 17). Les limites de prédiction de chaque point contiennent la première bissectrice, sauf pour une observation pour $d_{0.5rec}$ et deux observations pour σ_{rec} .

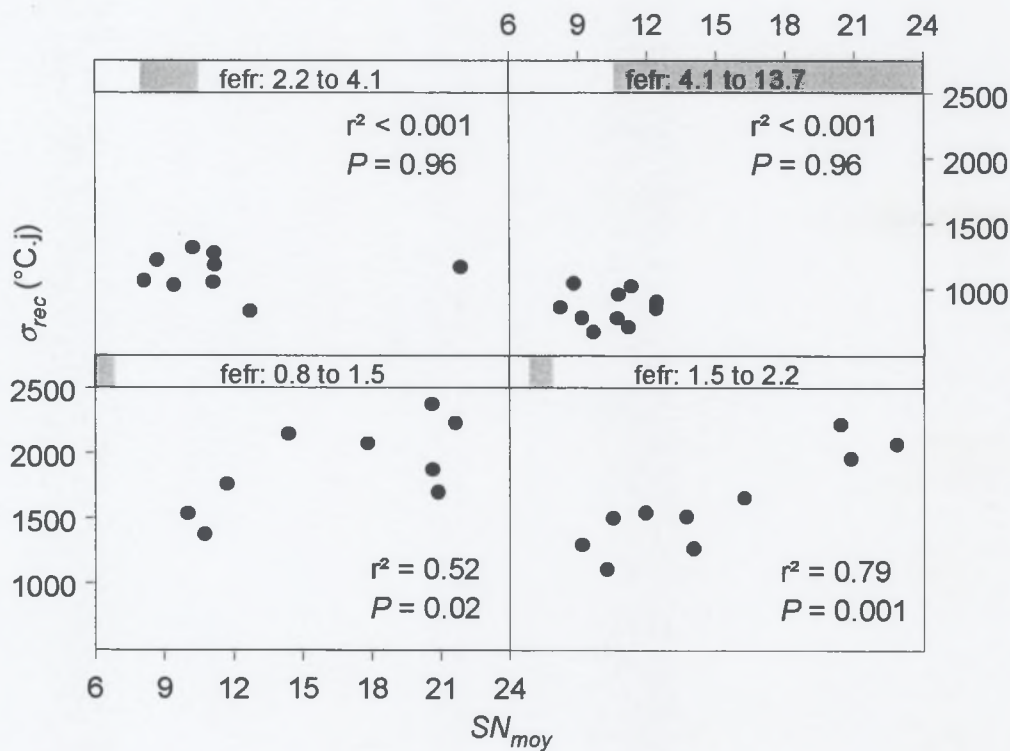


Figure 14: Relation entre l'étalement de la récolte (σ_{rec}) et le nombre moyen de graines par fruit (SN_{moy}) en fonction du rapport feuilles:fruits du plant (LFR) sur 38 goyaviers-fraises provenant de 2 parcelles différentes.

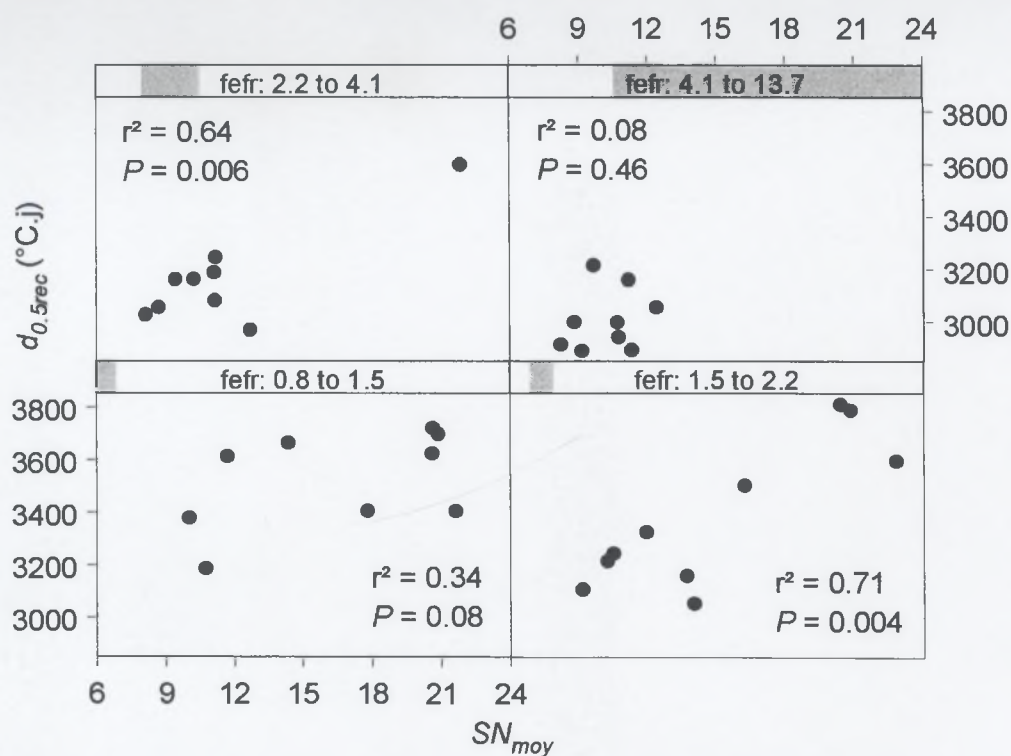


Figure 15: Relation entre la date de demi-récolte ($d_{0.5rec}$) et le nombre moyen de graines par fruit (SN_{moy}) en fonction du rapport feuilles:fruits du plant (LFR) sur 38 goyaviers-fraises provenant de 2 parcelles différentes.

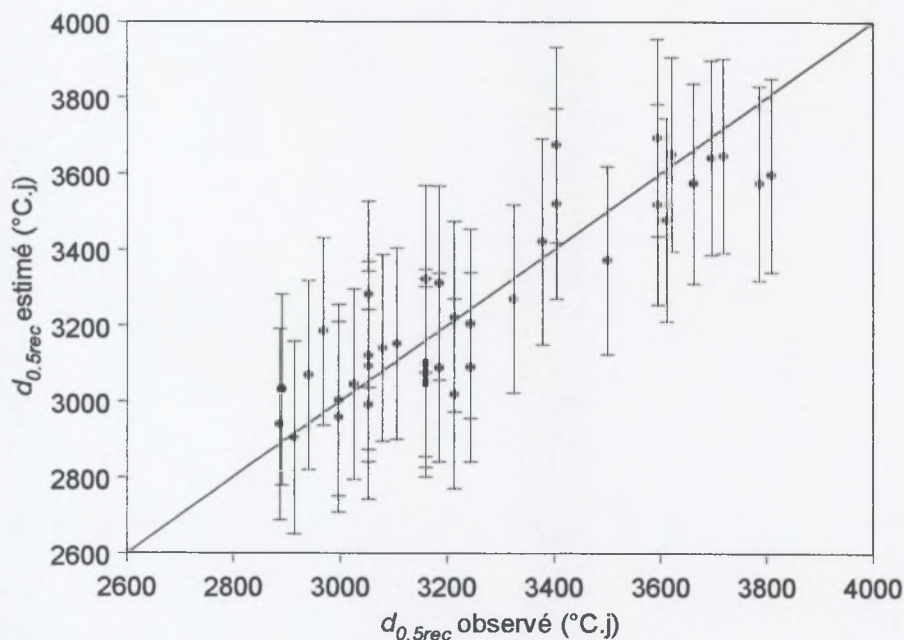


Figure 16: Relation entre les moyennes de la loi normale représentant la distribution des dates de récolte des fruits d'un plant ($d_{0.5rec}$) observées et estimées en fonction du rapport feuilles:fruits du plant et du nombre moyen de graines dans ses fruits. Les barres verticales représentent les limites de prédiction à 95% de chaque point. La ligne représente la bissectrice du quadrant.

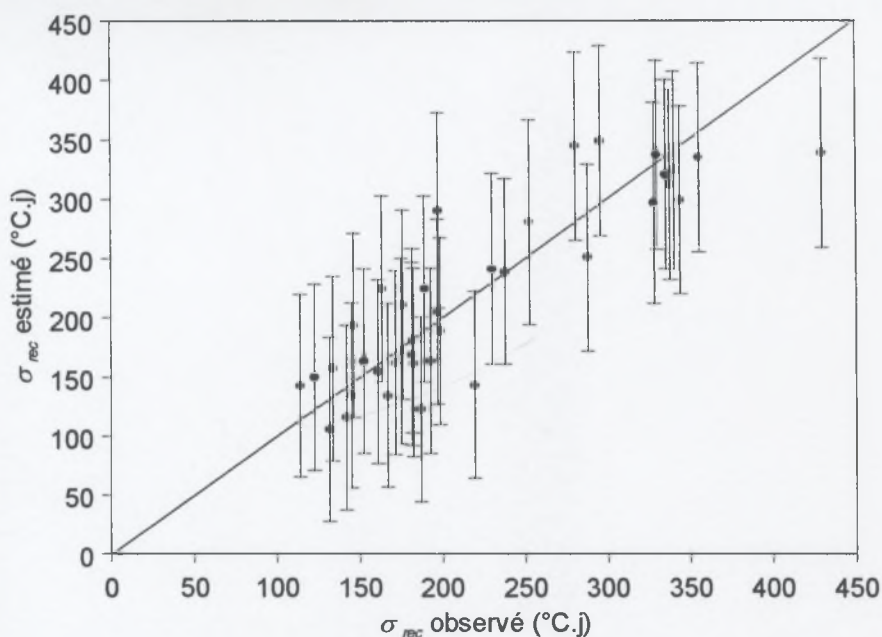


Figure 17: Relation entre les écart-types de la loi normale représentant la distribution des dates de récolte des fruits d'un plant (σ_{rec}) observées et estimées en fonction du rapport feuilles:fruits du plant et du nombre moyen de graines dans ses fruits. Les barres verticales représentent les limites de prédiction à 95% de chaque point. La ligne représente la bissectrice du quadrant.

Ces résultats sont en accord avec les travaux précédents pour ce qui concerne l'effet de *LFR* sur la date de récolte (Jones, 1931; Barone *et al.*, 1994; Inglese *et al.*, 1995). La relation hyperbolique avec *LFR* traduit l'effet plus marqué des petites valeurs de *LFR* sur $d_{0.5rec}$ et σ_{rec} , c'est à dire lorsque la disponibilité moyenne en assimilats par fruit est réduite, mais aussi la plus variable en valeur relative au sein du plant à cause des différences locales de rapport feuilles:fruits sur des plants très chargés. D'où probablement l'effet à la fois sur la date de demi-récolte, mais aussi sur l'étalement de la récolte. La période de faible croissance du fruit, dont la durée est affectée par *LFR* (Fig. 13), correspond à la maturation des graines avant la croissance finale liée au remplissage de la pulpe (données non présentées). Les effets de *LFR* et SN_{moy} pourraient être liés à une maturation des graines d'autant plus lente que la disponibilité globale en assimilats est faible (*LFR*: la source) et qu'il y a en moyenne un nombre élevé de graines par fruit (SN_{moy} : les puits). Cette hypothèse est confortée par l'effet différencié de SN_{moy} sur σ_{rec} , et dans une moindre mesure $d_{0.5rec}$ en fonction de *LFR* (Fig. 14 et 15). Les travaux montrant un effet positif du nombre de graines sur la précocité chez certaines espèces ont probablement été réalisés sur des plants en conditions non limitantes en assimilats (Ehlenfeldt, 2001). Dans de telles conditions chez le goyavier-fraise (*LFR* > 2.2), le nombre moyen de graines ne semble pas avoir d'effet sur la précocité du fruit.

Relativement aux valeurs observées, les résidus de la relation entre σ_{rec} et les prédictors $1/LFR$ et SN_{moy} sont plus élevés que ceux de la relation établie avec $d_{0.5rec}$, ce qui traduit une moins bonne explication de σ_{rec} par ces prédictors et l'influence probable d'autre(s) facteur(s) non pris en compte dans notre étude. Sur les 10 plants de l'essai JBG dont la floraison est connue, il n'a pas été mis en évidence de relation entre l'étalement de la floraison (en °C.j) sur

un échantillon de fleurs suivies et σ_{rec} sur les plants ($r^2=0.11$; $P=0.34$). Cela suggère que l'étalement de la floraison sur le plant est peu explicatif de la variabilité de σ_{rec} . Cette variabilité pourrait se construire à l'échelle du fruit ou du rameau sous l'influence de facteurs locaux comme le rapport feuilles:fruits du rameau (Jones, 1931; Inglese *et al.*, 1995), la vigueur du rameau ou son ensoleillement (Volz *et al.*, 1995). L'influence de la date individuelle de floraison sur cette variabilité n'est pas toujours claire comme le montrent les études réalisées sur la pomme (Sullivan, 1965; Volz *et al.*, 1995). La forte variabilité des dates individuelles de récolte des fruits de goyavier-fraise pour un même couple de valeurs (LFR , SN) (Fig. 11) pourrait aussi être expliquée par ces facteurs locaux de variabilité, et l'hypothèse de l'influence de la date individuelle de floraison est à vérifier chez le goyavier-fraise.

Par contre, une partie de la variabilité de $d_{0.5rec}$ est expliquée par la date de demi-floraison du plant. Les floraisons des plants sont asynchrones, mais l'on utilise la date de demi-floraison de l'ensemble des plants comme point de départ de l'accumulation des degrés.jours pour la croissance du fruit. Pour les 10 plants de l'essai JBG dont la floraison est connue, on a recalculé les $d_{0.5rec}$ à partir des dates de demi-floraison de chaque plant, et une réduction importante de la variance sur les $d_{0.5rec}$ est observée, passant de 20746 °C².j² à 15006 °C².j².

L'ajustement de la distribution des dates de récolte par plant à une loi normale est en général bon (données non présentées). L'inconvénient de la loi normale est sa définition sur $(-\infty, +\infty)$. Cependant, les probabilités deviennent très faibles et négligeables à l'extérieur des distributions observées, réduisant ainsi fortement le risque d'obtenir une date de récolte très précoce ou très tardive. Les distributions de dates de récolte à l'échelle du groupe de plants estimées par le modèle reproduisent de façon satisfaisante les distributions observées (Fig. 18 et 19), malgré quelques points de ces dernières en dehors des régions de prédiction. Le modèle reproduit bien en particulier le pic de récolte et l'étalement total de la récolte sur chacune des parcelles, avec des périodes de récolte différenciées par LFR et SN_{moy} , les températures moyennes ayant été similaires au cours de la croissance des fruits sur les deux parcelles. On note cependant un léger retard de la distribution estimée sur l'essai PR (Fig. 18) à partir du début du pic de récolte. Ce sont des indicateurs importants pour l'agriculteur. D'autres indicateurs pratiques peuvent être construits à partir de chaque distribution estimée: date de demi-récolte, étalement dans le temps du pic de récolte, ...

Le modèle a été établi sur un domaine expérimental large de couples (LFR , SN_{moy}), représentatifs des valeurs observées dans les parcelles. Cela permet de l'utiliser d'une façon générale pour des situations comprises dans le domaine expérimental. Sa validité externe reste à vérifier avec des données indépendantes, notamment sur la température moyenne au cours du cycle.

L'effet de LFR et SN_{moy} sur les dates de récolte pourrait en partie expliquer la variabilité résiduelle dans la relation établie entre la durée floraison-récolte et les sommes de températures à partir de données relevées sur différentes parcelles et durant plusieurs saisons (Annexe 1). En supposant que ces données soient représentatives de l'ensemble des couples (LFR , SN_{moy}) et donnent ainsi une relation moyenne, le modèle développé permet d'ajuster la prévision des dates de récoltes pour un cycle de production donné en fonction de ses caractéristiques LFR et SN_{moy} , comme cela est montré pour chacun des essais PR et JBG (Fig. 18 et 19).

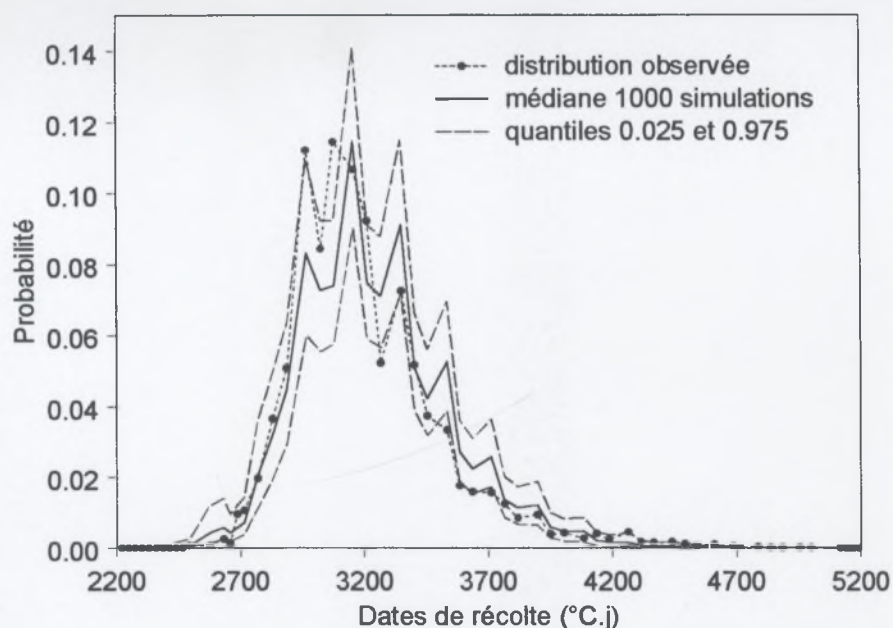


Figure 18: Validation sur l'essai PR du modèle de prédiction des dates de récolte. Distribution observée des dates de récolte sur 27 plants (19172 fruits), et médiane et région de prédiction à 95% de 1000 distributions simulées. Le rapport feuilles:fruits des plants varie de 0.8 à 13.4 et le nombre moyen de graines des fruits d'un plant varie de 8.2 à 14.4.

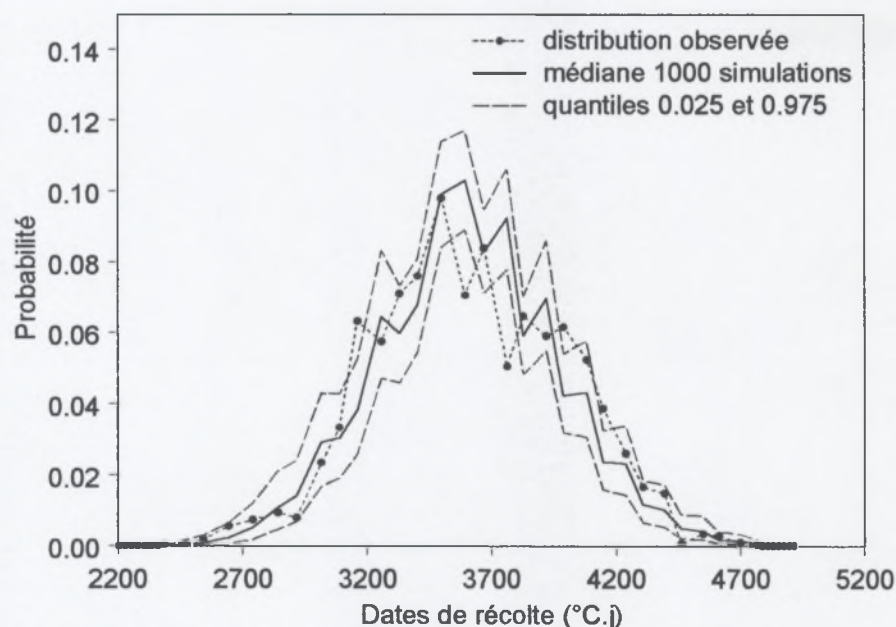


Figure 19: Validation sur l'essai JBG du modèle de prédiction des dates de récolte. Distribution observée des dates de récolte sur 11 plants (17360 fruits), et médiane et région de prédiction à 95% de 1000 distributions simulées. Le rapport feuilles:fruits des plants varie de 1.3 à 11.1 et le nombre moyen de graines des fruits d'un plant varie de 11.3 à 22.8.

Chapitre VI

Discussion et conclusion

I- La problématique

L'origine de ce travail est la domestication d'une nouvelle espèce fruitière, le goyavier-fraise, pour la diversification agricole dans des zones de climat très humide de l'île de la Réunion. La variabilité de caractères d'intérêt agronomique et économique tels que le nombre de fruits produits par plant ou leur calibre ne semble pas liée au génotype, et une approche agronomique a été retenue pour domestiquer cette espèce, en cherchant à expliquer comment s'élaborent ces caractères et leur variabilité, afin de pouvoir agir dessus par des techniques culturales appropriées. L'objectif de l'agriculteur est de diminuer cette variabilité tout en obtenant des rendements et des calibres de fruit correspondant à ses objectifs économiques.

L'analyse de la variabilité du nombre de fruits produits par plant et de leur calibre, qui caractérisent le rendement commercial de la culture, est au centre de notre problématique. L'élaboration du nombre de fruits et de leur calibre a été abordée dans une démarche inspirée de celle de l'élaboration du rendement en grandes cultures: identification des étapes importantes du cycle de production et recherche des facteurs qui les affectent. Un modèle d'élaboration du nombre de fruits et de leur calibre, structuré autour de ces étapes, est apparu comme un outil adapté pour représenter ce système complexe de processus dépendants et les effets de différents facteurs sur son déroulement. La variabilité des conditions climatiques dans la zone de culture du goyavier-fraise à la Réunion, qui sont des facteurs affectant le cycle de production, rend d'autant plus pertinente la construction d'un tel modèle. Son objectif appliqué est la simulation pour tester l'effet d'actes techniques sur les caractéristiques de la récolte, et la prédiction des caractéristiques de la récolte (nombre de fruits, calibres, dates de récolte) pour aider les agriculteurs à prendre leurs décisions techniques et commerciales.

Deux particularités, qui se retrouveront d'une façon générale dans toute démarche similaire de domestication abordée par une approche agronomique, découlent de cette approche et ont accompagné notre démarche:

- privilégier une approche agronomique amène à travailler avec du matériel végétal de semis. Il présente donc une certaine diversité génétique induisant des différences individuelles de comportement, en particulier au niveau des processus qui nous intéressent pour l'élaboration du nombre de fruits et de leur calibre. Un effet micro-environnemental peut également affecter localement dans la parcelle certaines étapes du cycle de production. Un effet plant est attendu, sur chaque étape, et *in fine* sur la récolte. L'origine génétique et/ou environnementale est toutefois difficile à distinguer sans étude spécifique. L'une ou l'autre des origines est prépondérante selon les étapes. Cette approche est une généralisation à une population hétérogène de ce qui aurait pu être fait en verger monoclinal pour une espèce horticole sélectionnée, où de plus l'environnement est généralement considéré homogène.

- domestiquer une espèce nouvelle sous-entend qu'il existe peu de références sur celle-ci. L'analyse bibliographique nous a confirmé cet état, montrant une faible quantité de

données, souvent ponctuelles et partielles. Il était donc nécessaire d'acquérir un minimum de connaissances sur l'espèce pour aborder ce travail. Une retombée méthodologique de cette étude est d'apporter des éléments pour bien cibler le champ des connaissances minimales nécessaires pour juger de l'intérêt et pour orienter une démarche de domestication basée sur une approche agronomique (voir partie VI).

II- La démarche

L'analyse critique de la démarche adoptée peut être abordée à deux niveaux. D'une part la démarche de domestication du goyavier-fraise à la Réunion basée sur une approche agronomique. D'autre part, la démarche expérimentale et de modélisation qui a été utilisée.

1- La domestication basée sur une approche agronomique

L'analyse de la démarche de domestication du goyavier-fraise à la Réunion basée sur une approche agronomique a pour objectif d'en préciser les avantages et les limites. La généralisation de cette démarche à d'autres espèces sera discutée dans la partie VI.

L'approche agronomique de la domestication, telle que nous l'avons conduite, consiste à déterminer comment se construit la variabilité sur des caractères d'intérêt, afin d'essayer d'améliorer ces derniers par des techniques culturales appropriées. Le goyavier-fraise présente des caractéristiques intéressantes pour cette démarche. Les caractères d'intérêt sont le nombre de fruits produits par plant et leur calibre. Ils sont très variables à différentes échelles: plant, parcelle, cycle de production, ce qui suggère un faible effet génétique. La couleur du fruit, qui a un intérêt pour la transformation, est liée aux deux variétés botaniques qui sont clairement distinctes, et celle à fruits rouges a été retenue.

L'étude de la phénologie en milieu naturel a permis d'identifier des étapes ou des facteurs limitants, et de tester l'effet de la température sur le développement et la productivité. Les résultats suggèrent que des conditions favorables à un bon développement des plants devraient améliorer la production, ce qui est prometteur pour la domestication. Plus généralement, l'étude de la phénologie de la plante en milieu naturel, si possible dans des conditions écologiques contrastées, permet d'acquérir rapidement (matériel végétal productif disponible) une connaissance du cycle de production et de certains facteurs qui peuvent l'affecter, et d'évaluer la variabilité existante. La mise en évidence du déclenchement de pousses florifères par un apport d'azote montre d'une part que cette espèce a un comportement et un potentiel intéressants en culture, et d'autre part qu'elle réagit favorablement à des pratiques culturales.

L'analyse de la variabilité du nombre de fruits produits par plant et de leur calibre nous a amené à découper le cycle de production en cinq étapes: l'élaboration du nombre de fleurs, la floraison, la pollinisation, la fécondation-fructification et la croissance du fruit. Seules la floraison, la fécondation-fructification et la croissance du fruit ont été abordées dans le cadre de cette thèse. Ce découpage est basé sur l'hypothèse que la croissance du fruit est affectée par le nombre de graines du fruit et le rapport feuilles:fruits du plant. Les étapes précédentes du cycle de production susceptibles d'affecter ces facteurs sont alors identifiées, et des hypothèses sont posées sur les facteurs susceptibles d'affecter ces étapes. Outre une connaissance globale de l'élaboration du rendement et du calibre, l'intérêt est de déterminer les

facteurs à travers lesquels l'agriculteur peut agir et ceux qui sont liés au plant avec un déterminisme génétique marqué.

Cette approche de l'élaboration du rendement est en général appliquée à des vergers monovariétaux et le problème de la diversité génétique ne se pose pas. Il se traduit seulement par une valeur des paramètres propre à chaque variété. Dans notre cas, le matériel végétal est hétérogène, et un effet marqué de cette diversité est observé sur la floraison et la fécondation. Il vient s'additionner aux autres sources de variabilité affectant ces étapes. Pour la floraison, cela se traduit par des lois d'acquisition de temps thermique de développement différentes entre plants. La variabilité est d'origine génétique et micro-environnementale, la première étant probablement prédominante compte tenu de l'homogénéité du dispositif expérimental sur chaque site d'observation. Pour la fécondation, l'effet plant est représenté par l'autocompatibilité qui a des retombées sur le nombre de graines formées et la fructification. Son déterminisme génétique n'a pas été formellement montré dans notre étude, mais il est fortement pressenti par référence aux données bibliographiques existant sur le sujet, et nous le considérerons acquis dans la suite de la discussion. La croissance du fruit est par contre largement déterminée par des facteurs de croissance, le nombre de graines du fruit et le rapport feuilles:fruits du plant, qui résultent des étapes précédentes. Ces facteurs dépendent donc à la fois du plant (effet génétique) et de l'effet des autres facteurs ayant affecté la floraison et la fécondation. Un effet génétique sur les processus mêmes de la croissance du fruit existe au niveau physiologique, sur le métabolisme carboné ou la division cellulaire par exemple. Cet effet n'a pas été considéré dans notre étude, et il apparaît toutefois faible par rapport à l'importance des conditions de croissance du fruit.

Ces résultats nous ont permis d'acquérir une connaissance de la façon dont s'élabore le rendement du goyavier-fraise, et des facteurs qui agissent sur les étapes étudiées. En particulier, les facteurs génétiques s'exprimant sur certains processus représentent une source de variabilité entre plants. Les facteurs environnementaux ou physiologiques affectent toutes les étapes, et l'agriculteur peut agir sur certains d'entre eux et ainsi affecter les caractéristiques de la récolte. Des hypothèses sur les processus sous-jacents ont pu être émises, par exemple pour la croissance du fruit (voir partie III-5). Des applications pratiques ou des expérimentations factorielles en découlent directement, contribuant ainsi à proposer des itinéraires techniques pour améliorer globalement la maîtrise de la production du goyavier-fraise en culture (voir partie IV). La domestication par une approche agronomique a donc permis de montrer rapidement la bonne aptitude du goyavier-fraise à la culture, et de proposer des itinéraires techniques adaptés.

Par contre, l'objectif de modélisation globale de l'élaboration du rendement, c'est à dire la quantification des processus, de leurs relations et de l'effet de facteurs, n'a pas été atteint dans le cadre de ce travail. Cela est dû à la difficulté rencontrée pour prendre en compte la variabilité génétique de certains caractères, ainsi que les conséquences de cette variabilité sur l'enchaînement des processus. Il s'agit là d'une limite de cette démarche de domestication du goyavier-fraise, imputable à l'hétérogénéité génétique du matériel végétal. Les étapes ayant posé problème sont celles dont le déterminisme génétique est marqué: la floraison et la fécondation. Il se pose le problème de la caractérisation de cette variabilité au sein d'une population pour l'intégrer dans le modèle, et celui de l'acquisition de données pour construire puis valider les modèles (exemple de la pollinisation). Le cas de la floraison a été résolu car la distribution de l'effet plant au sein d'une population peut être considérée comme normale, et un modèle général de floraison intégrant explicitement l'asynchronisme entre plants a été proposé. Par contre, le cas de la fécondation n'a pu être résolu car la variabilité de l'effet plant

au sein d'une population n'a pu être caractérisée, et l'autocompatibilité partielle et variable du goyavier-fraise rend très complexe à aborder l'étape de la pollinisation au sein d'une population hétérogène. Deux alternatives peuvent être envisagées pour contourner cette difficulté:

- rester dans la démarche actuelle d'une population hétérogène et aborder globalement les étapes de la pollinisation et de la fécondation-fructification, c'est à dire choisir de ne pas descendre à un niveau trop bas de représentation des processus car cette précision est difficile à appréhender. Des relations sont alors à rechercher entre des facteurs susceptibles d'affecter ces deux étapes et leur résultat. Ces facteurs peuvent être le déroulement de la floraison, la pluviométrie journalière, la température, la présence d'abeilles. Leur effet est à rechercher sur la distribution du nombre de graines par fruit à l'échelle du plant, la variabilité de cette distribution entre plants sur une parcelle, et le taux de fructification par plant. Cette démarche implique une base de données conséquente, c'est à dire de nombreuses observations de terrain dans des conditions contrastées.

- si une limite de l'approche agronomique est atteinte pour la modélisation de certaines étapes, on peut passer par une sélection préalable et travailler sur une population clonale pour supprimer la variabilité d'origine génétique. On se situe alors dans le cas d'une floraison avec faible asynchronisme entre plants (effet environnemental seul), et dans des conditions d'autopollinisation. La formation des graines serait alors déterminée par l'intensité de pollinisation (charge en pollen des stigmates), qui est à relier aux conditions de pollinisation. Deux expérimentations préalables peuvent permettre de préciser la démarche à adopter: 1) évaluer la variabilité de l'intensité de pollinisation et voir si elle peut être limitante, et 2) établir pour le clone sélectionné la relation entre intensité de pollinisation et distribution du nombre de graines par fruit (Falque *et al.*, 1995, 1996). Plusieurs modèles traitant des mécanismes de pollinisation et formation des graines en population clonale ont déjà été publiés (de Reffye *et al.*, 1978; Brain et Lansberg, 1981; DeGrandi-Hoffman *et al.*, 1989, Lescourret *et al.*, 1998).

2- La démarche expérimentale et la modélisation

La modélisation est apparue comme un outil privilégié pour représenter et étudier le système complexe de l'élaboration du rendement, et répondre aux objectifs de simulation et de prédiction. Dans notre cadre de travail, le but est de réaliser, dans un délai de court terme, un modèle qui répond à une problématique agronomique et qui est susceptible d'être utilisé aisément sur le terrain, c'est à dire avec des variables d'entrée facilement accessibles. Compte tenu du peu de connaissances disponibles, nous avons choisi des modèles de type descriptif, basés sur des relations statistiques. Les modèles statistiques comportent en général peu de variables, l'estimation des paramètres est relativement aisée, et leur valeur prédictive peut être élevée car leur estimation prend implicitement en compte les facteurs non considérés (Marcelis *et al.*, 1998). Cependant, les principales contraintes de ces modèles résident dans leur domaine de validité qui est limité à l'espèce considérée et aux conditions expérimentales dans lesquelles ont été ajustés les paramètres. L'intégration de nouveaux facteurs suppose de reconstruire un modèle à partir d'un jeu de données adapté (Marcelis *et al.*, 1998).

Un modèle de représentation de la plante et de son fonctionnement peut certes satisfaire cette problématique agronomique, mais à plus long terme et d'une façon sans doute plus lourde à mettre en œuvre, tant dans la construction du modèle que dans son utilisation pratique.

Le choix des étapes du cycle de production découle de l'hypothèse que la croissance du fruit dépend de son nombre de graines et du rapport feuilles:fruits du plant. Il est clair, à cause du peu de connaissances sur cette espèce, que les facteurs de variabilité *a priori* principaux ont été retenus pour chaque étape. Les résultats montrent qu'ils expliquent en général bien la variabilité observée, mais pas entièrement. La variabilité non expliquée est due à des facteurs non considérés. Cette variabilité non expliquée doit être prise en compte, d'une part car les facteurs retenus ne sont pas exhaustifs, et d'autre part car elle participe à la variabilité, et donc à la précision, des prédictions des modèles, ce qui est un élément d'aide à la décision dans les choix techniques. Des approches stochastiques ont été retenues, soit en considérant qu'une étape est une succession de processus aléatoires (cas des modèles de floraison et de fécondation-fructification), soit en couplant à la partie déterministe du modèle une partie probabiliste traduisant la variabilité observée et non expliquée (cas du modèle de croissance du fruit). Cette démarche peut amener à une amélioration progressive des modèles, au fur et à mesure de l'acquisition future de connaissances, par l'intégration de nouveaux facteurs dans la partie déterministe des modèles, réduisant ainsi la variabilité non expliquée.

Les objectifs d'analyse de la variabilité et de construction de modèles statistiques nous ont amenés dans nos expérimentations à considérer une large gamme de variabilité des facteurs environnementaux ou de croissance étudiés. Cela permet d'une part de mettre en évidence des relations entre ces facteurs et les réponses du processus étudié, et d'autre part d'estimer les paramètres des modèles sur un domaine expérimental large sensé représenter les conditions rencontrées en culture à la Réunion, conférant ainsi un domaine de validité correspondant à l'île. Cela s'est traduit par des sites expérimentaux sur une gamme altitudinale, pour tester l'effet de la température, large (200 - 900 m) et représentative des zones potentielles de culture de cette espèce à la Réunion. Les gammes expérimentales de variation du rapport feuilles:fruits du plant, du nombre de graines par fruit ou du nombre moyen de graines par fruit sur un plant se sont montrées représentatives de ce que l'on observe en verger.

L'extrapolation des modèles à la zone de culture du goyavier-fraise à la Réunion reste à être validée. Elle implique l'hypothèse que tous les facteurs non considérés au niveau de la parcelle expérimentale, et dont les effets sont pris en compte dans le paramétrage du modèle, soient eux aussi représentatifs des conditions de culture du goyavier-fraise à la Réunion, ou ne soient pas limitants dans ces mêmes conditions. Les facteurs *a priori* les plus importants sont la diversité génétique des plants suivis, qui doit être représentative si elle a un effet sur l'étape modélisée, le rayonnement et la disponibilité en eau, qui ne doivent pas être limitants. Le stress hydrique est toutefois peu probable dans nos écologies humides tout au long de l'année.

Pour la diversité génétique, on suppose que la diversité des plants sur lesquels ont été paramétrés les modèles est représentative, pour le ou les processus modélisés, de la diversité génétique des vergers. Cette hypothèse semble raisonnable pour les modèles de floraison et de croissance du fruit où une validation externe a pu être réalisée de façon satisfaisante sur une ou des parcelles différentes de celle ayant servi au paramétrage. Pour le modèle de fécondation-fructification, l'échantillon était réduit. La diversité des niveaux d'autocompatibilité est forte, mais leur distribution n'est pas connue. L'échantillon ne semble pas représentatif des conditions de production pour ce qui est du taux maximal de fructification.

Les parcelles reçoivent probablement des rayonnements différents du fait des écarts de nébulosité liés à l'altitude. Ce facteur peut avoir une influence sur la croissance du fruit qui dépend de l'assimilation carbonée, à moins que le rayonnement dans la zone de culture ne soit

pas limitant pour l'espèce. Son effet n'est pas mis en évidence dans la validation du modèle de croissance du fruit: même zone géographique et année différente; les facteurs explicatifs des différences entre valeurs estimées et observées sont plutôt à rechercher à l'échelle du plant ou du fruit. Il serait nécessaire de valider ce modèle dans des sites où des différences de rayonnement, qui serait alors à quantifier, sont attendues.

Enfin, un effet du dispositif de plantation n'est pas à exclure. Les densités de plantation des parcelles expérimentales sont semblables à celles utilisées en parcelle de production (1-1.5 m x 3-4 m). Les modèles réalisés s'adressent à des plants non encore jointifs sur la ligne de plantation, comme l'étaient les plants des parcelles expérimentales.

III- Analyse critique des principaux résultats

Une analyse critique des principaux résultats est présentée, et les pistes de recherche ouvertes par ce travail, sur la connaissance de la plante ou pour affiner les modèles, sont évoquées.

1- Phénologie du goyavier-fraise en conditions naturelles

L'étude de la phénologie du goyavier-fraise en conditions naturelles à la Réunion a permis d'acquérir des données de base sur son cycle phénologique et sur les facteurs susceptibles de l'affecter. En particulier, la température a un effet marqué sur le développement des pousses, des boutons floraux et des fruits, mais elle ne semble pas affecter les paramètres mesurés de l'élaboration du rendement car leurs valeurs entre sites ne s'ordonnent pas en fonction de l'altitude.

La morphologie des pousses apparaît relativement constante (90% des pousses émises ont entre 4 et 6 entre-nœuds), mais leur nombre émis par rameau, et surtout leur caractère florifère, sont variables. On observe en milieu naturel une variabilité liée à l'année, au site et au plant, qui est probablement à rapprocher de la variabilité mise en évidence entre plants et entre cycles de production en parcelle cultivée. Un effet du site et de l'année a été mis en évidence pour le nombre de fleurs et de fruits par rameau. A l'échelle du plant, la variabilité est également marquée, en particulier pour le taux de fructification.

Les contributions respectives du milieu et des caractéristiques du plant à la capacité de production n'ont pas pu être déterminées durant cette étude. Les résultats montrent que des conditions favorables à un bon développement des plants devraient améliorer la production: la faible productivité a pour origine avant tout une faible émission de pousses florifères; il existe une étroite relation entre le nombre de boutons floraux émis et le nombre de pousses émises; de plus, les paramètres de l'élaboration du rendement se sont montrés sensibles à un stress, lié au passage d'un cyclone. Ces résultats sont prometteurs pour la mise en culture du goyavier-fraise, avec des techniques culturales favorisant le bon développement des plants.

Le déterminisme du débourrement n'a pas pu être mis en évidence. Les décalages dans l'époque de débourrement entre 1992 et 1993, et entre sites en 1992, suggèrent qu'un facteur constant comme la photopériode n'est pas impliqué. Le débourrement serait alors échelonné entre les basses et les hautes altitudes en fonction des différences de température entre l'induction et le débourrement. Il pourrait s'agir d'une levée de stress sur chaque site, comme par exemple une augmentation de température ou une pluviométrie importante au cours de la

saison sèche et fraîche. Ces hypothèses demandent à être vérifiées. La connaissance du déterminisme du débourrement du cycle naturel est importante pour la construction du modèle d'élaboration du rendement afin d'initialiser une échelle de temps thermique entre le débourrement, ou son induction, et la floraison. Sans cette connaissance, la prédiction de la floraison pourrait être envisagée à partir du débourrement des bourgeons, stade phénologique repérable sur les plants (voir paragraphe 3- suivant).

2- Déclenchement de pousses florifères par un apport d'azote

La mise en évidence du déclenchement de pousses florifères par un apport d'azote chez le goyavier-fraise est un résultat important d'un point de vue agronomique et scientifique. Au niveau agronomique, cette caractéristique permet à l'agriculteur de réaliser deux récoltes par an sur une même parcelle, au lieu d'une seule, et de choisir une période de production où la valorisation des fruits est intéressante. L'étalement des dates de déclenchement sur plusieurs parcelles assure une récolte prolongée, intéressante en particulier pour la commercialisation ou la gestion de la main-d'œuvre. La variabilité du nombre de fruits par plant et de leur calibre entre les cycles de production (Fig. 4), qui est en partie à l'origine de ce travail, est maintenant mieux connue.

Au niveau scientifique, les résultats suggèrent que l'azote n'a qu'un rôle d'induction de la pousse, et que l'intensité végétative et reproductrice de celle-ci dépend surtout de la disponibilité en carbone. Cette hypothèse serait à vérifier, en particulier par des dosages de différentes formes carbonées (sucres solubles, substances de réserve) ou la mesure de l'activité photosynthétique avant et après l'apport d'azote, et la recherche de relations entre ces variables et l'intensité de la réaction de la plante. Ces résultats pourraient de plus amener à définir un indicateur de l'intensité végétative et reproductrice de la pousse émise, qui est intéressant pour la mise au point du modèle et son utilisation ultérieure pour des cycles de production avec interaction.

Le rôle principalement inducteur de l'azote est corroboré par des observations en parcelles où des pousses florifères sont émises, en dehors du cycle naturel de production, après une taille ou un effeuillage par exemple. Ces éléments viennent appuyer l'hypothèse d'une levée de stress déterminant l'émission de la pousse naturelle.

Le comportement du goyavier-fraise lors de pousses déclenchées fait penser à une forme d'alternance (Goldschmidt et Colomb, 1982; Monselise et Goldschmidt, 1982). En cycle naturel, diverses observations tendent à montrer que le goyavier-fraise n'est pas une espèce alternante (les résultats du suivi de la phénologie ne donnent pas d'indication sur ce point car un cyclone a perturbé le fonctionnement des plantes entre les deux années). Par contre, il se comporte comme une espèce alternante lorsque l'on considère des cycles de production successifs ou superposés. Le goyavier-fraise pourrait alors être un matériel de choix pour l'étude des mécanismes de l'alternance, liés très probablement dans ce cas à la disponibilité en carbone.

3- La floraison

Le processus de base du modèle de floraison est l'acquisition probabiliste par chaque bouton floral de petits incréments de temps de développement. La mise en évidence de

l'asynchronisme de floraison entre les plants, probablement liée à des aspects génétiques, et l'irrégularité de la floraison sur chaque plant, reflétant un asynchronisme de floraison au sein du plant, nous a conduit à scinder la population globale de boutons floraux de la parcelle en ensembles supposés suivre une même loi d'acquisition de temps de développement. Chez le goyavier-fraise, cet ensemble de base, appelé "unité florifère", est constituée par les boutons floraux des pousses florifères issues d'un même rameau. Chaque "unité florifère" est caractérisée par la différence entre son seuil de floraison spécifique et le seuil de floraison du plant. Ces différences de seuil reflètent une acquisition plus ou moins rapide de temps de développement, liée à leur micro-environnement, en partie à leur nombre de boutons floraux, et probablement à des facteurs génétiques. Les seuils de floraison des plants ont probablement une large détermination génétique (ce pourrait être testé par des suivis sur plusieurs années de la floraison individuelle de certains plants), et traduisent aussi des différences de l'environnement thermique des plants. Cette approche a permis de structurer la variabilité d'acquisition de temps de développement à l'échelle des boutons floraux de la parcelle en trois composantes:

- la variabilité d'acquisition de temps de développement des boutons floraux de l'"unité florifère" qui est la variabilité du processus stochastique,
- la variabilité d'acquisition moyenne de temps de développement entre les "unités florifères" au sein du plant (asynchronisme au sein du plant),
- la variabilité d'acquisition moyenne de temps de développement entre les plants de la parcelle (asynchronisme entre plants).

Chacune de ces composantes est quantifiée par un paramètre du modèle 'rameau', le quatrième paramètre étant la moyenne des seuils de floraison des plants sur la parcelle.

Compte tenu des objectifs de prédiction, le modèle intermédiaire 'plant' a été retenu. Il néglige l'asynchronisme de floraison au sein du plant, c'est à dire qu'il ne considère que la variabilité d'acquisition de temps de développement des boutons floraux de chaque plant, et celle moyenne entre les plants de la parcelle.

L'effet de la température sur la floraison est bien rendu par le modèle qui reproduit correctement la durée totale de la floraison à l'échelle du plant et à l'échelle de la parcelle, ainsi que la période du pic de floraison dans ce dernier cas.

Nous avons considéré la floraison au sens strict du terme, c'est à dire l'ouverture du bouton floral en une fleur susceptible de libérer son pollen et d'être pollinisée. Le modèle de floraison construit ne s'intéresse qu'à ce stade de l'ouverture du bouton floral et non aux événements phénologiques précédents la floraison comme cela a été fait chez le kiwi (Agostini, 1995). Chez le goyavier-fraise, la détermination du nombre de boutons floraux donnant une fleur est abordée dans une étape relative à l'élaboration du nombre de fleurs qui n'a pu être traitée dans le cadre de cette thèse (voir partie V).

D'un point de vue temporel se pose le problème du manque de date initiale pour démarrer l'accumulation de temps de développement pour une échelle de temps avant la floraison dans le cas d'un cycle naturel. Le déterminisme du débourrement n'étant pas connu, on ne peut utiliser une date avant le débourrement. Une date initiale pourrait être identifiée au moment du débourrement. Cela nécessite une étude de sa dynamique et l'établissement d'un modèle thermique en parcelle cultivée pour l'intervalle débourrement-floraison. Dans l'état actuel des connaissances, l'échelle de temps thermique associée au déroulement de la floraison naturelle

ne peut être déterminée qu'*a posteriori*, quand on connaît la date de demi-floraison, comme cela a été fait pour ajuster le modèle aux données. Dans le cas de la floraison de cycles déclenchés, qui n'ont pas été abordés dans ce travail, le point de départ de l'échelle de temps thermique sera la date de fertilisation (Annexe 1).

4- La fécondation et la fructification

Les résultats des expérimentations réalisées sur la biologie de la reproduction du goyavier-fraise permettent de mieux connaître cette étape et les facteurs qui l'affectent. Un modèle statistique qui reproduit les distributions de graines par fruit observées et prend en compte les facteurs affectant ces distributions a été construit. Ces facteurs sont les conditions de pollinisation, l'origine génétique du pollen et l'autocompatibilité de chaque plant. L'autocompatibilité variable mise en évidence dans les populations de goyavier-fraise est une conséquence directe de leur hétérogénéité génétique. Ce résultat est une des raisons de la construction d'un modèle de floraison capable de simuler la floraison de chaque plant relativement à l'ensemble des autres plants de la parcelle.

L'existence de deux populations de fleurs à l'échelle du plant ayant des comportements différents vis à vis de la fécondation reste à vérifier et à expliciter. Plusieurs hypothèses sur leur origine ont été écartées, et la voie d'une explication génétique est à explorer chez cette espèce fortement polyploïde et présentant des anomalies dans la formation des gamètes (Singhal *et al.*, 1985; Raseira et Raseira, 1996).

Malgré sa petite taille, notre échantillon de plants a montré une grande variabilité des taux d'autocompatibilité. L'estimation de leur distribution au sein d'une population demande toutefois des travaux complémentaires avec des dispositifs expérimentaux appropriés.

Une autre approche aurait pu être de mieux comprendre les mécanismes de la fécondation chez le goyavier-fraise par une étude plus poussée de la biologie de sa reproduction: viabilité du pollen, suivi de la germination des grains de pollen (auto et allo) et de la progression des tubes polliniques jusqu'aux ovules, fertilité et fécondation préférentielle des ovules, avortement post-zygotique, recherche de relations entre intensité de pollinisation et formation de graines. Ces travaux, d'ordre cognitif, auraient pu être menés sur quelques plants pour identifier à quel(s) niveau(x) se manifeste l'effet plant (e.g. Sedgley *et al.*, 1989; Ellis et Sedgley, 1993; Falque *et al.*, 1996). Ils n'étaient cependant pas réalisables dans le cadre de cette thèse. Ils auraient l'intérêt de clarifier la biologie de la reproduction chez cette espèce où de nombreuses anomalies, probablement liées à la forte polyploïdie, sont observées à différentes étapes de la reproduction (Singhal *et al.*, 1985; Raseira et Raseira, 1996). De ce fait, un schéma classique de déroulement de la fécondation comme celui que nous avons utilisé pour construire le modèle d'aide à l'analyse des résultats n'est peut-être pas adapté. Cette approche explicative pourrait permettre d'apporter des éléments pour construire un modèle de fécondation adapté. Mais elle ne résoudrait pas le problème de la prise en compte de la diversité génétique d'une population, ni le problème de la pollinisation.

5- La croissance du fruit

L'étude de la croissance du fruit du goyavier-fraise a montré qu'il a une courbe de croissance en double-sigmoïde peu commune car elle ne présente pas de plateau final: le fruit chute

spontanément de la plante alors qu'il est en pleine croissance exponentielle. La date correspondante est retenue comme la date de récolte, ou de maturité. Le nombre de graines par fruit (*SN*), et le rapport feuilles:fruits du plant (*LFR*) ont une influence sur les paramètres de la courbe de croissance qui se traduit par un effet sur le poids final du fruit à maturité et sur la date de récolte. La forme exponentielle finale de la courbe, et le fait que la date de récolte individuelle ne peut être correctement estimée à partir des facteurs *SN* et *LFR*, nous conduit à étudier séparément les effets des facteurs sur le poids final et sur la date de récolte, indépendamment de l'équation de la courbe de croissance.

Les résultats obtenus sur le poids final du fruit montrent que la croissance de ce dernier dépend fortement de la fourniture en assimilats. La séparation précoce des courbes de croissance sous l'effet de *LFR* suggère que la mobilisation des réserves du plant pour la croissance du fruit est faible, et que celle-ci est alimentée directement en assimilats par les feuilles. Nous avons choisi un indicateur synthétique de la disponibilité globale en carbone à l'échelle du plant, son rapport feuilles:fruits mesuré à la nouaison (*LFR*). Le choix de cette variable a été avant tout pratique car elle peut se calculer à partir des sorties des modèles des étapes précédentes (Fig. 5), et elle correspond également aux objectifs d'utilisation du modèle sur le terrain ne nécessitant pas d'entrées difficiles à acquérir. Le nombre de fruits du plant est fourni par l'étape de fécondation-fructification. Des études spécifiques nous ont montré que le nombre de feuilles peut être estimé à partir d'une part de la section du tronc avant l'émission de la pousse florifère (relation allométrique donnant le nombre de feuilles âgées) et d'autre part du nombre de pousses émises (jeunes feuilles), et enfin d'un taux de chute de feuilles avant la nouaison.

Le choix de *LFR* à l'échelle du plant est critiquable d'un point de vue physiologique pour plusieurs raisons:

- Il considère que toutes les feuilles apportent la même contribution à la croissance des fruits, ce qui n'est sans doute pas le cas, à cause de leur âge, de leur position plus ou moins externe dans l'architecture, et de leur distance respective par rapport aux fruits.

- Il donne le même poids aux sources (nombre de feuilles) et aux puits (nombre de fruits) dans leur effet sur la croissance. Or le nombre de fruits semble avoir un effet plus fort sur la croissance, en particulier lorsque $LFR < 2.5$, ce qui correspond aux plants les plus chargés. Cela est illustré dans le tableau I de l'article sur la croissance du fruit (Chapitre V) où pour des plants dont $LFR = 1.5$, le poids moyen des fruits à maturité est plus élevé (5.3 g contre 4.3 g) lorsque la charge en fruits est moyenne (623 fruits par plant en moyenne) que lorsqu'elle est élevée (1300 fruits par plant en moyenne). La prise en compte séparée des sources et des puits à l'aide des nombres de feuilles et de fruits comme variables explicatives aurait sans doute permis de mieux prendre en compte la contribution de chaque effet et d'affiner les sorties du modèle.

- *LFR* ne prend pas en compte la variabilité des rapports feuilles:fruits locaux, à l'échelle de la pousse, qui déterminent en grande partie les conditions de croissance du fruit. Nos résultats suggèrent que cet effet local serait à considérer pour affiner les sorties du modèle. Une part non négligeable de la variabilité résiduelle du modèle est probablement imputable à ces effets locaux. Dans notre démarche globale, leur intégration ne pourrait se faire que de façon statistique. Une approche plus mécaniste nécessiterait au moins un modèle architectural et un modèle écophysiologique.

- *LFR* est calculé à la nouaison et est considéré constant durant toute la croissance des fruits. Le nombre de fruits n'évolue pas, ou très peu, après la fructification. Dans notre expérimentation, l'apparition de nouvelles feuilles a été prévenue. Cependant, une chute des feuilles existe. Cette chute a tendance à faire baisser *LFR* entre la fructification et la récolte. Ces chutes ont été quantifiées: environ 13%, indépendamment de *LFR*, dans l'expérimentation (essai PR), et de 21 à 43% sur les 11 plants utilisés pour la validation externe (essai JBG). Ces chutes ne sont pas régulières et interviennent principalement au cours de la croissance exponentielle du fruit. Or nos résultats indiquent que la croissance dépend alors de l'alimentation carbonée des fruits. Ces chutes pourraient donc induire des erreurs dans la prédiction du poids final du fruit dans le cas où les chutes effectives diffèrent de celles relevées dans notre expérimentation. Ce peut être un facteur expliquant les écarts entre distributions observée et estimée sur certains plants en validation externe. Il serait intéressant de voir comment quantifier cette diminution de *LFR* et la considérer dans le modèle.

Malgré ces critiques, *LFR* à l'échelle du plant s'est révélé être une variable explicative qui rend compte d'une grande part de la variabilité du poids moyen des fruits entre les plants. Cela va dans le sens de l'hypothèse d'une contribution directe des feuilles à la croissance des fruits. Le nombre de graines explique lui une large part de la variabilité du poids individuel du fruit au sein du plant.

La recherche de relations entre les facteurs de croissance étudiés et les paramètres de la courbe de croissance du fruit ont permis de mettre en évidence l'effet différentiel de ces facteurs sur les grandeurs quantitatives et temporelles de la courbe de croissance. Cette approche est intéressante pour l'étude de la croissance d'un fruit nouveau, sous réserve de considérer des facteurs expliquant une grande partie de la variabilité observée. Les résultats obtenus peuvent avoir des retombées agronomiques immédiates (par exemple, l'efficacité de l'éclaircissage sera maximale s'il est pratiqué très tôt après la fructification), mais appellent à des études complémentaires pour approfondir les mécanismes en jeu. En terme de modélisation, cette approche est coûteuse en précision, et l'estimation directe des paramètres d'un modèle dérivé a permis d'améliorer la précision des sorties. Ces deux approches sont complémentaires et dépendantes: le modèle dérivé n'aurait pu être construit si les relations entre les paramètres de la courbe de croissance et les facteurs de croissance n'avaient été établies. La précision des sorties de tels modèles statistiques dépend de la pertinence des facteurs identifiés pour expliquer la variabilité observée.

Si *LFR* seul explique de façon satisfaisante la variabilité des dates de récolte entre les plants, la variabilité des dates individuelles de récolte au sein des plants n'a pu être expliquée avec *SN* comme variable explicative. De nombreux facteurs (température effective du fruit, ensoleillement, compétition locale pour les assimilats, date de floraison, ...) peuvent influencer la vitesse de développement du fruit. Nous avons alors opté pour la détermination de la distribution des dates de récolte directement à l'échelle du plant. Il serait toutefois intéressant de tester l'effet de la date de floraison et du rapport feuilles:fruits local à l'échelle de la pousse sur la date de récolte du fruit. Si ces facteurs s'avéraient suffisamment explicatifs, leur intégration dans le modèle sous sa forme actuelle ne pourrait se faire que de façon probabiliste et indépendante, alors que ces facteurs sont probablement liés en réalité, le nombre de fruits d'un rameau dépendant entre autres des conditions de pollinisation et donc des dates de floraison sur le rameau.

Les deux variables explicatives retenues pour la distribution des dates de récolte à l'échelle du plant ont été *LFR* et le nombre moyen de graines des fruits du plant (*SN_{moy}*, accessible par les

sorties de l'étape de fécondation-fructification). Ces variables ont un effet à la fois sur la date de demi-récolte du plant, et sur l'étalement de la récolte, avec une augmentation de ces réponses quand LFR diminue et SN_{moy} augmente. Les critiques relatives au choix de LFR à l'échelle du plant comme variable explicative restent en partie valables pour la prédiction de la distribution des dates de récolte (Tableau 9), et l'utilisation indépendante du nombre de feuilles et du nombre de fruits du plant comme variables explicatives aurait été plus opportune. L'estimation de la date de demi-récolte du plant ($d_{0.5rec}$) peut être affinée en calculant l'échelle de temps thermique à partir de la date de demi-floraison du plant. Nos calculs sur 10 plants dont la floraison était connue ont alors montré une chute de la variance de $d_{0.5rec}$ de plus de 27%. La date de demi-floraison de chaque plant est une sortie du modèle de floraison, et il sera possible d'avoir une meilleure précision sur $d_{0.5rec}$ dans le modèle global. Cela nécessite toutefois de réestimer, avec un nouveau jeu de données adapté, les paramètres de la relation liant $d_{0.5rec}$ à ses deux variables explicatives.

La construction du modèle de prédiction de la distribution des dates de récolte a utilisé nos deux jeux de données, ce qui n'a pas permis de le tester sur des données indépendantes. Il serait intéressant de le faire, sur des parcelles à des altitudes contrastées (par exemple sur les parcelles à 200 et 890 m) afin de tester le comportement du modèle par rapport à la température.

Les facteurs de croissance, et en particulier LFR , ont une influence sur la durée de développement du fruit. Une approche multiplicative pour aborder la modélisation de la croissance du fruit n'aurait pas permis de rendre compte de cet effet sur les courbes individuelles de croissance car elle n'affecte que la variable modélisée, comme par exemple le poids cumulé. La recherche de relations entre les paramètres de la courbe de croissance et les facteurs de croissance permet de rendre compte de leurs effets sur la durée de développement du fruit. Ainsi, l'allongement de la phase de faible croissance du fruit sous l'effet de LFR (Chapitre V-II, Fig. 13) est traduit au niveau de l'équation de la courbe de croissance par le paramètre t_E de l'exponentielle. Il correspond à la date où le taux de croissance de cette phase est égal à β , paramètre peu variable et faiblement corrélé à t_E ($r = -0.26$; $p < 0.001$). Sa valeur dépend donc en partie de la durée de la phase de faible croissance du fruit, et l'on vérifie qu'il est affecté significativement par LFR : $t_E = 2394.6 + 849.3/LFR$ $r^2=0.66$; $n=131$; $P=0$.

Dans notre étude sur la croissance du fruit, nous n'avons pas recherché d'éventuel effet plant. Or un tel effet n'est pas à exclure à cause des éléments qui peuvent avoir un effet sur la capacité photosynthétique du plant ou sur les processus de croissance du fruit: les différences génétiques entre plants et les différences environnementales locales (sol, ensoleillement). Compte tenu de l'effet marqué des variables explicatives étudiées sur la croissance du fruit, l'expression des différences génétiques entre plants sur les processus mêmes de croissance apparaît plus faible que sur les étapes de la floraison et de la fécondation.

IV- Les retombées pratiques de cette étude

Les retombées de cette étude sont scientifiques, par l'acquisition de connaissances sur le goyavier-fraise et leur synthèse en modèles qui ont été présentés précédemment, et par la délimitation des intérêts et limites de la domestication par une approche agronomique (voir ce chapitre, partie VI). Nous ne reviendrons pas sur ces points. Les retombées sont également pratiques pour les agriculteurs. Il est clair que les travaux présentés ne permettent pas d'atteindre l'objectif concret de l'outil de prédiction sur le cycle complet de production, mais

les connaissances acquises et certains modèles sont déjà utiles pour l'appui technique auprès des agriculteurs ou pour orienter des expérimentations sur les techniques culturales.

Les agriculteurs conduisent régulièrement deux cycles de production par an sur leurs parcelles, dont un est déclenché par un apport d'azote, et l'autre est le cycle naturel. Les relations entre vitesse de développement et température sont couramment utilisées pour prévoir la date de demi-récolte sur une parcelle en fonction de la date observée de demi-floraison ou de la date d'apport d'azote, ou à l'inverse pour déterminer une date pour l'apport d'azote en fonction d'une période de récolte souhaitée.

Les connaissances sur les facteurs affectant la croissance du fruit et le déclenchement d'un cycle de production suggèrent des recommandations pour la conduite de la culture, qui demandent toutefois à être validées expérimentalement. Ainsi, l'apport de ruches durant la floraison peut favoriser la pollinisation croisée dans les zones où il y a peu d'abeilles, ou si le risque climatique sur la pollinisation est fort. Si les plants sont très chargés à la nouaison, un éclaircissage précoce va favoriser la croissance et augmenter le calibre à la récolte (voir ci-dessous). La date de fertilisation est choisie avant tout par rapport à une période de production. Ce choix peut être modulé par le stade phénologique du plant et sa charge au moment prévu pour l'apport, et par les conditions climatiques probables lors de la floraison (qui seront favorables ou non à la pollinisation) et lors de la récolte (pluie et chaleur affectent la qualité des fruits).

Le rapport feuilles:fruits du plant est un facteur prépondérant de la croissance du fruit. Un antagonisme apparaît entre la charge du plant et le calibre des fruits et un équilibre est à trouver afin de maximiser le revenu de l'agriculteur. Compte tenu de l'importance économique du calibre du fruit, sur le plan commercial et des coûts de production, il est préférable de favoriser le calibre du fruit plutôt que la charge du plant. Le couplage d'un modèle micro-économique à l'échelle de la parcelle (Fusillier et Normand, 1999) avec le modèle de croissance du fruit peut aider à déterminer une gamme de rapports feuilles:fruits optimale.

L'éclaircissage des jeunes fruits après la nouaison sur des plants chargés peut améliorer le calibre des fruits à la récolte et favoriser l'intensité végétative et florifère d'un cycle de production déclenché au moment de, ou juste après, la récolte, ce qui est la pratique courante. En effet, le statut carboné du plant est *a priori* meilleur si le rapport feuilles:fruits du plant est élevé. Un essai a été mis en place sur ce thème en 1999, et les résultats espérés ont été atteints, tant sur l'amélioration du calibre du fruit que sur l'émission du cycle de production suivant. Un équilibre durable entre la production du cycle naturel et celle du cycle déclenché a été observé sur deux ans, soit quatre cycles de production, sur les plants éclaircis. Sur les plants non éclaircis, l'alternance de production entre les cycles naturels et déclenchés est maintenue (Normand et Huet, 2001).

Le contrôle du rapport feuilles:fruits du plant est un enjeu fort pour la croissance du fruit et la gestion des cycles de production. Cela suggère que l'autocompatibilité du plant est un critère intéressant pour une future phase de sélection. Seule l'autopollinisation est présente dans une parcelle monovariétale, et une autocompatibilité moyenne à faible peut avoir un effet d'auto-éclaircissage des plants conduisant à des rapports feuilles:fruits élevés. Un modèle global d'élaboration du rendement serait alors d'une aide précieuse pour déterminer par simulation les valeurs de taux d'autocompatibilité à rechercher en fonction d'objectifs de production définis.

Bien que le modèle global d'élaboration du nombre de fruits et de leur calibre n'ait pu être achevé, les modèles relatifs à la croissance du fruit peuvent être utilisés, après validation, pour prédire les caractéristiques de la récolte. En effet, les valeurs des facteurs de croissance, *LFR* et *SN*, et leur variabilité peuvent être estimées sur le terrain après la nouaison, selon un protocole qui reste à déterminer. Le modèle de croissance du fruit permet alors d'estimer environ 4 mois à l'avance le nombre de fruits et leur calibre. Le modèle de prévision des dates de récolte permet de donner la distribution des dates de récolte, affinant ainsi les prédictions des modèles thermiques.

Le modèle de floraison décrit le déroulement de la floraison, à différentes échelles (unité florifère, plant, parcelle) sous l'influence de la température. Il a des utilisations possibles dans divers domaines où la simulation du déroulement de la floraison est utile: phénologie, interactions entre la biologie florale et la reproduction (échange de gènes, taux de fécondation croisée, comportement des pollinisateurs, formation des graines et des fruits), émission de pollen allergène dans l'atmosphère, contrôle des maladies et ravageurs de la floraison.

V- La construction du modèle global d'élaboration du nombre de fruits et de leur calibre

Notre travail a contribué à la construction d'un modèle global d'élaboration du nombre de fruits et de leur calibre chez le goyavier-fraise dans le cadre d'un cycle de production naturel grâce à la construction de modèles de floraison et de croissance du fruit, et de l'étude de la fécondation-fructification. Pour atteindre l'objectif du modèle global, il manque la modélisation de l'élaboration du nombre de fleurs et de la pollinisation-fécondation-fructification, la validation individuelle de ces modèles, puis l'accrochage des différents modèles, avant l'étude et la validation du modèle global. Pour la prédiction du déroulement de la floraison naturelle, il est de plus nécessaire de construire un modèle thermique entre le débourrement et la floraison (voir partie III-3 de ce chapitre).

La modélisation de l'élaboration du nombre de fleurs peut prendre différentes formes. Une approche architecturale serait longue à mettre en place, et nécessiterait une liaison avec un modèle physiologique ou de fonctionnement afin de relier l'intensité végétative et florifère des pousses à des caractéristiques propres des rameaux, du plant, ou à des facteurs externes. L'approche que nous privilégions est de type élaboration du rendement et a été utilisée dans l'étude de l'effet de l'azote sur le déclenchement de pousses florifères (Chapitre II). Elle distingue, dans le nombre de fleurs produites par plant, les caractéristiques du plant et les caractéristiques végétatives et florifères des pousses émises. L'effet de différents facteurs est recherché sur les paramètres de la décomposition. Des études préliminaires ont ainsi montré que le développement du plant, estimé par la section du tronc au niveau du sol, a un effet sur certains de ces paramètres. Une approche probabiliste semble la plus opportune pour restituer la variabilité observée.

L'étude et la modélisation séparées de la pollinisation et de la fécondation-fructification sont complexes pour le goyavier-fraise, et il a été proposé de traiter globalement ces deux étapes si l'on reste dans une démarche de matériel végétal hétérogène. Les grandes lignes de cette approche sont présentées dans ce chapitre, partie II-1 sur l'analyse de la démarche. L'autre alternative, également présentée dans cette partie, est de travailler sur du matériel végétal homogène, c'est à dire après sélection. Dans les deux cas, un modèle global d'élaboration du rendement n'est envisageable que sur le moyen terme compte tenu des travaux nécessaires.

Après validation individuelle de chaque modèle dans des conditions représentatives de la zone de culture du goyavier-fraise à la Réunion, leur accrochage en un modèle global d'élaboration du nombre de fruits et de leur calibre est à réaliser, et des jeux de données sont à prévoir pour la validation. Il est indispensable d'avoir accès à des sorties intermédiaires des modèles, qui seront comparées aux données du terrain, afin de déterminer les origines des écarts sur les sorties finales. L'accrochage des modèles va entraîner l'accumulation des erreurs statistiques provenant de chaque modèle, et la variabilité des sorties est à considérer.

VI- La démarche est-elle généralisable?

La question de la généralisation de la démarche présentée dans ce travail peut être abordée à deux niveaux: d'une part la généralisation de la domestication d'une espèce végétale par une approche agronomique, et d'autre part la généralisation des modèles proposés.

1- La domestication par une approche agronomique

La domestication d'une espèce sauvage par une approche agronomique consiste à identifier les facteurs environnementaux et/ou physiologiques affectant des caractères d'intérêt définis, afin de pouvoir avoir une action sur ces derniers par des pratiques culturelles adaptées. La méthodologie est basée sur l'étude de la variabilité de ces caractères d'intérêt, et la façon dont elle s'élabore, sous l'effet de facteurs environnementaux, physiologiques et génétiques, au cours des différents processus dépendants conduisant à ces caractères.

Cette approche agronomique n'est pas à mettre en opposition ou ne se substitue pas à une approche traditionnelle de la domestication basée sur l'étude de la diversité du matériel végétal et la sélection. Cette phase est souvent longue et déconnectée des problèmes de la production. La mise au point des itinéraires techniques est ensuite réalisée sur ce matériel végétal sélectionné. L'approche par la sélection nécessite des structures capables de multiplier et diffuser le matériel végétal sélectionné.

L'approche agronomique est complémentaire de cette approche traditionnelle, et permet d'apporter à court terme une connaissance générale de la plante conduisant à proposer des itinéraires techniques pour améliorer globalement la conduite d'une parcelle de matériel végétal hétérogène par rapport à des objectifs agronomiques. Elle n'est pas exclusive de phases de sélection destinées à limiter la variabilité génétique, sans toutefois rechercher la sélection de cultivars. Elle peut de plus contribuer à définir des critères de sélection. Leakey et Simons (1998) distinguent deux stratégies extrêmes de domestication d'espèces pérennes. L'une, tournée vers les petits agriculteurs, est l'amélioration incrémentielle de l'espèce par la gestion culturelle sur la ferme. Un exemple typique est l'agroforesterie. L'autre, orientée par les marchés, est la sélection et l'amélioration génétique, conduisant à des cultures à grande échelle. L'approche proposée se situe entre ces extrêmes, tant au niveau de la méthodologie que des utilisations. De plus, la stratégie peut évoluer en fonction des objectifs, sous l'effet notamment de l'importance commerciale que prend l'espèce.

Cette approche n'est envisageable que si les caractères d'intérêt apparaissent peu liés au génotype. Une étude préalable de leur variabilité et de sa structure est donc nécessaire. D'autre part, la diversité sur des caractères non retenus, mais qui peuvent avoir des retombées

économiques, doit être réduite. Par exemple, améliorer le calibre des fruits n'a peut-être qu'un intérêt limité si leur forme et leur couleur sont très variables.

La variabilité au sein de chaque variété botanique de goyavier-fraise (*cattleianum* et *lucidum*) est limitée à la Réunion, probablement grâce à une introduction récente (200 ans), et à une base génétique introduite réduite. Au Brésil, dans sa zone d'origine, il existe pour chaque variété botanique une variabilité sur des caractères végétatifs et pomologiques en fonction de la zone d'origine du matériel végétal (Raseira et Raseira, 1996). Dans des situations où il existe une forte variabilité liée par exemple à des populations d'origines différentes, une phase préalable de sélection de populations peut être nécessaire. Les espèces sauvages exotiques seraient ainsi plus prédisposées à une approche agronomique de la domestication sans sélection préalable. On peut aussi supposer que c'est le cas d'espèces à multiplication végétative dominante pour lesquelles la variabilité spontanée peut être réduite.

De même, si des contraintes phytosanitaires fortes sont présentes sur l'espèce à domestiquer, une sélection préalable pour rechercher des résistances est nécessaire. L'intérêt d'une telle espèce apparaît toutefois limité.

Une approche de type élaboration du rendement permet de répondre au manque de données sur l'espèce, d'apporter sur un pas de temps relativement court une connaissance globale de la façon dont se construisent les caractères d'intérêt, et d'identifier les étapes et facteurs limitants. Afin de déterminer les étapes à considérer pour l'élaboration du rendement, une connaissance des facteurs affectant les caractères d'intérêt, ou des hypothèses fortes sur leur nature, sont nécessaires. Les facteurs sur lesquels l'agriculteur peut agir pour affecter les caractères d'intérêt sont précisés. Ces résultats permettent d'une part d'évaluer le comportement de l'espèce en culture, et d'autre part de tester et de proposer des techniques culturales allant dans le sens des objectifs recherchés sur les caractères d'intérêt. Des avancées rapides peuvent ainsi être réalisées.

La variabilité génétique se traduit par une variabilité entre plants sur certains processus et/ou sur les caractères d'intérêt. Elle peut toutefois poser un problème si elle est trop forte, par exemple sur le nombre de fruits par plant, ou sur le calibre maximal du fruit, car un objectif de l'agriculteur est de limiter cette variabilité. Une démarche de sélection peut alors être engagée sur la base de critères pertinents par rapport à l'élaboration des caractères d'intérêt.

Cependant, la variabilité génétique peut entraîner des difficultés dans l'objectif de représenter l'élaboration des caractères d'intérêt à travers un modèle. Pour la plupart des espèces fruitières, le calibre est un caractère d'intérêt. La croissance du fruit est affectée par le nombre de fruits sur le plant et parfois par le nombre de graines pour les fruits polyspermes. On doit alors considérer la floraison et/ou la fécondation, étapes où se construisent en partie ces facteurs, et qui sont sous un contrôle génétique marqué. Par exemple, les études sur la biologie de la reproduction d'espèces sauvages montrent des effets plants marqués sur différents processus de la fécondation (Snow, 1982; Sedgley *et al.*, 1989; Ellis et Sedgley, 1993; Kimura *et al.*, 1997). De même, l'asynchronisme de floraison est marqué au sein de populations sauvages (Primack, 1980; Bawa, 1983; Pors et Werner, 1989; Carthew, 1993). L'élaboration du nombre de fleurs, qui n'a pas été traité dans le cadre de ce travail, peut également être sous l'effet d'un déterminisme génétique. A part dans des cas particuliers, comme par exemple l'autocompatibilité complète, la floraison, la fécondation, et leur lien la pollinisation, poseront des problèmes de modélisation dont les solutions dépendront des caractéristiques propres à chaque espèce pour ces étapes. Afin de mener à bien un tel travail,

une sélection préalable pour travailler sur du matériel végétal homogène est souhaitable. Par rapport à la typologie de Leakey et Simons (1998), la recherche d'un modèle d'élaboration du rendement place plutôt la stratégie vers la domestication en vue d'une culture intensive, où la sélection est un moyen de réaliser des avancées. Une démarche conjointe sélection/modélisation apparaît cohérente par rapport à cette classification.

Une modélisation ne concernant que la croissance du fruit peut toutefois amener à des applications pratiques si l'effet des facteurs environnementaux et de croissance est prédominant sur l'effet de facteurs génétiques. Les facteurs de croissance relatifs au fruit et au plant sont alors mesurés sur le terrain et ces valeurs sont utilisées comme entrées du modèle. La prise en compte de leur variabilité est indispensable car elle résulte en partie de la variabilité génétique sur les étapes précédant celle de la croissance, et des modèles stochastiques sont particulièrement indiqués.

Ces considérations sur la domestication par une approche agronomique nous conduisent à préciser le champ des connaissances minimales pour évaluer et orienter une telle démarche pour une espèce fruitière sauvage dont l'intérêt agricole et économique a déjà été identifié:

- variabilité des caractères d'intérêt retenus et des caractères économiques (intérêt de la démarche agronomique pour cette espèce; besoin d'une sélection préalable)
- cycle phénologique de la plante en milieu naturel (acquisition rapide de données sur le cycle et identification de certains facteurs qui l'affectent)
- facteurs affectant les caractères d'intérêt (point de départ de l'approche de l'élaboration de ces caractères)
- biologie de la reproduction (connaissance nécessaire car effet génétique marqué)

Il est peu probable que ces connaissances soient présentes dans la littérature compte tenu du caractère sauvage de l'espèce, et des études et expérimentations doivent être engagées pour traiter de ces aspects.

2- Généralisation des modèles proposés

La généralisation des modèles proposés est leur application à d'autres espèces et/ou d'autres conditions de culture, avec bien sûr les contraintes propres aux modèles statistiques présentées précédemment. Les approches utilisées pour les modèles de floraison et de croissance du fruit sont suffisamment générales pour permettre de les appliquer à d'autres espèces, à condition de vérifier les hypothèses posées pour la construction de chaque modèle. Les différents types de floraison pouvant être simulés par le modèle de floraison en sont une bonne illustration.

L'approche descriptive statistique qui a conduit au modèle de fécondation-fructification était justifiée par le manque d'hypothèse pour expliquer la forme des distributions du nombre de graines par fruit, et le modèle n'est pas généralisable. Par contre, la méthodologie consistant à construire un modèle pour aider à analyser des résultats issus de mécanismes complexes et peu connus peut être utilisée dans d'autres situations. D'une façon générale, la fécondation est un ensemble de processus complexes, rarement connus entièrement, et dont le déroulement et les facteurs qui les affectent sont propres à l'espèce. Elle est étroitement liée au mode de pollinisation, lui-même dépendant de l'espèce. La modélisation des étapes de pollinisation,

fécondation et fructification doit donc se construire à partir d'hypothèses considérant les caractéristiques propres de l'espèce considérée.

Littérature citée

(y compris celle citée dans les annexes 1, 2 et 3)

- [1] Agostini, D. (1995). *La floraison du kiwi (Actinidia deliciosa cv. Hayward): analyse de la variabilité et simulation par un modèle stochastique*. Thèse, Université Claude Bernard - Lyon I, Lyon, France.
- [2] Akamine, E.K. et Goo, T. (1979). Respiration and ethylene production in fruits of species and cultivars of *Psidium* and species of *Eugenia*. *Journal of the American Society for Horticultural Science*, 104, 632-635.
- [3] Anonyme, 1981. Pasture-range weed problems in Hawaii. *Newsletter of the Asian-Pacific Weed Science Society*, 1-2.
- [4] Arnold, C. Y. (1959). The determination and significance of the base temperature in a linear heat unit system. *Proceedings of the American Society for Horticultural Science*, 74, 430-445.
- [5] Atchison, E. (1947). Chromosome numbers in the Myrtaceae. *American Journal of Botany*, 34, 159-164.
- [6] Atkins, T.A. (1990). Using crop loading models to predict orchard profitability. *Acta Horticulturae*, 276, 363-370.
- [7] Baborun, T. (1998). Substances naturelles actives: la flore mauricienne, une source d'approvisionnement potentielle. *Proceedings of the Second Annual Meeting of Agricultural Scientists*, 12-13 Aug. 1997, Réduit, Mauritius, Lalouette, J.A., Bachraz, D.Y., Sukurdeep, N. et Seebaluck, B.D. (eds), Food and Agricultural Research Council, Réduit, Maurice, 83-94.
- [8] Baker, R.P., Hasenstein, K.H. et Zavada, M.S. (1997). Hormonal changes after compatible and incompatible pollination in *Theobroma cacao* L.. *HortScience*, 32, 1231-1234.
- [9] Barone, E., Gullo, G., Zappia, R. et Inglese, P. (1994). Effect of crop load on fruit ripening and olive oil (*Olea europea* L.) quality. *Journal of Horticultural Science*, 69, 67-73.
- [10] Bawa, K.S. (1983). Patterns of flowering in tropical plants. In: Jones CE, Little RJ, eds. *Handbook of experimental pollination*. Scientific and Academic Editions, New York, USA, 394-410.
- [11] Becwar, M.R., Stanwood, P.C. et Leonhardt, K.W. (1983). Dehydration effects on freezing characteristics and survival in liquid nitrogen of desiccation-tolerant and desiccation-sensitive seeds. *Journal of the American Society for Horticultural Science*, 108, 613-618.
- [12] Bittenbender, H.C. et Kobayashi K.D. (1990). Predicting the harvest of cycled 'Beaumont' guava. *Acta Horticulturae*, 269, 197-204.

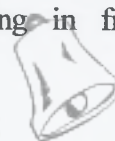
- [13] Bonhomme, R. (2000). Bases and limits to using 'degree.day' units. *European Journal of Agronomy*, 13, 1–10.
- [14] Brain, P. et Landsberg, P.P. (1981). Pollination, initial fruit set and fruit drop in apples: analysis using mathematical models. *Journal of Horticultural Science*, 56, 41–54.
- [15] Broschat, T.K. et Donselman, H.M. (1983). Effect of photoperiod on growth of seven species of tropical and subtropical ornamental trees. *HortScience*, 18, 571.
- [16] Brown, B.I. et Wills, R.B.H. (1983). Post-harvest changes in guava fruit of different maturity. *Scientia Horticulturae*, 19, 237–243.
- [17] Brun, L.O. et Chazeau, J. (1986). *Catalogue des ravageurs d'intérêt agricole de Nouvelle-Calédonie*. Orstom, Paris, France.
- [18] Bustan, A., Goldschmidt, E.E. et Erner, Y. (1996). Integrating temperature effects on fruit growth into a Citrus productivity model. *Proceedings of the International Society of Citriculture*, vol. 2, VIII International Citrus Congress, 12-17 May 1996, Sun City, South Africa, 938–944.
- [19] Cadet, L.J.T. (1980). *La végétation de l'île de la Réunion. Etude phytoécologique et phytosociologique*. Cazal, Saint Denis de la Réunion, France.
- [20] Carthew, S.M. (1993). Patterns of flowering and fruit production in a natural population of *Banksia spinulosa*. *Australian Journal of Botany*, 41, 465–480.
- [21] Casagrande Junior, J.G., Bianchi, V.J., Strelow, E.Z., Bacarin, M.A. et Fachinello, J.C. (1999). Influência do sombreamento sobre os teores de carboidratos e fenóis em estacas semilenhosas de araçazeiro. *Pesquisas Agropecuarias Brasileira*, 34, 2219–2223.
- [22] Cecy, C. et Yassumoto, Y. (1973). Observações sobre a fitoterapia popular nos arredores de Curitiba. III. *Tribuna Farmaceutica*, XLI (1-2), 14–22.
- [23] Chezhiyan, N. (1988). Stigma receptivity, flower shedding, flower abnormality and pollination studies in *Psidium* sp.. *Madras Agricultural Journal*, 75, 29–32.
- [24] Chezhiyan, N. et Shanker, C. (1983). Fruit setting, parthenocarpy, flower and fruit drop in guava and its relatives. *Progressive Horticulture*, 15, 195–199.
- [25] Clement, C.R. et da Silva Filho, D.F. (1994). Amazonian small fruits with commercial potential. *Fruit Varieties Journal*, 48, 152–158.
- [26] Cuadra, R. et Quincosa, A. (1982). Comportamiento de diferentes especies de *Psidium* como patrones para guayabos resistentes a *Meloidogyne* (Nematoda: *Heteroderidae*). *Ciencias de la Agricultura*, 13, 19–26.
- [27] Cuevas, J. et Polito, V.S. (1997). Compatibility relationships in 'Manzanillo' olive. *HortScience*, 32, 1056–1058.

- [28] Dacünha-Castelle, D. et Duflo, M. (1982). *Probabilités statistiques. I. Problèmes à temps fixe*. Masson, Paris, France.
- [29] Dagnelie, P. (1998a). *Statistique théorique et appliquée. tome 1, inférence statistique à une et à deux dimensions*. De Boeck & Larcier S.A., Brussels, Belgium.
- [30] Dagnelie, P. (1998b). *Statistique théorique et appliquée. tome 2, inférence statistique à une et à deux dimensions*. De Boeck & Larcier S.A., Brussels, Belgium.
- [31] Dattee, Y. (1972). Hérité quantitative du nombre d'ovules par ovaire dans quelques familles de luzerne cultivée (*Medicago sativa* L.). *Annales de l'Amélioration des Plantes*, 22, 201–209.
- [32] DeGrandi-Hoffman, G., Roth, S.A. et Loper, G.M. (1989). ALMOPOL: A cross-pollination and nut set simulation model for almond. *Journal of the American Society for Horticultural Science*, 114, 170–176.
- [33] Delph, L.F. (1986). Factors regulating fruit and seed production in the desert annual *Lesquerella gordonii*. *Oecologia* (Berlin), 69, 471–476.
- [34] Devine, W.T. (1977). A programme to exterminate introduced plants on Raoul Island. *Biological Conservation*, 11, 193–207.
- [35] Diong, C.H. (1998). The strawberry guava *Psidium cattleianum* Sabine 1821. *The Malaysian Naturalist*, 52(1), 8–10.
- [36] Do Ngoc, L., Brillard, M., Hoebeke, J. et Aucouturier, P. (1995). Nouvelle chaîne α de jacaline chez deux espèces sauvages de jaquier. *Comptes Rendus de l'Académie des Sciences de Paris*, 318, 167–172.
- [37] Dodds, P.N., Clarke, A.E. et Newbigin, Ed. (1996). A molecular perspective on pollination in flowering plants. *Cell*, 85, 141–144.
- [38] Doyle, C.J., Moore, W.B. et Henzell, R.F. (1989). Modelling the economic consequences of potential management changes in a mature kiwifruit orchard in New Zealand. *Agricultural Systems*, 31, 321–347.
- [39] Dufault, R.J. (1997). Determining heat units requirements for broccoli harvest in coastal South Carolina. *Journal of the American Society for Horticultural Science*, 122, 169–174.
- [40] Dulai, S., Molnar, I., Lehocski, E. et Pocs, T. (1998). The role of photosynthetic activity in the vulnerability of an insular biome to invasion by alien species. In: *Photosynthesis: mechanisms and effects*, Garab, G. (ed), Kluwer Academic Publishers, Netherland, V, 4057–4060.
- [41] Ehlenfeldt, M.K. (2001). Self- and cross-fertility in recently released highbush blueberry cultivars. *HortScience*, 36, 133–135.

- [42] Ellis, M.F. et Sedgley, M. (1993). Gynodioecy and male sterility in *Eucalyptus leucoxylon* F. Muell. (Myrtaceae). *International Journal of Plant Science*, 154, 314–324.
- [43] Etienne, J. (1982). *Etude systématique, faunistique et écologique des Tephritidae de la Réunion*. Thèse, Ecole Pratique des Hautes Etudes, Paris, France.
- [44] Falque, M., Vincent, A., Vaissière, B.E. et Eskes, A.B. (1995). Effect of pollination intensity on fruit set and seed set in cacao (*Theobroma cacao* L.). *Sexual Plant Reproduction*, 8, 354–360.
- [45] Falque, M., Lesdalons, C. et Eskes, A.B. (1996). Comparison of two cacao (*Theobroma cacao* L.) clones for the effect of pollination intensity on fruit set and seed content. *Sexual Plant Reproduction*, 9, 221–227.
- [46] Fleisch, H. et Bartholomew, D.P. (1987). Development of a heat unit model of pineapple ('Smooth Cayenne') fruit growth from field data. *Fruits*, 42, 709–715.
- [47] Fleischmann, K. (1997). Invasion of alien woody plants on the island of Mahé and Silhouette, Seychelles. *Journal of Vegetation Science*, 8, 5–12.
- [48] Fosberg, F.R. (1941). Varieties of the strawberry guava. *Proceedings of the Biological Society of Washington*, 54, 179–180.
- [49] Fouqué, A. (1972). Espèces fruitières d'Amérique tropicale: Famille des Myrtacées. *Fruits*, 27, 120–139.
- [50] Fusillier, J.L. et Normand, F. (1999). *Note méthodologique sur l'évaluation micro-économique de cultures pérennes. Application d'un modèle de simulation au cas d'une plantation de goyavier-fraise*. Note CIRAD-TERA n°19/99, Cirad St Denis, La Réunion.
- [51] Galan, F. (1995). *Etude du système agraire de la commune de Saint-Benoit: évolutions historiques et actuelles, typologie, modélisations économiques*. Mémoire de DAA, Institut National Agronomique Paris Grignon, Paris, France.
- [52] Gebhardt, S.E., Cutrufelli, R. et Matthews, R.H. (1982). *Composition of foods. Fruits and fruit juices. Raw. Processed. Prepared*. Agriculture Handbook n° 8-9, USDA-HNIS, 144.
- [53] Génard, M. et Bruchou, C. (1993). A functional and exploratory approach to studying growth: the example of the peach fruit. *Journal of the American Society for Horticultural Science*, 118, 317–323.
- [54] Goldschmidt, E. E. et Golomb, A. (1982). The carbohydrate balance of alternate-bearing citrus trees and the significance of reserves for flowering and fruiting. *Journal of the American Society for Horticultural Science*, 107, 206–208.
- [55] Gurusamy, C. (1999). The pattern of seed formation in cauliflower (*Brassica oleracea* L. var. *botrytis*) with special reference to seed abortion. *Canadian Journal of Botany*, 77, 1189–1192.

- [56] Habib, R., Tisné-Agostini, D., Vannière, M.P. et Monestiez, P. (1991). Geostatistical method for independant sampling in kiwifruit vine to estimate yield components. *New Zealand Journal of Crop and Horticultural Science*, 19, 329–335.
- [57] Harrison, R.E., Luby, J.J. et Ascher, P.D. (1994). Pollen source affects yield components and reproductive fertility of four half-high blueberry cultivars. *Journal of the American Society for Horticultural Science*, 119, 84–89.
- [58] Hirano, R.T. et Nakasone, H.Y. (1969a). Chromosome numbers of ten species and clones in the genus *Psidium*. *Journal of the American Society for Horticultural Science*, 94, 83–86.
- [59] Hirano, R.T. et Nakasone, H.Y. (1969b). Pollen germination and compatibility studies of some *Psidium* species. *Journal of the American Society for Horticultural Science*, 94, 287–289.
- [60] Hof, L., Keiser, L.C.P., Elberse, I.A.M. et Dolstra, O. (1999). A model describing the flowering of single plants, and the heritability of flowering traits of *Dimorphotheca pluvialis*. *Euphytica*, 110, 35–44.
- [61] Hoyos, J.F. (1989). *Frutales en Venezuela (Nativos y exóticos)*. Sociedad de Ciencias Naturales La salle, monografía n° 26, Caracas, Venezuela.
- [62] Huenneke, L.F. et Vitousek, P.M. (1990). Seedling and clonal recruitment of the invasive tree *Psidium cattleianum*: implications for management of native Hawaiian forests. *Biological Conservation*, 53, 199–211.
- [63] Inglese, P., Barbera, G., La Mantia, T. et Portolano, S. (1995). Crop production, growth, and ultimate size of cactus pear fruit following fruit thinning. *HortScience*, 30, 227–230.
- [64] Jankiewicz, L.S., Colinas Leon, M.T. et Martinez Albores, V.M. (1986). Cupapé (*Cordia dodecandra* DC., Boraginaceae) a fruit tree in the process of domestication in Mexico. *Acta Societatis Botanicorum Poloniae*, 55, 115–127.
- [65] Jefferies, C.J. et Brain, P. (1984). A mathematical model of pollen-tube penetration in apple styles. *Planta*, 160, 52–58.
- [66] Jenni, S., Cloutier, D.C., Bourgeois, G. et Stewart, K.A. (1996). A heat unit model to predict growth and development of muskmelon to anthesis of perfect flowers. *Journal of the American Society for Horticultural Science*, 121, 274–280.
- [67] Jones, I.D. (1931). Preliminary report on relation of soil moisture and leaf area to fruit development of the Georgia Belle peach. *Proceedings of the American Society for Horticultural Science*, 28, 6–14.
- [68] Kimura, P.H., Okamoto, G. et Hirano, K. (1997). Flower types, pollen morphology and germination, fertilization, and berry set in *Vitis coignetiae* Pulliat. *American Journal of Enology and Viticulture*, 48, 323–327.

- [69] Kupzow, A.J. (1980). Theoretical basis of the plant domestication. *Theoretical and Applied Genetics*, 57, 65–74.
- [70] Lang, G.A. et Danka R.G. (1991). Honey-bee-mediated cross- versus self-pollination of 'Sharpblue' blueberry increases fruit size and hastens ripening. *Journal of the American Society for Horticultural Science*, 116, 770–773.
- [71] Lapiere, F. (1996). *Organisation de la production et de la mise en marché du goyavier (Psidium cattleianum) à la Réunion. Perspectives d'avenir*. Mémoire de DAA, Ecole nationale supérieure des Industries Alimentaires, Section Industrie alimentaires régions chaudes (Ensia-Siarc), Montpellier, France.
- [72] Lavergne, R. (1981). Goyave la suave. In: *Fleurs de Bourbon*, tome 3. Cazal, Saint-Denis de la Réunion, France.
- [73] Leakey, R.R.B. et Simons, A.J. (1998). The domestication and commercialization of indigenous trees in agroforestry for the alleviation of poverty. *Agroforestry Systems*, 38, 165–176.
- [74] Lehman, E.L. (1983). *Theory of point estimation*. Wiley and Sons, New York, USA.
- [75] Lescourret, F., Habib, R., Génard, M., Agostini, D. et Chadœuf, J. (1998). Pollination and fruit growth models for studying the management of kiwifruit orchards. I. Models description. *Agricultural Systems*, 56, 67–89.
- [76] Lorence, D.H. et Sussman, R.W. (1986). Exotic species invasion into Mauritius wet forest remnants. *Journal of Tropical Ecology*, 2, 147–162.
- [77] Luckwill, L.C., Weaver, P. et MacMillan, J. (1969). Gibberellins and other growth hormones in apple seeds. *Journal of Horticultural Science*, 44, 413–424.
- [78] MacDonald, I.A.W., Thébaud, C., Strahm, W.A. et Strasberg, D. (1991). Effects of alien plant invasions on native vegetation remnants on La Réunion (Mascarene Island, Indian Ocean). *Environmental Conservation*, 18, 51–61.
- [79] Maggs, D.H. (1985). Myrtaceae. In: *Tropical tree fruits for Australia*. Queensland Department of Primary Industries, Brisbane, Australia, 108–112.
- [80] Marcelis, L.F.M., Heuvelink, E. et Goudriaan, J. (1998). Modelling biomass production and yield of horticultural crops: a review. *Scientia Horticulturae*, 74, 83–111.
- [81] Mathsoft, (1999). *S-Plus® 2000 Guide to Statistics, volume 1*. Data Analysis Products Division, Mathsoft, Seattle, WA, USA.
- [82] Meynard, J.M. et Sébillote, M. (1994). L'élaboration du rendement du blé, base pour l'étude des autres céréales à talles. In *Un point sur...élaboration du rendement des principales cultures annuelles*, Combe, L., Picard, D. (coords.), INRA, Paris.
- [83] Monselise, S.P. et Goldschmidt, E.E. (1982). Alternate bearing in fruit trees. *Horticultural Review*, 4, 128–173.



- [84] Motooka, P., Nagai, G. et Ching, L. (1983). Cut-surface application of glyphosate to control tropical brush species. *Proceedings of the Meeting of the Weed Science Society of America*, 1983, Hawaii, USA, University of Hawaii, 96.
- [85] Motooka, P., Ching, L., Nagai, G., Powley, J. et Yamamoto, T. (1989). Control of Hawaii brush species with tebuthiuron. *Proceedings of the 12th Asian-Pacific Weed Science Society Conference*, Tapei, Taiïwan, 1, 203–206.
- [86] Munoz, C., Sepulveda, G., Garcia-Huidobro, J. et Sherman, W. B. (1986). Determining thermal time and base temperature required for fruit development in low-chilling peaches. *HortScience*, 21, 520–522.
- [87] Nachtigal, J.C. et Fachinello, J.C. (1995). Efeito de substratos e do acido indolbutirico no enraizamento de estacas de araçazeiro (*Psidium cattleianum* Sabine). *Revista Brasileira de Agrociencia*, 1, 34–39.
- [88] Nerd, A., Aronson, J.A. et Mizrahi, Y. (1990). Introduction and domestication of rare and wild fruit and nut trees for desert areas. *Advances in new crops: Proceedings of the First National Symposium NEW CROPS, research, development, economics*, 23-26 October 1988, Indianapolis, USA, Janick, J. et Simon, J.E. (Eds), Timber Press, Portland, USA, 355–363.
- [89] Nguyen, R., Poucher, C. et Brazzel, J.R. (1992). Seasonal occurrence of *Anastrepha suspensa* (Diptera: Tephritidae) in Indian River County, Florida, 1984-1987. *Journal of Economic Entomology*, 85, 813–820.
- [90] Nic Lughadha, E. et Proença, C. (1996). A survey of the reproductive biology of the Myrtoideae (Myrtaceae). *Annals of the Missouri Botanical Garden*, 83, 480–503.
- [91] Nik-Masdek, N.H. et Ahmad-Kamil, J. (1994). Response of *Psidium littorale* var. *longipes* (strawberry guava) to root-knot nematode (*Meloidogyne incognita*). *Proceedings of the 4th International Conference on Plant Protection in the Tropics*, 28-31 March 1994, Rajan, A. et Yusof, I. (eds), Malaysian Plant Protection Society, Kuala Lumpur, Malaisie, 435–436.
- [92] Normand, F. (1993). *Rapport Technique Annuel, 1993*. Programme d'expérimentation dans les Hauts de l'Est et du Nord-est, Cirad-Flhor, Saint-Pierre de la Réunion, France.
- [93] Normand, F. et Huet, P. (2000). *Fiche technique goyavier (2^{ème} version)*. Cirad-Flhor, Saint-Pierre de la Réunion, France.
- [94] Normand, F., Quilici, S., et Simiand C. (2000). Seasonal occurrence of fruit flies in strawberry guava (*Psidium cattleianum* Sabine) in Réunion Island: host phenology and fruit infestation. *Fruits*, 55, 271–281.
- [95] Normand, F. et Huet, P. (2001). *Rapport technique annuel 2000*. Programme d'expérimentation dans les Hauts humides, Cirad-Flhor, Saint-Pierre de la Réunion, France.

- [96] Paniandy, J-C. (1999). *Etude d'un fruit tropical: le goyavier-fraise (Psidium cattleianum Sabine). Caractérisation physico-chimique, biochimique et conservation en frais en vue de sa valorisation*. Thèse, Université de la Réunion, Saint Denis, France.
- [97] Paniandy, J-C., Normand, F. et Reynes, M. (1999). Facteurs intervenant sur la conservation en frais de la goyave-fraise à l'île de la Réunion. *Fruits*, 54, 49–56.
- [98] Pedrosa-Macedo, J.H. (2000). Biology and behavior of the strawberry guava sawfly, *Haplostegus epimelas* Konow 1901 (Hymenoptera: Pergidae), in Southern Brazil. *Proceedings of the Entomological Society of Washington*, 102, 129–134.
- [99] Popenoe, W. (1974). Fruits of the Myrtle family. In: *Manual of tropical and subtropical fruits, excluding the banana, coconut, pineapple, citrus fruits, olive and fig*. (publication originale: 1920, par Macmillan Publishing Co, New York) Hafner Press, New York., 272–311.
- [100] Pors, B., Werner, P.A. (1989). Individual flowering time in a goldenrod (*Solidago canadensis*): field experiment shows genotype more important than environment. *American Journal of Botany*, 76, 1681–1688.
- [101] Primack, R.B. (1980). Variation in the phenology of natural populations of montane shrubs in New Zealand. *Journal of Ecology*, 68, 849–862.
- [102] Proença, C. et Gibbs, P.E. (1994). Reproductive biology of eight sympatric Myrtaceae from Central Brazil. *New Phytologist*, 126, 343–354.
- [103] Quevedo, E. et Zuluaga, L. (1994). Una aproximación al proceso de domesticación de vegetales con referencia especial a los frutales en la Amazonia colombiana. *Agronomía Colombiana*, 11, 190–205.
- [104] Raseira, A. et Raseira, M.C.B. (1994). "Ya-cy", cultivar de araçazeiro lançada pela EMBRAPA/CPACT. *Horti Sul*, 3(1), 37–39.
- [105] Raseira, M.C.B. et Raseira, A. (1996). *Contribuição ao estudo do araçazeiro, Psidium cattleianum*. EMBRAPA-CPACT, Pelotas, Brazil.
- [106] Raunet, M. (1991). *Le milieu physique et les sols de l'île de la Réunion. Conséquences pour la mise en valeur agricole*. Centre de Coopération Internationale en Recherche Agronomique pour le Développement (Cirad), Montpellier, France.
- [107] Raw, A. (2000). Foraging behaviour of wild bees at hot pepper flowers (*Capsicum annum*) and its possible influence on cross pollination. *Annals of Botany*, 85, 487–492.
- [108] de Reffye, P., Parvais, J.P., Mossu, G. et Lucas, P. (1978). Influence des aléas de la pollinisation sur les rendements du cacaoyer. Modèle mathématique et simulation. *Café Cacao Thé*, 22, 251–274.

- [109] Reid, W. et Gast, K.L.B. (1993). The potential for domestication and utilization of native plums in Kansas. *Proceedings of the Second National Symposium on New Crops: Exploration, research and commercialization*, 6-9 October 1991, Indianapolis, USA, Janick, J. et Simon, J.E. (Eds), John Wiley and Sons Inc., New York, USA, 520-523.
- [110] Rodrigo, J. et Herrero, M. (1996). Evaluation of pollination as the cause of erratic fruit set in apricot 'Monique'. *Journal of Horticultural Science*, 71, 801-805.
- [111] Ruiz, R., Garcia-Luis, A., Monerri, C. et Guardiola, J.L. (2001). Carbohydrate availability in relation to fruitlet abscission in *Citrus*. *Annals of Botany*, 87, 805-812.
- [112] Salinero Corral, M.C. et Aguin Casal, O. (1993). Experiencias en cultivo de guayabo. *Fruticultura Profesional*, 54 (especial cultivos frutales alternativos), 60-63.
- [113] Salinero Corral, M.C. et Aguin Casal, O. (1996). Estudio en plantacion de *Feijoa sellowiana*, *Psidium guayaba* y *Psidium cattleianum*. *Nuevos cultivos, nuevos usos, nuevas alternativas*, ITEA, 17, 215-228.
- [114] Sanzol, J. et Herrero, M. (2001). The "effective pollination period" in fruit trees. *Scientia Horticulturae*, 90, 1-17.
- [115] Schneider, D., Stern, R.A., Eisikowitch, D. et Goldway, M. (2001). Determination of the self-fertilization potency of 'Golden Delicious' apple. *Journal of Horticultural Science & Biotechnology*, 76, 259-263.
- [116] Schultes, R.E. (1989). *Pouteria ucuqui* (Sapotaceae), a little-known amazonian fruit tree worthy of domestication. *Economic Botany*, 43 125-127.
- [117] Scott, A.J. (1990). Myrtacées. In: *Flore des Mascareignes, La Réunion, Maurice, Rodrigues*. Tome 92, ORSTOM, Paris, France.
- [118] Sébillotte, M. (1990). Système de culture, un concept opératoire pour les agronomes. In: *Un point sur ... Les systèmes de culture*, Combe, L. et Picard, D. (coord.), INRA, Paris, France, 165-196.
- [119] Sedgley, M. et Smith R.M. (1989). Pistil receptivity and pollen tube growth in relation to the breeding system of *Eucalyptus woodwardii* (*Symphyomyrtus*: Myrtaceae). *Annals of Botany*, 64, 21-31.
- [120] Sedgley, M., Hand, F.C., Smith R.M. et Griffin, A.R. (1989). Pollen tube growth and early seed development in *Eucalyptus regnans* F. Muell. (Myrtaceae) in relation to ovule structure and preferential outcrossing.. *Australian Journal of Botany*, 37, 397-411.
- [121] Seth, J.N. (1960). Varietal cross-incompatibility in guava (*Psidium guajava* L.). *Horticultural Advance*, 4, 161-164.
- [122] Seth, J.N. (1970). Storage of pollen-grains in *Psidium* species. *Indian Journal of Agricultural Science*, 40, 741-745.

- [123] Shinozaki, K., Yoda, K., Hozumi, K. et Kira, T. (1964a). A quantitative analysis of plant form. The pipe model theory. I. Basic analysis. *Japanese Journal of Ecology*, 14, 97-105.
- [124] Shinozaki, K., Yoda, K., Hozumi, K. et Kira, T. (1964b). A quantitative analysis of plant form. The pipe model theory. II. Further evidence of the theory and its application in forest ecology. *Japanese Journal of Ecology*, 14, 133-139.
- [125] Shore, J.S. et Barrett, S.C.H. (1984). The effect of pollination intensity and incompatible pollen on seed set in *Turnera ulmifolia* (Turneraceae). *Canadian Journal of Botany*, 62, 1298-1303.
- [126] Singh, U.R., Pandey, I.C., Upadhyay, N.P. et Tripathi, B.M. (1976). Effect of different rootstocks on the growth, yield and quality of guava. *The Punjab Horticultural Journal*, 16, 121-124.
- [127] Singhal, V.K., Gill, B.S. et Bir, S.S. (1985). Cytology of woody species. *Proceedings of the Indian Academy of Science (Plant Science)*, 94, 607-617.
- [128] Smith-White, S. (1948). Cytological studies in the Myrtaceae I. Chromosome numbers in the Leptospermoideae and Myrtoideae. *Proceedings of the Linnean Society of New South Wales*, 73, 16-36.
- [129] Snow, A.A. (1982). Pollination intensity and potential seed set in *Passiflora vitifolia*. *Oecologia*, 55, 231-237.
- [130] Sprent, P. (1989). *Applied nonparametric statistical methods*. Chapman & Hall, London, United Kingdom.
- [131] Srivastava, H.C. (1978). Cytogenetic studies of *Psidium cattleianum* Sabine. *National Academy Science Letters*, 1, 2.
- [132] Stover, E., Wirth, F. et Robinson, T. (2001). A method for assessing the relationship between crop load and crop value following fruit thinning. *HortScience*, 36, 157-161.
- [133] Subramanyam, M. D., Dinesh, M. R. et Braganza, M. (1992). Varietal evaluation and floral biology studies in the genus *Psidium*. *Acta Horticulturae*, 321, 211-219.
- [134] Sullivan, D.T. (1965). The effect of time of bloom of individual flowers on the size, shape and maturity of apple fruits. *Proceedings of the American Society for Horticultural Science*, 87, 41-46.
- [135] Sweet, C. (1986). Guavas: new popularity reported. *California Grower*, 10(10), 32-33.
- [136] Teatitia, S.S., Phogat, K.P.S. et Srivastava, V.S. (1970). Blossom biology studies in *Psidium* species. *Progressive Horticulture*, 2(3), 101-112.
- [137] Teatitia, S.S. et Phogat, K.P.S. (1971). Effect of rootstocks on growth, yield and quality in guava (*Psidium guajava*). *Progressive Horticulture*, 2(4), 37-45.

- [138] Vaissière, B.E., Malabœuf, F. et Rodet, G. (1996a). Viability of Cantaloupe pollen carried by honeybees *Apis mellifera* varies with foraging behavior. *Naturwissenschaften*, 83, 84-86.
- [139] Vaissière, B.E., Rodet, G., Cousin, M., Botella, L. et Torrè Grossa J-P. (1996b). Pollination effectiveness of honey bees (Hymenoptera: Apidae) in kiwifruit orchard. *Horticultural Entomology*, 89, 453-461.
- [140] Vargas, R.I., Harris, E.J. et Nishida, T. (1983). Distribution and seasonal occurrence of *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) on the island of Kauai in the Hawaiian Islands. *Environmental Entomology*, 12, 303-310.
- [141] Vargas, R.I., Stark, J.D. et Nishida, T. (1990). Population dynamics, habitat preference, and seasonal distribution patterns of oriental fruit fly and melon fly (Diptera: Tephritidae) in an agricultural area. *Environmental Entomology*, 19, 1820-1828.
- [142] Vogel, R. (1982). L'acclimatation en Corse de quelques fruitiers exotiques. *Fruits*, 37(3), 157-165.
- [143] Voltolini, J.A. et Fachinello, J.C. (1997). Effect of shading cattley guava stock plant (*Psidium cattleianum* Sabine) on propagation by cuttings. *Acta Horticulturae*, 452, 59-62.
- [144] Volz, R.K., Palmer, J.W. et Gibbs, H.M. (1995). Within-tree variability in fruit quality and maturity for 'Royal Gala' apple. *Acta Horticulturae*, 379, 67-74.
- [145] Wagenvoort, W.A., van de Vooren, J.G. et Brandenburg, W.A. (1989). Plant domestication and the development of sea starwort (*Aster tripolium* L.) as a new vegetable crop. *Acta Horticulturae*, 242, 115-122.
- [146] Wang, J.Y. (1960). A critique of the heat unit approach to plant response studies. *Ecology*, 41, 785-790.
- [147] Wielgolaski, F. E. (1999). Starting dates and basic temperatures in phenological observations of plants. *International Journal of Biometeorology*, 42, 158-68.
- [148] Wikler, C., Vitorino, M.D. et Pedrosa-Macedo, J.H. (1994). Projeto de controle biologico do araçazeiro - *Psidium cattleianum* Sabine, 1820. *Agrarias, Curitiba*, 13(1-2), 247-253.
- [149] Wikler, C., Smith, C.W. et Pedrosa-Macedo, J.H. (1996). The stem-gall wasp *Eurytoma* sp. (Hymenoptera: Eurytomidae): a potential biological control agent against *Psidium cattleianum*. *Proceedings of the IX International Symposium on Biological Control of Weeds*, 19-26 Jan. 1996, Stellenbosch, Moran, V.C. et Hoffmann, J.H. (Eds), University of Cape Town, Cape Town, Afrique du sud, 219-221.
- [150] Wilson, B.J. et Nishimoto, R.K. (1975). Ammonium sulfate enhancement of picloram absorption by detached leaves. *Weed Science*, 23, 297-301.

- [151] Wilson, C.W. (1980). Guava. In: *Tropical and subtropical fruits*, Nagy, S. et Shaw, P.E. (Eds), Westport, USA, 279–299.
- [152] von Witzke, S. et Röbellen, G. (1988). Domestication of *Cuphea*, a potential oil crop plant for industrial uses. *Proceedings du Colloque sur les Cultures Alternatives non Excédentaires*, Barbera, G. (Ed), 21 au 24 juillet 1987, Milazzo, Italie, Office des Publications Officielles des Communautés Européennes, 46–56.
- [153] Wutscher, H.K. et Shull, A.V. (1975). Subtropical fruits in South Texas. *Proceedings of the Florida State Horticultural Society*, 1974, 394–400.
- [154] Yang, S., Logan, J. et Coffey, D.L. (1995). Mathematical formulae for calculating the base temperature for growing degree days. *Agricultural and Forest Meteorology*, 74, 61–74.

Annexes

Annexe 1: Relations entre la température et la durée des intervalles fertilisation - demi floraison et demi floraison - demi récolte

Annexe 2: Détermination du nombre d'ovules dans les fleurs

Annexe 3: Calcul de la vraisemblance d'un déroulement de floraison observé

Annexe 4: Tableaux de présentation synthétique des modèles

Annexe 5: Planches photos

Annexe 1: Relations entre la température et la durée des intervalles fertilisation - demi floraison et demi floraison - demi récolte

La prévision de la période de récolte d'une parcelle est un élément important pour préparer au mieux le chantier de récolte et la commercialisation. Cela est particulièrement vrai pour la culture du goyavier-fraise dont la récolte nécessite une main-d'œuvre importante (environ 588 heures pour 10 t de fruits récoltés (Normand et Huet, 2000)), et l'agriculteur doit être très disponible à cette période et/ou prévoir de faire appel à de la main-d'œuvre extérieure. De plus, c'est un fruit fragile dont la durée de conservation après récolte est courte (Paniandy *et al.*, 1999), ce qui implique une commercialisation rapide et donc des débouchés préparés.

Le développement des pousses, des boutons floraux et du fruit du goyavier-fraise est très lié à la température (voir chapitre II-I). On peut donc s'attendre à une grande variabilité dans la durée des cycles de production en fonction de l'altitude (gamme altitudinale de culture potentielle du goyavier-fraise à la Réunion: 0 - 1200 m) et de la saison sur un site donné pour les cycles déclenchés par un apport d'azote. Une prédiction basée sur un nombre de jours depuis la floraison apparaît alors inappropriée, et la recherche de relations entre la durée du cycle de production et la température est indispensable pour construire un modèle de prévision de la période de récolte basé sur les températures. Les modèles thermiques construits à l'aide des données du suivi phénologique (Chapitre II-I) ne sont pas adaptés pour ces prévisions car ils ont été paramétrés en conditions naturelles et non en parcelles cultivées. D'autre part, ils ne concernent pas l'intervalle entre l'apport d'azote (fertilisation) et la floraison pour les cycles déclenchés.

Le cycle de production du goyavier-fraise peut être divisé en deux phases: fertilisation-floraison et floraison-récolte. Pour un cycle de production naturelle, seule cette dernière partie est considérée. A l'échelle d'une parcelle, la date de demi-floraison peut être choisie comme date de floraison. De même, la date de demi-récolte peut être retenue comme date de récolte. On se propose de construire des modèles thermiques permettant de prévoir la date de demi-récolte à partir de la date de fertilisation ou de demi-floraison, basés sur des relations entre les durées des intervalles fertilisation-demi floraison et demi floraison-demi récolte (appelés par la suite fertilisation-floraison et floraison-récolte respectivement) et les températures. Ces relations seront également utiles pour définir des échelles de temps thermique, ou temps de développement, pour les modèles de floraison et de croissance du fruit construits dans le cadre de cette thèse.

1- Matériels et méthodes

1.1- Le modèle thermique

Les relations entre durée de développement et température utilisent généralement le système des sommes de températures (Wang, 1960; Bonhomme, 2000), avec l'hypothèse que la vitesse de développement d'un organe ou d'une plante varie linéairement avec la température. Sous cette hypothèse, la somme des températures moyennes journalières, diminuées d'un seuil appelé température de base (T_B), est constante au cours d'une même phase de développement. Soit, pour une phase de développement comprise entre les dates t_0 et t_f (n jours):

$$(1) \quad ST = \sum_{i=i_0}^{i=i_f} (T_{mi} - TB) = n \cdot (T_m - TB)$$

où: ST est la somme de températures au cours de la phase de développement,
 T_{mi} est la température moyenne au jour i ((mini + maxi)/2),
 T_m est la température moyenne au cours de la phase de développement (moyenne des températures moyennes journalières),
 TB est la température de base pour la phase de développement.

La relation (1) a été appliquée aux deux phases fertilisation-floraison et floraison-récolte du goyavier-fraise. TB est le paramètre à estimer pour chacune de ces phases.

1.2- Les données sur les cycles de production

Les observations ont été réalisées sur des parcelles situées à 200, 490, 515 et 890 m d'altitude dans l'est de l'île de la Réunion, et sur une parcelle située à 920 m dans le sud de l'île. Les dates de demi-floraison et de demi-récolte ont été relevées sur plusieurs cycles sur chaque parcelle de la façon suivante:

- date de demi-floraison: c'est la date du pic de floraison, déterminé visuellement en fonction de la densité de fleurs ouvertes, de fleurs passées et de boutons floraux. A cette date, 40 à 60 rameaux sont repérés sur 20 à 30 plants (2 rameaux par plant), pris au hasard dans la parcelle, de façon à avoir une distribution de boutons floraux, fleurs ouvertes et fleurs passées sur les pousses issues de ces rameaux proche de 25% - 50% - 25%, représentative d'une demi-floraison.

- date de demi-récolte: elle correspond à la date de demi-récolte des fruits portés par les rameaux repérés à la demi-floraison, c'est à dire lorsque 50% de ces fruits ont atteint leur pleine maturité (chute spontanée).

Pour les cycles déclenchés, la date de fertilisation est aussi relevée.

Sur chaque parcelle, les cycles suivis étaient:

- des cycles naturels,
- des cycles déclenchés commerciaux,
- une série de fertilisations mensuelles consistant en un apport d'azote chaque mois sur des plants différents et permettant l'obtention de données tout au long de l'année et dans des conditions contrastées de température.

Un taux de fructification moyen a été calculé pour chaque cycle à partir des nombres totaux de fruits et de fleurs portés par les rameaux repérés.

1.3- Les données thermiques

Les parcelles n'étaient pas équipées d'enregistreurs de température, et les températures moyennes journalières ont été déterminées en utilisant le gradient thermique quotidien calculé entre des stations de Météo France (voir chapitre II-I), sauf à 200 m d'altitude (Tableau 10).

L'approximation du gradient est acceptable compte tenu de la régularité de la pente entre les deux stations météo impliquées, de la localisation des parcelles à proximité de ce transect, et de la même exposition cardinale des parcelles et des stations météo.

Localisation de la parcelle		Méthode d'estimation de la température	Stations Météo France impliquées
commune	altitude (m)		
Sainte Rose	200	données brutes	Gros Piton Ste Rose (181 m)
Saint-Benoit	490	gradient	St Benoit (40 m) - Pl. des Palmistes (1025 m)
Saint-Benoit	515	gradient	St Benoit (40 m) - Pl. des Palmistes (1025 m)
Plaine des Palmistes	890	gradient	St Benoit (40 m) - Pl. des Palmistes (1025 m)
Saint-Joseph	920	gradient	Le Baril (130 m) - Grand Coude (1080 m)

Tableau 10: Localisation des parcelles suivies, méthode d'estimation de la température moyenne journalière et stations Météo France impliquées dans cette estimation.

1.4- Analyses statistiques

1.4.1- Estimation de la température de base

La température de base TB est un paramètre important dont dépend la qualité de prédiction du modèle, et différentes méthodes ont été proposées pour déterminer TB à partir de séries d'observations de durées de phases de développement dans des conditions thermiques contrastées (Arnold, 1959; Yang *et al.*, 1995). Ces méthodes ont été utilisées pour estimer la température de base pour chacune des phases fertilisation-floraison et floraison-récolte. La température de base recherchée est celle qui:

- minimise l'écart-type des sommes de températures d'une série de cycles,
- minimise l'écart-type, exprimé en jours, des sommes de températures d'une série de cycles,
- minimise le coefficient de variation des sommes de températures d'une série de cycles,
- annule le coefficient de régression entre les températures moyennes durant la phase de développement et les sommes de températures correspondantes,
- annule la vitesse de développement dans la relation linéaire entre la température moyenne au cours de la phase de développement et la vitesse de développement.

Pour les quatre premières méthodes, l'estimation de TB a été réalisée par un programme de minimisation de fonctions non linéaires du logiciel Splus 2000 (Mathsoft, 1999). Yang *et al.* (1995) ont traduit ces méthodes en formules donnant directement la température de base. Il a été vérifié la cohérence des résultats obtenus par minimisation et par ces formules.

Pour la dernière méthode, une relation linéaire est estimée entre la température moyenne durant la phase de développement et la vitesse de développement (l'inverse de la durée de la phase), et on en déduit la température de base recherchée qui annule la vitesse de développement.

L'erreur standard et l'intervalle de confiance à 95% de la température de base estimée ont été calculés par la méthode du bootstrap (1000 tirages). Un jackknife après le bootstrap a été utilisé pour identifier les observations ayant une forte influence sur le résultat.

1.4.2- Validation des modèles

Les températures de base ont été estimées sur les cycles naturels et déclenchés suivis du 17/06/92 (première fertilisation) au 06/04/98 (dernière récolte) sur les parcelles à 200, 490, 515, et 890 m d'altitude.

Une validation interne des modèles a été réalisée avec les données ayant servi à estimer les températures de base. Une validation externe a été réalisée avec des données indépendantes constituées des cycles de production suivis à partir du 11/07/97 (date de demi-floraison) sur les parcelles précédentes, et de 4 cycles de production observés sur une parcelle à 920 m d'altitude dans le sud de l'île. Chaque validation consiste à estimer la date de demi-floraison ou de demi-récolte à partir de la date de fertilisation ou de demi-floraison respectivement, et des températures observées. Pour les deux niveaux de validation, les écarts entre dates observées et dates estimées ont été calculés et des corrélations (τ de Kendall (Sprent, 1989)) ont été recherchées entre ces écarts et certaines caractéristiques des cycles: altitude, durée de la phase de développement, température moyenne au cours de la phase de développement, taux de fructification, mois de fertilisation.

Enfin, les modèles ont été testés, sur l'ensemble des cycles déclenchés, dans une utilisation courante pour prévoir la date de demi-récolte à partir de la date de fertilisation. Le calcul se déroule en deux étapes: 1) estimation de la date de demi-floraison à partir de la date de fertilisation, puis 2) estimation de la date de demi-récolte à partir de la date de demi-floraison estimée. Les écarts entre dates de récolte observées et estimées calculés par cette méthode sont comparés à ceux obtenus par une estimation de la date de demi-récolte à partir de la date de demi-floraison observée (et non plus estimée, c'est à dire que la date de récolte n'est estimée que depuis la date de floraison observée).

2- Résultats

Les caractéristiques des cycles de production utilisés pour l'estimation des températures de base pour les phases fertilisation-floraison et floraison-récolte et la validation interne, et pour la validation externe sont résumées dans le tableau 11.

Objectif des données	Phase ^x	Gamme alt. (m)	Tmoy ^y mini (°C)	Tmoy ^y maxi (°C)	Durée moyenne (j) de la phase (amplitude)	Nb de valeurs
Estimation et validation interne	fert-flo	200 - 890	14.6	25.3	82.8 (52 - 160)	76
	flo-rec	200 - 890	14.8	25.2	129.9 (99 - 184)	105
Validation externe	fert-flo	200 - 920	15.4	25.4	85.0 (36 - 137)	22
	flo-rec	200 - 920	15.0	24.8	134.2 (102 - 186)	35

^x: fert-flo: fertilisation-floraison; flo-rec: floraison-récolte

^y: température moyenne au cours de la phase de développement

Tableau 11: Caractéristiques des cycles de production utilisés pour l'estimation des températures de base et la validation interne, et pour la validation externe.

Les amplitudes des températures moyennes au cours des phases de développement sont larges, et les durées correspondantes sont très variables. La durée, en jours, de la phase fertilisation-floraison varie d'un facteur de presque 4, alors que la phase floraison-récolte varie d'un facteur à peine égal à 2.

La vitesse de développement pour les phases fertilisation-floraison et floraison-récolte peut être exprimée de façon linéaire en fonction de la température moyenne (Fig. 20). Il n'est pas observé d'aplatissement aux extrémités des nuages de points, ce qui indique que les températures moyennes minimales et maximales qui ont été rencontrées ne sont pas limitantes. Cependant, cela ne montre pas que des températures moyennes journalières limitant le développement n'aient pas été atteintes au cours de certains cycles.

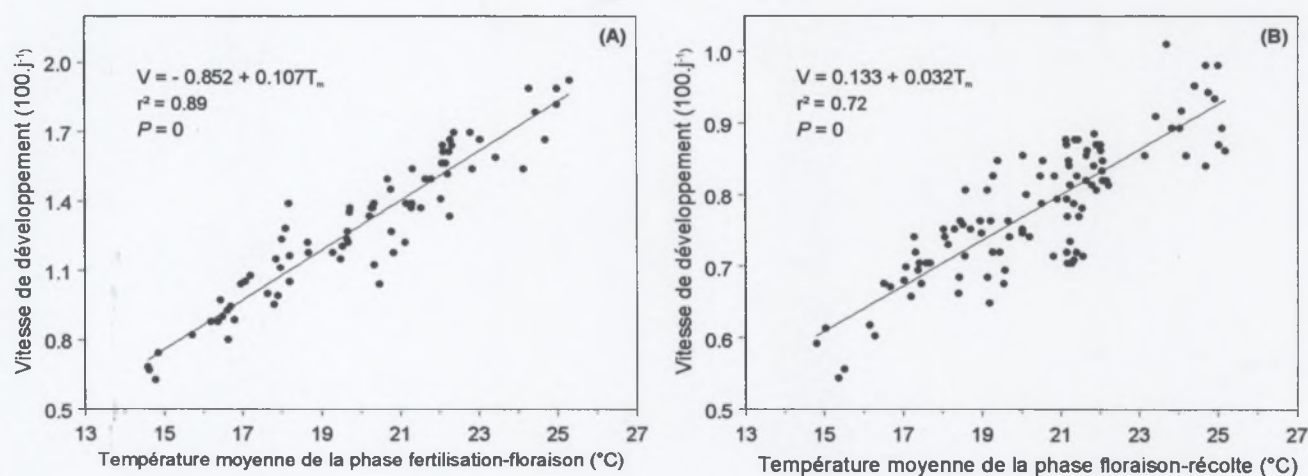


Figure 20: Relation entre la vitesse de développement et la température moyenne au cours des phases fertilisation-floraison (A) et floraison-récolte (B).

Pour chaque phase de développement, les températures de base estimées par différentes méthodes sont proches, sauf celle estimée par minimisation de l'écart-type des sommes de températures qui est plus élevée que les autres (Tableaux 12 et 13). Bonhomme (2000) signale que la valeur de la température de base dépend de la méthode avec laquelle elle est estimée. D'une façon générale, la méthode de minimisation de l'écart-type des sommes de températures donne des températures de base plus élevées que celles obtenues par les autres méthodes (Arnold, 1959; Yang *et al.*, 1995; Bonhomme, 2000). Yang *et al.* (1995) notent de plus que les coefficients de variation des sommes de températures et des durées de phase de développement sont plus élevés avec la température de base ainsi estimée qu'avec celles estimées par les autres méthodes et déconseillent ce mode de calcul.

Les erreurs-standards de la température de base estimée et l'amplitude des intervalles de confiance à 95% sont plus faibles pour la phase fertilisation-floraison que pour la phase floraison-récolte.

Méthode d'estimation (Arnold, 1959)	TB (°C)	s.e. (°C)	IC 95%
Minimisation de l'écart-type des ST^x	9.2	0.29	8.6 : 9.8
Minimisation de l'écart-type des ST^x en jours	8.3	0.36	7.6 : 9.0
Minimisation du coefficient de variation des ST^x	8.2	0.37	7.4 : 8.9
Annulation du coefficient de régression	8.1	0.37	7.3 : 8.7
Annulation de la vitesse de développement	7.9	0.44	6.8 : 8.7

^x: ST : sommes de température

Tableau 12: Température de base (TB), erreur standard (s.e.) et intervalle de confiance à 95% (IC95%) estimés pour la phase fertilisation-floraison par différentes méthodes de calcul.

Méthode d'estimation (Arnold, 1959)	TB (°C)	s.e. (°C)	IC 95%
Minimisation de l'écart-type des ST^x	2.7	0.94	0.6 : 4.3
Minimisation de l'écart-type des ST^x en jours	- 3.7	1.53	- 6.6 : - 0.8
Minimisation du coefficient de variation des ST^x	- 3.8	1.53	- 6.8 : - 0.9
Annulation du coefficient de régression	- 4.0	1.42	- 6.8 : - 1.2
Annulation de la vitesse de développement	- 4.2	1.43	- 7.0 : - 1.6

^x: ST : sommes de température

Tableau 13: Température de base (TB), erreur standard (s.e.) et intervalle de confiance à 95% (IC95%) estimés pour la phase floraison-récolte par différentes méthodes de calcul.

Nous retenons comme température de base la moyenne des valeurs estimées par les 4 dernières méthodes, soit 8.1°C pour la phase fertilisation-floraison et - 4°C pour la phase floraison-récolte. Les sommes de températures correspondantes sont respectivement 914 et 3134°C.j.

La validation interne du modèle donne des écarts entre valeurs observées et estimées qui varient de - 17 à + 25 jours (moyenne: 0.8 jours) pour la phase fertilisation-floraison, et de - 19 à + 20 jours (moyenne: - 0.4 jours) pour la phase floraison-récolte. Pour la phase fertilisation-floraison, les écarts sont corrélés faiblement, mais significativement, à la durée de la phase exprimée en jours. Ils ne sont pas corrélés à l'altitude, à la température moyenne, ni au mois de fertilisation (Tableau 14). Pour la phase floraison-récolte, les écarts sont corrélés de façon significative à l'altitude, au taux de fructification et à la durée de la phase; mais ils ne sont pas corrélés à la température moyenne (Tableau 15).

Une analyse de variance montre que les écarts moyens du modèle de la phase floraison-récolte pour les cycles fertilisés (moy = - 0.2 j) et naturels (moy = - 0.9 j) de la validation interne ne sont pas significativement différents ($P = 0.69$).

	Ecart observé- estimé	Altitude	Durée fert-flo	Température moyenne
Altitude	- 0.14 ^{ns}			
Durée fert-flo	0.24 ^{**}	0.27 ^{***}		
Température moyenne	- 0.04 ^{ns}	- 0.38 ^{***}	- 0.81 ^{***}	
Mois de fertilisation	- 0.04 ^{ns}	0.00 ^{ns}	0.10 ^{ns}	- 0.13 ^{ns}

^{ns} : non significatif à 5%; * : $P < 0.05$; ** : $P < 0.01$; *** : $P < 0.001$

Tableau 14: Corrélations (τ de Kendall) entre les écarts du modèle lors de la phase fertilisation-floraison en validation interne et des variables caractéristiques des cycles étudiés.

	Ecart observé- estimé	Altitude	Taux de fructification	Durée flo-rec
Altitude	0.18 ^{**}			
Taux de fructification	0.32 ^{***}	- 0.15 [*]		
Durée flo-rec	0.35 ^{***}	0.51 ^{***}	0.01 ^{ns}	
Température moyenne	0.02 ^{ns}	- 0.45 ^{***}	0.16 [*]	- 0.63 ^{***}

^{ns} : non significatif à 5%; * : $P < 0.05$; ** : $P < 0.01$; *** : $P < 0.001$

Tableau 15: Corrélations (τ de Kendall) entre les écarts du modèle lors de la phase floraison-récolte en validation interne et des variables caractéristiques des cycles étudiés.

La validation externe du modèle donne des écarts entre valeurs observées et estimées qui varient de - 18 à + 14 jours (moyenne: 0.6 jour) pour la phase fertilisation-floraison, et de - 11 à + 24 jours (moyenne: 2.9 jours) pour la phase floraison-récolte. Une analyse de variance montre que les écarts moyens du modèle de la phase floraison-récolte en validation externe pour les cycles fertilisés (moy = 0.6 j) et naturels (moy = 6.3 j) ne sont pas significativement différents ($P = 0.07$), mais la probabilité est proche du seuil de 5%.

Les distributions des écarts en validation externe et interne ne sont pas significativement différentes: tests de Kolmogorov-Smirnov, $P=0.95$ pour fertilisation-floraison et $P=0.23$ pour floraison-récolte. Pour la première phase de développement, 86 à 88% des écarts se trouvent entre - 10 et + 10 jours. Pour la seconde, 74 à 77% des écarts se trouvent dans cet intervalle.

En validation externe, pour la phase fertilisation-floraison, les écarts sont corrélés significativement à la durée de la phase exprimée en jours ($\tau = 0.53$) et à la température moyenne ($\tau = - 0.41$). Ils ne sont pas corrélés à l'altitude ($\tau = 0.23$), ni au mois de fertilisation ($\tau = 0.07$) (données non montrées). Pour la phase floraison-récolte, les écarts sont corrélés de façon significative à la durée de la phase ($\tau = 0.44$), à l'altitude ($\tau = 0.37$), et au taux de fructification ($\tau = 0.32$), mais ils ne sont pas corrélés à la température moyenne ($\tau = - 0.10$) (données non montrées).

Dans le test d'utilisation du modèle avec l'ensemble des cycles déclenchés, les écarts entre dates de récolte observées et estimées en deux étapes depuis la fertilisation varient de -25 à +29 jours (moyenne: - 0.1 jour), et 58% des écarts se trouvent entre - 10 et + 10 jours. La différence (y) de ces écarts avec ceux sur la date de récolte estimée à partir de la date de demi-floraison observée est fortement expliquée par les écarts entre les dates de demi-floraison observée et estimée (x): $y = - 0.58 + 0.95 \cdot x$; $n=95$; $r^2=0.97$; $P=0$.

3- Discussion

L'utilisation de parcelles à des altitudes différentes, représentatives de la zone de culture du goyavier-fraise à la Réunion, et la possibilité de déclencher des cycles de production par un apport d'azote nous ont permis d'obtenir en quelques années un jeu de données conséquent sur les durées des phases fertilisation-floraison et floraison-récolte dans des conditions contrastées de température (Tableau 11). La variabilité des durées observées confirme l'impossibilité d'utiliser un décompte de jours pour prévoir les dates de demi-récolte, et justifie la construction de modèles basés sur la température.

Les relations linéaires entre vitesse de développement et températures moyennes (Fig. 20) confirment la relation entre durée de développement et température, et permettent de vérifier l'hypothèse de linéarité préliminaire à l'application du système des sommes de températures. Les températures de base calculées paraissent faibles pour une plante subtropicale, en particulier celle négative de la phase floraison-récolte. Il n'a cependant pas été trouvé d'observation aberrante dans les calculs, et les résultats obtenus par deux approches différentes (minimisation et formules analytiques) sont similaires. Ils sont de plus cohérents avec ceux de la bibliographie au regard de l'effet de la méthode d'estimation sur la valeur estimée de la température de base. Des températures de base négatives ont déjà été rapportées, mais pour des plantes de climat froid (Wielgolaski, 1999). Dans le système des sommes de températures, la température de base a une valeur statistique et non physiologique (Arnold, 1959; Yang *et al.*, 1995; Bonhomme, 2000). Elle est estimée sous l'hypothèse de linéarité entre vitesse de développement et température, et dépend de la méthode d'estimation, de la mesure de la température, et de la présence de facteurs autres que la température affectant la croissance (Yang *et al.*, 1995; Bonhomme, 2000). Les exemples cités par Yang *et al.* (1995) et Bonhomme (2000) montrent que les températures de base ainsi estimées ont souvent des valeurs plus basses que les zéros de développement physiologique.

Si l'on considère qu'une précision de ± 10 jours sur la date de demi-floraison ou de demi-récolte est acceptable compte-tenu de la durée de ces phases et de l'étalement de la floraison (voir chapitre III-I) et de la récolte (voir chapitre V-III) sur une parcelle, alors les modèles sont relativement précis, malgré certains écarts importants entre valeurs observées et estimées. Les distributions des écarts entre dates observées et estimées par les modèles sont semblables en validation interne et externe, ce qui montre la pertinence des modèles pour la prédiction dans la gamme de températures moyennes étudiées.

Pour les deux phases de développement, les écarts entre dates observées et estimées sont corrélés à la durée de la phase de développement, mais pas à la température moyenne, qui est un déterminant important de la durée de développement. Cela suggère que les corrélations entre écarts et durées de développement reflètent l'effet de facteurs autres que la température ayant une influence sur la durée de développement. Ces facteurs pourraient être des facteurs de croissance. En particulier, pour la phase floraison-récolte, des corrélations significatives

sont de plus notées entre les écarts du modèle et l'altitude et le taux de fructification. Dans le chapitre V-III, il est montré que le rapport feuilles-fruits du plant et le nombre moyen de graines par fruit affectent la durée de développement du fruit. En altitude, les goyaviers-fraises perdent une grande partie de leurs feuilles en saison fraîche, ce qui diminue le rapport feuilles:fruits des plants et augmente très probablement la durée de développement des fruits. On ne peut cependant exclure, en altitude, un effet limitant des faibles températures sur le développement du fruit (températures moyennes journalières les plus faibles comprises entre 11 et 12°C). De même, la corrélation positive entre les écarts du modèle et les taux de fructification est à rapprocher de l'effet du rapport feuilles:fruits, et probablement du nombre moyen de graines par fruit, sur la durée de développement du fruit. L'absence de corrélation entre la durée de la phase floraison-récolte et le taux de fructification (Tableau 15) est probablement due à la forte variabilité des températures moyennes et à leur effet prépondérant sur la durée de la phase. Le taux de fructification n'est cependant qu'une composante de la charge en fruits des plants, et des variables plus pertinentes, mais plus laborieuses à quantifier, auraient été le rapport feuilles:fruits moyen des plants ou la production totale de la parcelle pour chaque cycle. Les écarts positifs les plus élevés sur la phase floraison-récolte sont d'ailleurs relevés en saison fraîche et en altitude et/ou sur des cycles à forte production.

La dispersion plus forte des points dans la régression de la vitesse de développement sur la température moyenne (Fig. 20), et les erreurs-standards plus élevées sur l'estimation de la température de base durant la phase floraison-récolte (Tableau 13) vont également dans le sens de l'effet d'autres facteurs que la seule température sur cette phase, alors que ce serait moins le cas pour la phase fertilisation-floraison.

Les remarques précédentes suggèrent que les écarts du modèle pour la phase floraison-récolte, et dans une moindre mesure fertilisation-floraison, sont probablement liés en partie à l'effet de facteurs de croissance, et donc que ces derniers ont une influence sur la durée de développement. Or une base de la méthode des sommes de températures est qu'il n'y ait pas de facteur limitant autre que la température sur la vitesse de développement (Wang, 1960; Bonhomme, 2000). L'effet des facteurs de croissance pourrait ainsi expliquer la valeur très basse de la température de base estimée pour la phase floraison-récolte.

Compte tenu du jeu de données utilisé pour estimer les températures de base, il est attendu un comportement satisfaisant des modèles en prévision dans la gamme de températures moyennes et de conditions de culture testées. On s'attend, à l'échelle de la région de validité définie par les conditions expérimentales, à une moyenne des écarts entre valeurs observées et estimées proche de zéro. Les conditions de culture testées sont les conditions environnementales et de conduite des parcelles, et le caractère déclenché ou non du cycle de production. Ce dernier ne semble pas avoir un effet direct sur la précision du modèle. Les résultats de la validation externe suggèrent que ce sont les conditions de croissance induites par le déclenchement du cycle (charge en fruits plus faible en général sur les cycles déclenchés que sur les cycles naturels) qui déterminent la précision du modèle. Sur un site donné, la précision attendue des modèles est moyenne, et un modèle paramétré à partir de données relatives à ce site serait plus précis.

Les modèles thermiques pourraient être sophistiqués par l'introduction de températures seuils minimale et/ou maximale dans le calcul des sommes de températures (voir par exemple Bustan *et al.*, 1996; Jenni *et al.*, 1996; Dufault, 1997). Cependant, l'effet probable des facteurs de croissance sur le développement du fruit et les données thermiques disponibles risquent de limiter l'intérêt de tels modèles plus complexes.

Concernant l'utilisation des modèles pour prévoir en deux étapes une date de demi-récolte dès la date de fertilisation, il a été montré que la précision sur la date de demi-récolte estimée est très moyenne, à cause principalement des écarts entre les dates de demi-floraison observées et estimées à partir de la fertilisation. La recherche de relations directes entre la durée complète du cycle fertilisation-demi récolte et la température n'a pas été envisagée car la segmentation du cycle de production en phases, et l'établissement d'un modèle thermique pour chacune, donne en général des prédictions plus précises qu'une seule relation sur tout le cycle (Wang, 1960). En pratique, une première estimation de la date de demi-récolte est réalisée dès la fertilisation, en utilisant des données thermiques historiques. Puis la date de demi-récolte est ré-estimée à partir de la date de demi-floraison observée, soit 3 à 6 mois avant la demi-récolte, et non pas à partir de celle estimée depuis la fertilisation. Si les facteurs de croissance sont connus (rapport feuilles:fruits des plants et nombre moyen de graines dans les fruits), le modèle de prédiction de la distribution des dates de récolte (Chapitre V-III) donnera des sorties plus précises et plus complètes tenant compte de l'effet des ces facteurs de croissance.

Littérature citée:

- Arnold, C. Y. (1959). The determination and significance of the base temperature in a linear heat unit system. *Proceedings of the American Society for Horticultural Science*, 74, 430-45.
- Bonhomme, R. (2000). Bases and limits to using 'degree.day' units. *European Journal of Agronomy*, 13, 1-10.
- Bustan, A., Goldschmidt, E.E. et Erner, Y. (1996). Integrating temperature effects on fruit growth into a Citrus productivity model. *Proceedings of the International Society of Citriculture*, vol. 2, VIII International Citrus Congress, 12-17 May 1996, Sun City, South Africa, 938-944.
- Dufault, R.J. (1997). Determining heat units requirements for broccoli harvest in coastal South Carolina. *Journal of the American Society for Horticultural Science*, 122, 169-174.
- Jenni, S., Cloutier, D.C., Bourgeois, G. et Stewart, K.A. (1996). A heat unit model to predict growth and development of muskmelon to anthesis of perfect flowers. *Journal of the American Society for Horticultural Science*, 121, 274-280.
- Mathsoft, (1999). *S-Plus® 2000 Guide to Statistics, volume 1*. Data Analysis Products Division, Mathsoft, Seattle, WA, USA.
- Normand, F. et Huet, P. (2000). *Fiche technique goyavier (2^{ème} version)*. Cirad-Flhor, Saint-Pierre de la Réunion, France.
- Paniandy, J-C., Normand, F. et Reynes, M. (1999). Facteurs intervenant sur la conservation en frais de la goyave-fraise à l'île de la Réunion. *Fruits*, 54, 49-56.
- Sprent, P. (1989). *Applied nonparametric statistical methods*. Chapman & Hall, London, United Kingdom.

- Wang, J.Y. (1960). A critique of the heat unit approach to plant response studies. *Ecology*, 41, 785-790.
- Wielgolaski, F. E. (1999). Starting dates and basic temperatures in phenological observations of plants. *International Journal of Biometeorology*, 42, 158-68.
- Yang, S., Logan, J. et Coffey, D.L. (1995). Mathematical formulae for calculating the base temperature for growing degree days. *Agricultural and Forest Meteorology*, 74, 61-74.

Annexe 2: Détermination du nombre d'ovules dans les fleurs

L'objectif de cette étude est de déterminer la distribution du nombre d'ovules dans les fleurs, afin de savoir si ce nombre peut être un facteur limitant pour le nombre de graines formées. L'influence de deux facteurs d'intérêt dans le cadre de notre étude sur l'élaboration du rendement du goyavier-fraise dans les conditions réunionnaises a été testée: le site et le type de floraison, naturelle ou déclenchée par un apport d'azote.

1- Matériels et méthodes

Des comptages d'ovules dans les fleurs ont été réalisés sur trois parcelles situées à 200, 520 et 890 m d'altitude au cours de deux floraisons sur chaque parcelle: l'une naturelle (oct. 98 à 200 m; déc. 98 à 520 et 890 m), et l'autre déclenchée par un apport d'azote (juin 98 à 200 m; nov. 98 à 520 m; avr.-mai 98 à 890 m). A 520 m, les floraisons naturelle et déclenchée étaient localisées sur des plants différents. Pour chaque floraison, 20 fleurs ont été récoltées au hasard sur la parcelle. Les ovules ont été comptés par dissection de l'ovaire sous une loupe binoculaire.

Les effets du site et du type de floraison ont été recherchés par analyse de variance bifactorielle sur les données non transformées. L'ajustement des distributions observées à des lois normales a été vérifié par le test de Kolmogorov-Smirnov (Sprent, 1989). La distribution globale du nombre d'ovules par fleur (120 données) a été comparée graphiquement à une distribution du nombre de graines par fruit construite à partir des comptages faits sur 3723 fruits issus de différents cycles de production sur les 3 sites de cette étude et un autre site localisé à 480 m d'altitude. Les fruits de l'essai de pollinisation contrôlée (Chapitre IV-II), en particulier ceux issus des allopollinisations saturantes, font partie de ces données. On suppose que cette distribution est représentative des nombres de graines par fruit rencontrés en parcelle.

2- Résultats et discussion

Le nombre d'ovules par fleur n'est pas constant. Il a varié de 75 à 122 à 200 m, de 69 à 114 à 520 m, et de 63 à 113 à 890 m. La moyenne générale est de 92.5. Ces valeurs sont similaires à celles trouvées au Brésil par Raseira et Raseira (1996) (moyenne=94; mini=76; maxi=119) pour la population 'Roxo sudeste' dont les caractéristiques végétatives et des fruits correspondent à celles des goyaviers-fraises présents à la Réunion.

Les distributions du nombre d'ovules par fleur peuvent être ajustées de façon satisfaisante par des lois normales, que ce soit à l'échelle du site, ou globalement sur l'ensemble des sites (Fig. 21). Ce type d'ajustement a également été obtenu par de Reffye *et al.* (1978) sur des clones de cacao. Un effet site significatif est constaté sur la moyenne du nombre d'ovules par fleur, avec des valeurs plus élevées sur le site à 200 m d'altitude que sur les deux autres (Tableau 16). Le type de floraison n'a en moyenne pas d'effet sur le nombre d'ovules par fleur, sauf à 890 m d'altitude où les fleurs de la floraison déclenchée ont significativement plus d'ovules que celles de la floraison naturelle ($P=0.02$).

Les distributions du nombre de graines par fruit et du nombre d'ovules par fleur montrent que les ovules sont en nombre nettement supérieur aux graines formées (Fig. 22), y compris en pollinisation saturante. Le nombre maximum de graines par fruit a été de 46, et le nombre minimum d'ovules par fleur de 63, soit un nombre supérieur de plus d'un tiers au nombre

maximal de graines. L'existence d'ovules surnuméraires est un caractère présent chez de nombreuses espèces de Myrtoideae (Nic Lughadha et Proença, 1996). Bien qu'il n'ait pas été observé de différence morphologique entre les ovules à la loupe binoculaire, certains d'entre eux pourraient être anormaux, stériles ou avoir tendance à avorter après fécondation (Sedgley et al., 1989; Gurusamy, 1999). Si cette proportion devient importante, alors le nombre d'ovules viables peut être limitant. La vérification de cette hypothèse demanderait des études spécifiques. En l'absence de tels résultats, et compte tenu de la différence entre les distributions des nombres d'ovules et de graines, et de l'effet marqué de facteurs tels que l'origine du pollen ou la pluie sur la distribution du nombre de graines, nous considérons que le nombre d'ovules n'est pas un facteur limitant la formation de graines dans l'ovaire.

Site	Flo. déclenchée	Flo. naturelle	Moyenne/site
200 m	100.0	102.7	101.3 ^a
520 m	87.4	88.6	88.0 ^b
890 m	92.7	83.7	88.2 ^b
Moyenne/floraison	93.4	91.6	92.5

Tableau 16: Effet du site et du type de floraison (déclenchée par un apport d'azote ou naturelle) sur le nombre d'ovules par fleur de goyavier-fraise. Les moyennes par site suivies de lettres différentes sont significativement différentes (test de Tukey de comparaisons multiples de moyennes). Effet site: $P < 0.001$; effet type de floraison: $P = 0.39$; interaction: $P = 0.04$.

Cette étude met en évidence les effets du site, et dans une moindre mesure du type de floraison sur le nombre d'ovules par fleur, mais ne permet pas de les expliquer. Des facteurs génétiques sont à l'origine de différences dans le nombre d'ovules par fleur (Dattee, 1972; de Reffye et al., 1978; Snow, 1982). Bien que le facteur génétique soit à considérer dans les différences observées, nos résultats suggèrent que d'autres facteurs d'ordre pédoclimatique ou physiologique pourraient intervenir. Les différences relevées entre sites et types de floraison ne modifient cependant pas l'hypothèse précédente.

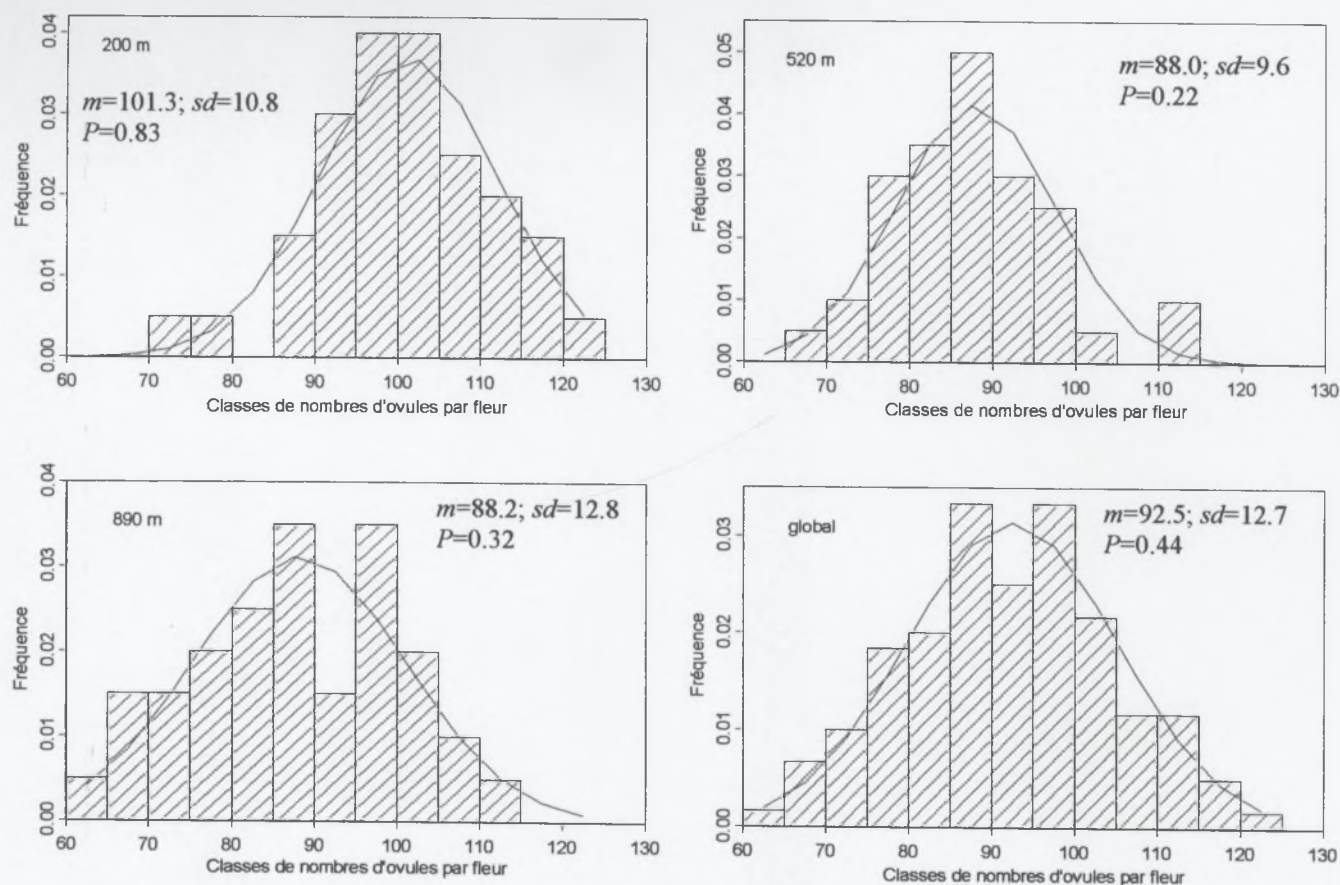


Figure 21: Distributions observées et estimées par une loi normale $N(m, sd^2)$ du nombre d'ovules dans les fleurs de goyavier-fraise sur 3 sites localisés à 200, 520 et 890m, et globalement sur les 3 sites. P : probabilité du test d'ajustement de Kolmogorov-Smirnov.

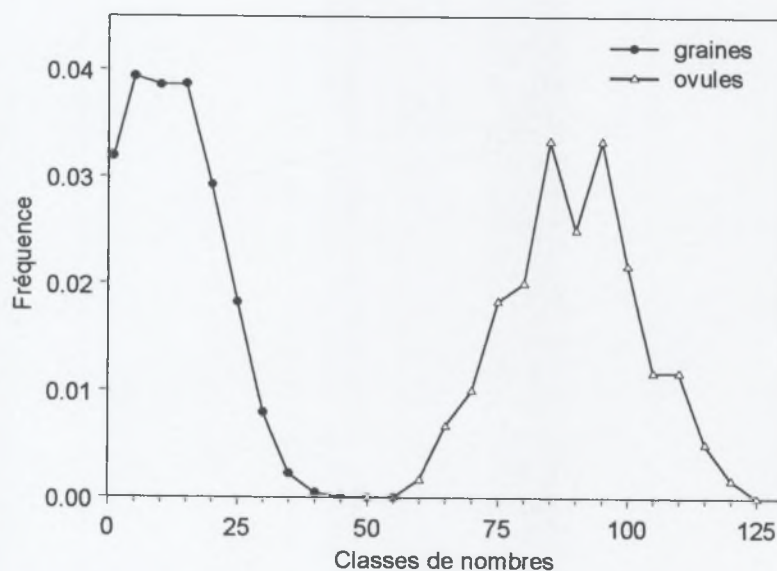


Figure 22: Distributions du nombre de graines dans les fruits ($n=3723$) et du nombre d'ovules dans les fleurs ($n=120$) chez le goyavier-fraise à la Réunion.

Littérature citée:

- Dattee, Y. (1972). Hérité quantitive du nombre d'ovules par ovaire dans quelques familles de luzerne cultivée (*Medicago sativa* L.). *Annales de l'Amélioration des Plantes*, 22, 201–209.
- Gurusamy, C. (1999). The pattern of seed formation in cauliflower (*Brassica oleracea* L. var. *botrytis*) with special reference to seed abortion. *Canadian Journal of Botany*, 77, 1189–1192.
- Nic Lughadha, E. et Proença, C. (1996). A survey of the reproductive biology of the Myrtoideae (Myrtaceae). *Annals of the Missouri Botanical Garden*, 83, 480–503.
- Raseira, M.C.B. et Raseira, A. (1996). *Contribuição ao estudo do araçazeiro*, *Psidium cattleianum*. EMBRAPA-CPACT, Pelotas, Brazil.
- de Reffye, P., Parvais, J.P., Mossu, G. et Lucas, P. (1978). Influence des aléas de la pollinisation sur les rendements du cacaoyer. Modèle mathématique et simulation. *Café Cacao Thé*, 22, 251–274.
- Sedgley, M., Hand, F.C., Smith R.M. et Griffin, A.R. (1989). Pollen tube growth and early seed development in *Eucalyptus regnans* F. Muell. (Myrtaceae) in relation to ovule structure and preferential outcrossing.. *Australian Journal of Botany*, 37, 397–411.
- Snow, A.A. (1982). Pollination intensity and potential seed set in *Passiflora vitifolia*. *Oecologia*, 55, 231–237.
- Sprent, P. (1989). *Applied nonparametric statistical methods*. Chapman & Hall, London, United Kingdom.

Annexe 3: Calcul de la vraisemblance d'un déroulement de floraison observé

Nous avons utilisé une distribution multinomiale pour calculer la probabilité d'un déroulement de floraison observé, i.e. sa vraisemblance (modèle de floraison, Chapitre III-II). Cependant, une approche par les probabilités conditionnelles est aussi possible, et nous allons montrer que les deux approches amènent au même résultat.

Soit un échantillon de N boutons floraux d'une unité florifère dont on observe le déroulement complet de la floraison, c'est à dire les nombres de boutons floraux n_1, \dots, n_D ($\sum_{i=1}^{i=D} n_i = N$) qui s'ouvrent au cours des intervalles de temps consécutifs A_1, \dots, A_D .

Soient p_1, \dots, p_D les probabilités théoriques d'ouverture d'un bouton floral sur cette unité florifère au cours des intervalles de temps A_1, \dots, A_D ($\sum_{i=1}^{i=D} p_i = 1$). D'après le modèle de floraison, les N boutons floraux de l'unité florifère sont indépendants par rapport à la floraison, c'est à dire que les événements correspondant à l'ouverture de chaque bouton floral sont indépendants, soit chaque bouton floral a une même probabilité de s'ouvrir au cours de l'intervalle A_i .

Soient X_1, \dots, X_D les D variables aléatoires qui représentent le nombre de boutons floraux qui s'ouvrent au cours des intervalles de temps A_1, \dots, A_D respectivement.

La probabilité d'observer n_1, \dots, n_D boutons floraux qui s'ouvrent au cours des intervalles de temps A_1, \dots, A_D , et que l'on cherche à calculer, est: $P = \pi(X_1=n_1 \text{ et } X_2=n_2 \text{ et } \dots \text{ et } X_D=n_D)$.

1- Utilisation d'une distribution multinomiale.

L'indépendance des N boutons floraux par rapport à la floraison implique que chaque bouton floral a une probabilité théorique p_i de s'ouvrir au cours de l'intervalle de temps A_i . La détermination des intervalles de floraison des N boutons floraux est une série de N tirages indépendants et identiques dans la loi théorique p_1, \dots, p_D . La variable π suit donc une loi multinomiale de paramètres N et p_1, \dots, p_D (Dagnelie, 1998), d'où:

$$P = \pi(X_1=n_1 \text{ et } X_2=n_2 \text{ et } \dots \text{ et } X_D=n_D) = \frac{N!}{\prod_{i=1}^{i=D} n_i!} \prod_{i=1}^{i=D} p_i^{n_i}$$

2- Utilisation des probabilités conditionnelles

Cette approche consiste à dire que la probabilité d'ouverture d'un bouton floral au cours d'un intervalle de temps A_k dépend du nombre de boutons floraux qui se sont ouverts au cours des intervalles de temps précédents, A_1, \dots, A_{k-1} , et donc du nombre de boutons floraux restant. Alors P s'écrit:

$$(1) \quad P = \pi(X_1=n_1 \text{ et } X_2=n_2 \text{ et } \dots \text{ et } X_D=n_D) = p(X_1=n_1) p(X_2=n_2|X_1=n_1) \dots p(X_D=n_D|X_1=n_1, X_2=n_2, \dots, X_{D-1}=n_{D-1})$$

Calcul individuel des D termes de ce produit de probabilités:

• Premier terme: X_1 est la variable aléatoire représentant le nombre de boutons floraux, parmi N , qui s'ouvrent durant l'intervalle A_1 . La probabilité théorique d'ouverture d'un bouton floral durant A_1 est p_1 . Les N boutons floraux sont indépendants par rapport à la floraison. La variable aléatoire X_1 suit une loi binomiale de paramètres N et p_1 , d'où:

$$p(X_1=n_1) = C_N^{n_1} p_1^{n_1} (1-p_1)^{N-n_1}$$

• Deuxième terme: $(X_2|X_1=n_1)$ est la variable aléatoire représentant le nombre de boutons floraux qui s'ouvrent au cours de l'intervalle A_2 parmi les $N-n_1$ boutons floraux restants. Ces $N-n_1$ boutons floraux sont indépendants par rapport à la floraison. La probabilité théorique d'ouverture d'un bouton floral au cours de A_2 est alors $p'_2 = \frac{n_2}{N-n_1} = \frac{p_2}{1-p_1}$. La variable aléatoire $(X_2|X_1=n_1)$ suit une loi binomiale de paramètres $N-n_1$ et p'_2 , d'où:

$$p(X_2=n_2|X_1=n_1) = C_{N-n_1}^{n_2} p'^{n_2}_2 (1-p'_2)^{N-n_1-n_2} = C_{N-n_1}^{n_2} \frac{p_2^{n_2} (1-p_1-p_2)^{N-n_1-n_2}}{(1-p_1)^{N-n_1}}$$

• k^{ième} terme: $(X_k | X_1=n_1, X_2=n_2, \dots, X_{k-1}=n_{k-1})$ est la variable aléatoire représentant le nombre de boutons floraux qui s'ouvrent au cours de l'intervalle A_k parmi les $N - \sum_{i=1}^{i=k-1} n_i$ boutons floraux restants. La probabilité théorique d'ouverture d'un bouton floral au cours de l'intervalle A_k est: $p'_k = \frac{n_k}{N - \sum_{i=1}^{i=k-1} n_i} = \frac{p_k}{1 - \sum_{i=1}^{i=k-1} p_i}$. La variable aléatoire $(X_k | X_1=n_1, X_2=n_2, \dots, X_{k-1}=n_{k-1})$ suit une loi

binomiale de paramètres $N - \sum_{i=1}^{i=k-1} n_i$ et p'_k , d'où:

$$\begin{aligned} p(X_k=n_k | X_1=n_1, \dots, X_{k-1}=n_{k-1}) &= C_{N-\sum_{i=1}^{i=k-1} n_i}^{n_k} p'^{n_k}_k (1-p'_k)^{N-\sum_{i=1}^{i=k} n_i} \\ &= C_{N-\sum_{i=1}^{i=k-1} n_i}^{n_k} \frac{p_k^{n_k}}{(1-\sum_{i=1}^{i=k-1} p_i)^{n_k}} \left(\frac{1-\sum_{i=1}^{i=k} p_i}{1-\sum_{i=1}^{i=k-1} p_i} \right)^{N-\sum_{i=1}^{i=k} n_i} = C_{N-\sum_{i=1}^{i=k-1} n_i}^{n_k} \frac{p_k^{n_k} (1-\sum_{i=1}^{i=k} p_i)^{N-\sum_{i=1}^{i=k} n_i}}{(1-\sum_{i=1}^{i=k-1} p_i)^{N-\sum_{i=1}^{i=k-1} n_i}} \end{aligned}$$

• Avant-dernier terme:

$$p(X_{D-1}=n_{D-1} | X_1=n_1, X_2=n_2, \dots, X_{D-2}=n_{D-2}) = C_{N-\sum_{i=1}^{i=D-2} n_i}^{n_{D-1}} \frac{p_{D-1}^{n_{D-1}} (1-\sum_{i=1}^{i=D-1} p_i)^{N-\sum_{i=1}^{i=D-1} n_i}}{(1-\sum_{i=1}^{i=D-2} p_i)^{N-\sum_{i=1}^{i=D-2} n_i}}$$

$$p(X_{D-1}=n_{D-1} | X_1=n_1, X_2=n_2, \dots, X_{D-2}=n_{D-2}) = C_{N-\sum_{i=1}^{i=D-2} n_i}^{n_{D-1}} \frac{p_{D-1}^{n_{D-1}} p_D^{n_D}}{\left(1 - \sum_{i=1}^{i=D-2} p_i\right)^{N-\sum_{i=1}^{i=D-2} n_i}}$$

• Dernier terme: $p(X_D=n_D | X_1=n_1, X_2=n_2, \dots, X_{D-1}=n_{D-1})$ est la probabilité que n_D boutons floraux s'ouvrent durant l'intervalle de temps A_D parmi les n_D boutons floraux restants et qui doivent s'ouvrir au cours de A_D . Cette probabilité vaut 1.

L'expression (1) s'écrit donc:

$$\begin{aligned} P &= \pi(X_1=n_1 \text{ et } X_2=n_2 \text{ et } \dots \text{ et } X_D=n_D) \\ &= \Gamma \cdot p_1^{n_1} (1-p_1)^{N-n_1} \cdot \frac{p_2^{n_2} (1-p_1-p_2)^{N-n_1-n_2}}{(1-p_1)^{N-n_1}} \cdot \dots \cdot \frac{p_k^{n_k} \left(1 - \sum_{i=1}^{i=k} p_i\right)^{N-\sum_{i=1}^{i=k} n_i}}{\left(1 - \sum_{i=1}^{i=k-1} p_i\right)^{N-\sum_{i=1}^{i=k-1} n_i}} \cdot \dots \cdot \frac{p_{D-1}^{n_{D-1}} p_D^{n_D}}{\left(1 - \sum_{i=1}^{i=D-2} p_i\right)^{N-\sum_{i=1}^{i=D-2} n_i}} \\ &= \Gamma \cdot \prod_{i=1}^{i=D} p_i^{n_i} \end{aligned}$$

$$\text{avec: } \Gamma = C_N^{n_1} \cdot C_{N-n_1}^{n_2} \cdot \dots \cdot C_{N-\sum_{i=1}^{i=k-1} n_i}^{n_k} \cdot \dots \cdot C_{N-\sum_{i=1}^{i=D-2} n_i}^{n_{D-1}}$$

$$\Gamma = \frac{N!}{n_1! (N-n_1)!} \cdot \frac{(N-n_1)!}{n_2! (N-n_1-n_2)!} \cdot \dots \cdot \frac{\left(N - \sum_{i=1}^{i=k-1} n_i\right)!}{n_k! \left(N - \sum_{i=1}^{i=k} n_i\right)!} \cdot \dots \cdot \frac{\left(N - \sum_{i=1}^{i=D-2} n_i\right)!}{n_{D-1}! n_D!} = \frac{N!}{\prod_{i=1}^{i=D} n_i!}$$

Soit:

$$P = p(X_1=n_1) p(X_2=n_2 | X_1=n_1) \dots p(X_D=n_D | X_1=n_1, X_2=n_2, \dots, X_{D-1}=n_{D-1}) = \frac{N!}{\prod_{i=1}^{i=D} n_i!} \prod_{i=1}^{i=D} p_i^{n_i}$$

3- Conclusion

Les deux approches conduisent donc au même résultat. Ces calculs, réalisés à l'échelle de l'unité florifère, sont valables à l'échelle de l'arbre ou de la population, en utilisant les effectifs observés à ces échelles et les probabilités théoriques correspondantes (Chapitre III-II).

Littérature citée:

Dagnelie, P. (1998). *Statistique théorique et appliquée. tome 1, inférence statistique à une et à deux dimensions*. De Boeck & Larcier S.A., Brussels, Belgium.

Annexe 4: Tableaux de présentation synthétique des modèles

Modèle	Objectif	Caractéristiques	Equations	Paramètres
Modèle de floraison	Prédiction	<p>Calcule la probabilité de floraison $P(t_i)$ d'un bouton floral entre les dates t_{i-1} et t_i exprimées en temps de développement.</p> <p>Trois modèles de floraison sont dérivés d'un même modèle général:</p> <ul style="list-style-type: none"> -modèle 'rameau': considère l'asynchronisme de floraison intra et inter plants. - modèle 'plant': ne considère que l'asynchronisme de floraison inter plants. - modèle 'population': ne considère pas d'asynchronisme de floraison intra plant, ni inter plants. <p>Variables d'entrée: dates thermiques: t_{i-1} et t_i (°C.j)</p> <p>Variable de sortie: probabilité de floraison d'un bouton floral entre t_{i-1} et t_i: $P(t_i)$</p>	<p>Notation: $\Psi(z) = \int_z^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \cdot dx$</p> <p>Modèle 'rameau':</p> $P(t_i) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} \left[\Psi \left[\frac{(a+u)-t_i}{\sigma_s \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{(a+u)-t_{i-1}}{\sigma_s \cdot \sqrt{t_{i-1}}} \right] \right] \cdot g(u) \cdot du \right] \cdot f(a) \cdot da$ <p>où: $g(u) = \frac{1}{\delta \cdot \sqrt{2\pi}} \cdot e^{-\frac{u^2}{2\delta^2}}$</p> $f(a) = \frac{1}{\beta_s \cdot \sqrt{2\pi}} \cdot e^{-\frac{(a-\alpha_s)^2}{2\beta_s^2}}$ <p>Modèle 'plant':</p> $P(t_i) = \int_{-\infty}^{+\infty} \left[\Psi \left[\frac{a-t_i}{\sigma_i \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{a-t_{i-1}}{\sigma_i \cdot \sqrt{t_{i-1}}} \right] \right] \cdot f(a) \cdot da$ <p>où: $f(a) = \frac{1}{\beta_i \cdot \sqrt{2\pi}} \cdot e^{-\frac{(a-\alpha_i)^2}{2\beta_i^2}}$</p> <p>Modèle 'population':</p> $P(t_i) = \Psi \left[\frac{a-t_i}{\sigma \cdot \sqrt{t_i}} \right] - \Psi \left[\frac{a-t_{i-1}}{\sigma \cdot \sqrt{t_{i-1}}} \right]$	<p>Modèle 'rameau': α_s: moyenne de la loi normale des seuils de floraison des plants sur une parcelle. β_s: écart-type de la loi normale des seuils de floraison des plants sur une parcelle. δ: écart-type de la loi normale centrée des décalages entre les seuils de floraison des unités florifères d'un plant et celui du plant. σ_s: paramètre de variance associé à la floraison à l'échelle de l'unité florifère.</p> <p>Modèle 'plant': α_i: moyenne de la loi normale des seuils de floraison des plants sur une parcelle. β_i: écart-type de la loi normale des seuils de floraison des plants sur une parcelle. σ_i: paramètre de variance associé à la floraison à l'échelle du plant.</p> <p>Modèle 'population': a: seuil de floraison d'un bouton floral de la population σ: paramètre de variance associé à la floraison à l'échelle de la parcelle.</p>

Modèle	Objectif	Caractéristiques	Equations	Paramètres
Modèle de fécondation-fructification	Aide à l'analyse de résultats	<p>Exprime la probabilité $P(n)$ qu'une fleur donne un fruit à n ($n > 0$) graines en fonction de l'origine du pollen, et de l'autocompatibilité du plant.</p> <p>Permet, à partir des distributions des nombres de graines par fruit provenant de fleurs auto- et allopollinisées sur un même plant, d'estimer l'autocompatibilité de ce plant.</p> <p>Les conditions expérimentales sont des auto- et allopollinisations saturantes sur un même plant, sans émasculatation préalable des fleurs.</p> <p>Variable d'entrée: nombre de graines par fruit: n</p> <p>Variable de sortie: Probabilité $P(n)$ qu'une fleur donne un fruit à n graines</p>	$P(n) = \left(p \cdot \frac{e^{-\theta_1} \cdot \theta_1^n}{n!} + (1-p) \cdot \frac{e^{-\theta_2} \cdot \theta_2^n}{n!} \right) \cdot q$ <p>($n > 0$) et</p> $P(0) = 1 - q \cdot \left(1 - p \cdot e^{-\theta_1} - (1-p) \cdot e^{-\theta_2} \right)$ <p>avec, en autopolinisation:</p> $\theta_1 = u_1 \cdot \left[1 + \lambda_M / \lambda_m \right] \cdot f_{11}$ $\theta_2 = u_2 \cdot \left[1 + \lambda_M / \lambda_m \right] \cdot f_{21}$ <p>et en allopollinisation:</p> $\theta_1 = u_1 \cdot \left[f_{11} + \lambda_M / \lambda_m \right]$ $\theta_2 = u_2 \cdot \left[f_{21} + \lambda_M / \lambda_m \right]$	<p>p : coefficient de partage entre les deux lois de Poisson de paramètres θ_1 et θ_2 décrivant la distribution bimodale des nombres de graines par fruit à l'échelle du plant.</p> <p>q : taux de fructification, constant</p> <p>u_1 : paramètre de réduction associé à la première loi de Poisson.</p> <p>u_2 : paramètre de réduction associé à la deuxième loi de Poisson.</p> <p>f_{11}: autocompatibilité des fleurs définies par la loi de Poisson de paramètre θ_1</p> <p>f_{21}: autocompatibilité des fleurs définies par la loi de Poisson de paramètre θ_2</p> <p>λ_M / λ_m: rapport, constant, entre le nombre de grains de pollen provenant de la fleur donneuse et le nombre de grains de pollen issus de la fleur receveuse déposés sur le stigmate lors de la pollinisation.</p>

Modèle	Objectif	Caractéristiques	Equations	Paramètres
Modèle de la courbe de croissance du fruit	Représentation	<p>Equation de la courbe de croissance du fruit en poids frais, W, en fonction du temps thermique t.</p> <p>Variable d'entrée: temps thermique: t (°C.j)</p> <p>Variable de sortie: poids frais du fruit au temps t: W (g)</p>	$W(t) = \frac{P}{1 + e^{-\alpha(t-t_I)}} + e^{\beta(t-t_E)}$	<p>P: poids du fruit à la fin de la première phase de croissance (g)</p> <p>t_I: date du point d'inflexion de la première phase de croissance (°C.j)</p> <p>α: ordonnée à l'origine du taux de croissance du fruit ($\text{g.g}^{-1}.\text{°C}^{-1}.\text{j}^{-1}$)</p> <p>$\beta$: taux de croissance au cours de la seconde phase de croissance ($\text{g.g}^{-1}.\text{°C}^{-1}.\text{j}^{-1}$)</p> <p>$t_E$: date à laquelle la vitesse de croissance de la seconde phase de croissance est égale à β (°C.j)</p>
Modèle de prédiction des dates de récolte à l'échelle du plant	Prédiction	<p>Simule une distribution normale $N(d_{0.5rec}, \sigma_{rec}^2)$ des dates individuelles de récolte des fruits à l'échelle d'un plant. Les paramètres $d_{0.5rec}$ et σ_{rec}^2 sont des relations linéaires du rapport feuilles:fruits du plant et du nombre moyen de graines dans les fruits du plant. Les paramètres simulés $\hat{d}_{0.5rec}$ et $\hat{\sigma}_{rec}^2$ sont tirés aléatoirement dans les variables ci-contre</p> <p>Variables d'entrée: nombre moyen de graines dans les fruits du plant: SN_{moy} rapport feuilles:fruits du plant: LFR</p> <p>Variable de sortie: Distribution estimée des dates de récolte individuelles des fruits d'un plant</p>	<p>$V = (LFR, SN_{moy})$</p> <p>$\hat{d}_{0.5rec}(V) = X_d(V) \hat{\Omega}_d^* + X_d(V) \Delta_d + \varepsilon_d$</p> <p>avec: $X_d(V) = (1, 1/LFR, SN_{moy})$</p> <p>$\hat{\sigma}_{rec}(V) = X_\sigma(V) \hat{\Omega}_\sigma^* + X_\sigma(V) \Delta_\sigma + \varepsilon_\sigma$</p> <p>avec: $X_\sigma(V) = (1/LFR, SN_{moy})$</p>	<p>$\hat{\Omega}_d^*$: paramètres de la régression linéaire entre $d_{0.5rec}$ et $X_d(V)$ estimés à partir des données.</p> <p>Δ_d: erreur sur l'estimation des paramètres de la régression, suit une loi multinormale centrée de matrice de variance-covariance celle associée à l'estimation des paramètres $\hat{\Omega}_d^*$.</p> <p>ε_d: erreur résiduelle (°C.j) de la régression linéaire entre $d_{0.5rec}$ et $X_d(V)$, suit une loi normale centrée de variance la variance résiduelle de la régression.</p> <p>$\hat{\Omega}_\sigma^*$: paramètres de la régression linéaire entre σ_{rec} et $X_\sigma(V)$ estimés à partir des données.</p> <p>Δ_σ: erreur sur l'estimation des paramètres de la régression, suit une loi multinormale centrée de matrice de variance-covariance celle associée à l'estimation des paramètres $\hat{\Omega}_\sigma^*$.</p> <p>ε_σ: erreur résiduelle (°C.j) de la régression linéaire entre σ_{rec} et $X_\sigma(V)$, suit une loi normale centrée de variance la variance résiduelle de la régression.</p>

Modèle	Objectif	Caractéristiques	Equations	Paramètres
Modèle emboîté de prédiction du poids du fruit à maturité	Prédiction	<p>Prédit un poids individuel du fruit à maturité en fonction des facteurs de croissance $U=(SN, LFR)$ [SN: nombre de graines du fruit; LFR: rapport feuilles:fruits du plant] à partir des relations entre les paramètres de la courbe de croissance (ici K et P) et les facteurs de croissance U.</p> <p>Les poids estimés servent à calculer la distribution des poids de fruit à maturité.</p> <p>Une prédiction $\hat{W}_f(U)$ est un tirage aléatoire dans la variable ci-contre. La variabilité de $\hat{W}_f(U)$ et des intervalles de confiance sont estimés par un grand nombre de tirages</p> <p>Variables d'entrée: nombre de graines du fruit: SN rapport feuilles:fruits du plant: LFR</p> <p>Variable de sortie: poids estimé du fruit à maturité: $\hat{W}_f(U)$ (g)</p>	$W_f = K \cdot P + \varepsilon_f$ <p>Le poids individuel d'un fruit caractérisé par les facteurs de croissance $U=(SN, LFR)$ est tiré dans la variable aléatoire</p> $\hat{W}_f(U) = (X_K(U) \hat{\Omega}_K^* + X_K(U) \Delta_K + \varepsilon_K) \cdot (X_P(U) \hat{\Omega}_P^* + X_P(U) \Delta_P + \varepsilon_P) + \varepsilon_r$ <p>avec: $X_K(U) = (1, \log(LFR))$ $X_P(U) = (LFR, SN)$.</p>	<p>K: paramètre d'échelle entre le poids du fruit à maturité et son poids P après la première phase de croissance. P: poids du fruit à la fin de la première phase de croissance (g) ε_f: terme aléatoire</p> <p>$\hat{\Omega}_K^*$: paramètres de la régression linéaire entre K et $X_K(U)$. estimés à partir des données. Δ_K: erreur sur l'estimation des paramètres de la régression, suit une loi multinormale centrée de matrice de variance-covariance celle associée à l'estimation des paramètres $\hat{\Omega}_K^*$. ε_K: erreur résiduelle de la régression linéaire entre K et $X_K(U)$, suit une loi normale centrée de variance la variance résiduelle de la régression $\hat{\Omega}_P^*$: paramètres de la régression linéaire entre P et $X_P(U)$ estimés à partir des données. Δ_P: erreur sur l'estimation des paramètres de la régression, suit une loi multinormale centrée de matrice de variance-covariance celle associée à l'estimation des paramètres $\hat{\Omega}_P^*$. ε_P: erreur résiduelle de la régression linéaire entre P et $X_P(U)$, suit une loi normale centrée de variance la variance résiduelle de la régression ε_r: variable aléatoire normale associée à l'erreur résiduelle du modèle de la courbe de croissance individuelle du fruit.</p>

Modèle	Objectif	Caractéristiques	Equations	Paramètres
Modèle simplifié de prédiction du poids du fruit à maturité	Prédiction	<p>Prédit directement un poids individuel du fruit à maturité en fonction d'un prédicteur $SMlog(LFR)$ dépendant des facteurs de croissance $U=(SN, LFR)$ [SN: nombre de graines du fruit; LFR: rapport feuilles:fruits du plant].</p> <p>Les poids estimés servent à calculer la distribution des poids de fruit à maturité</p> <p>Variables d'entrée: nombre de graines du fruit: SN rapport feuilles:fruits du plant: LFR</p> <p>Variable de sortie: poids estimé du fruit à maturité: $\hat{W}_f(U)$ (g)</p>	$\hat{W}_f(U) = X(U) \hat{\Omega}^*$ <p>avec: $X(U) = (1, SN \cdot \log(LFR))$</p>	$\hat{\Omega}^*$: paramètres de la régression linéaire entre W_f et $X(U)$ estimés à partir des données.

Annexe 5: Planches photos



P. cattleianum var. *cattleianum*



P. cattleianum var. *lucidum*



Bouton floral et fleurs de goyavier-fraise sur deux pousses issues d'un même rameau terminal



Stades de maturité différenciés des fruits sur un même plant



Plant de goyavier-fraise en pleine période de récolte

Liste des figures (hors figures incluses dans les articles)

Figure 1: Distribution en classes de 1 g du calibre de l'ensemble des fruits récoltés sur 2 plants d'une même parcelle et un même cycle de production.

Figure 2: Le cycle phénologique naturel du goyavier-fraise à la Réunion (500 m d'altitude).

Figure 3: Phénologie du goyavier-fraise en parcelle cultivée (500 m d'altitude).

Figure 4: Variabilité de la production et du poids moyen d'un fruit par plant lors de 2 cycles naturels et 2 cycles déclenchés par fertilisation en 1999 et 2000.

Figure 5: Organisation du modèle global d'élaboration du nombre de fruits, de leurs calibres et de leurs dates de récolte pour un cycle de production naturel chez le goyavier-fraise.

Figure 6: Distributions observées du nombre de graines par fruit après une pollinisation libre sur 4 plants de goyavier-fraise.

Figure 7: Relation entre le nombre moyen de graines par fruit sur un rameau de goyavier-fraise et le taux de fructification sur ce rameau, en autopolinisation saturante, en allopollinisation saturante, et en pollinisation libre.

Figure 8: Distributions observées et ajustées par la somme de 2 lois de Poisson, du nombre de graines par fruit sur 11 plants en pollinisation libre.

Figure 9: Distributions observées, ajustées avec le modèle général, et avec le modèle simplifié en λ du nombre de graines par fruit après une autopolinisation saturante sur 4 plants de goyavier-fraise.

Figure 10: Distributions observées, ajustées avec le modèle général, et avec le modèle simplifié en λ du nombre de graines par fruit après une allopollinisation saturante sur 5 plants de goyavier-fraise.

Figure 11: Distribution des dates individuelles de récolte de 165 fruits d'un plant ($LFR = 1.9$) en fonction de leur nombre de graines.

Figure 12: Evolution cumulée de la totalité de la récolte de 9 plants dont le rapport feuilles:fruits varie de 0.8 à 12.6.

Figure 13: Croissance cumulée en poids frais de deux fruits contenant chacun 17 graines et provenant de plants ayant des rapports feuilles:fruits différents.

Figure 14: Relation entre l'étalement de la récolte et le nombre moyen de graines par fruit en fonction du rapport feuilles:fruits du plant sur 38 goyaviers-fraises représentant 2 parcelles différentes.

Figure 15: Relation entre la date de demi-récolte et le nombre moyen de graines par fruit en fonction du rapport feuilles:fruits du plant sur 38 goyaviers-fraises représentant 2 parcelles différentes.

Figure 16: Relation entre les moyennes de la loi normale représentant la distribution des dates de récolte des fruits d'un plant observées et estimées en fonction du rapport feuilles:fruits du plant et du nombre moyen de graines dans ses fruits.

Figure 17: Relation entre les écart-types de la loi normale représentant la distribution des dates de récolte des fruits d'un plant observées et estimées en fonction du rapport feuilles:fruits du plant et du nombre moyen de graines dans ses fruits.

Figure 18: Validation sur l'essai PR du modèle de prédiction des dates de récolte. Distribution observée des dates de récolte sur 27 plants (19172 fruits), et médiane et région de prédiction à 95% de 1000 distributions simulées.

Figure 19: Validation sur l'essai JBG du modèle de prédiction des dates de récolte. Distribution observée des dates de récolte sur 11 plants (17360 fruits), et médiane et région de prédiction à 95% de 1000 distributions simulées.

Figure 20: Relation entre la vitesse de développement et la température moyenne au cours des phases fertilisation-floraison et floraison-récolte.

Figure 21: Distributions observées et estimées par une loi normale du nombre d'ovules dans les fleurs de goyavier-fraise sur 3 sites localisés à 200, 520 et 890m, et globalement sur les 3 sites.

Figure 22: Distributions du nombre de graines dans les fruits ($n=3723$) et du nombre d'ovules dans les fleurs ($n=120$) chez le goyavier-fraise à la Réunion.

Liste des tableaux (hors tableaux inclus dans les articles)

Tableau 1: Analyses physico-chimiques de fruits de goyavier-fraise selon différents auteurs.

Tableau 2: Effet du type de pollinisation et du plant sur le taux de fructification du goyavier-fraise.

Tableau 3: Effet du type de pollinisation et du plant sur le nombre moyen de graines par fruit du goyavier-fraise.

Tableau 4a: Estimation des paramètres du modèle général complet pour l'allopollinisation

Tableau 4b: Estimation des paramètres du modèle général complet pour l'autopollinisation

Tableau 4c: Estimation des paramètres du modèle général complet pour la pollinisation libre

Tableau 5: Estimation des paramètres pour le modèle simplifié à θ_1 , α , p et q constants en allopollinisation

Tableau 6a: Estimation des paramètres pour le modèle simplifié à θ_2 constant en autopollinisation

Tableau 6b: Estimation des paramètres pour le modèle simplifié à p constant en autopollinisation

Tableau 7: Estimation des paramètres pour le modèle en λ .

Tableau 8: Estimation des paramètres pour le modèle en λ simplifié avec p constant en auto- et en allopollinisation, et q constant en allopollinisation

Tableau 9: Effet du rapport feuilles:fruits du plant, obtenu par différents niveaux de la charge en fruits et du nombre de feuilles, sur les paramètres de la loi normale représentant la distribution des dates de récolte des fruits du plant.

Tableau 10: Localisation des parcelles suivies, méthode d'estimation de la température moyenne journalière et stations Météo France impliquées dans cette estimation.

Tableau 11: Caractéristiques des cycles de production utilisés pour l'estimation des températures de base et la validation interne, et pour la validation externe.

Tableau 12: Température de base, erreur standard et intervalle de confiance à 95% estimés pour la phase fertilisation-floraison par différentes méthodes de calcul.

Tableau 13: Température de base, erreur standard et intervalle de confiance à 95% estimés pour la phase floraison-récolte par différentes méthodes de calcul.

Tableau 14: Corrélations entre les écarts du modèle lors de la phase fertilisation-floraison en validation interne et des variables caractéristiques des cycles étudiés.

Tableau 15: Corrélations entre les écarts du modèle lors de la phase floraison-récolte en validation interne et des variables caractéristiques des cycles étudiés.

Tableau 16: Effet du site et du type de floraison (déclenchée par un apport d'azote ou naturelle) sur le nombre de d'ovules par fleurs du goyavier-fraise.

Résumé

La domestication du goyavier-fraise (*Psidium cattleianum*) est envisagée dans le cadre d'un projet de diversification agricole dans les Hauts humides de l'île de la Réunion. Le nombre de fruits par plant et leur calibre sont deux caractères qui déterminent le rendement commercial. Ils sont très variables et semblent peu liés au génotype. Une approche agronomique de la domestication de cette espèce est proposée. Elle consiste à rechercher comment s'élaborent ces caractères et leur variabilité à travers différentes étapes du cycle de production, sous l'effet de facteurs environnementaux, physiologiques et génétiques. Le but est de pouvoir agir sur certains facteurs par des techniques culturales appropriées. Les particularités générales de cette démarche sont la diversité génétique du matériel végétal et le peu de références sur l'espèce, confirmé par une analyse bibliographique.

Le cycle phénologique du goyavier-fraise a été caractérisé à travers un suivi en milieu naturel à la Réunion. En culture, un apport d'azote modifie ce cycle phénologique naturel en déclenchant l'émission d'une pousse florifère. Le cadre de notre travail est un cycle de production naturel en parcelle cultivée, sans interaction avec des cycles déclenchés. Le schéma d'un modèle global d'élaboration du nombre de fruits et de leur calibre est proposé, structuré autour de 5 étapes du cycle de production: l'élaboration du nombre de fleurs, la floraison, la pollinisation, la fécondation-fructification et la croissance du fruit. Les sorties attendues sont les distributions des calibres et des dates de récolte à l'échelle de la parcelle. Ses objectifs sont la simulation et la prédiction pour aider les agriculteurs dans leurs choix techniques. L'étude et la modélisation des étapes de floraison, de fécondation-fructification et de croissance du fruit sont présentées.

La variabilité génétique des plants entraîne un asynchronisme de floraison qui n'est pas pris en compte par un modèle de floraison conçu pour un verger monovariétal. Un modèle général de floraison, stochastique, est alors construit. Il décrit la floraison à différentes échelles et prend en compte l'asynchronisme de floraison entre les plants et au sein de chaque plant. Un essai de pollinisation contrôlée montre que le goyavier-fraise présente une autocompatibilité partielle et variable entre les plants. Les distributions des nombres de graines par fruit ont été modélisées. Elles dépendent du plant, de l'origine génétique du pollen et des conditions de pollinisation. La modélisation de la pollinisation et de la fécondation dans une population hétérogène apparaît complexe et nécessite des études complémentaires. Une approche générale de la modélisation de l'effet de facteurs de croissance sur la courbe de croissance du fruit est présentée puis appliquée au goyavier-fraise, avec comme facteurs de croissance le nombre de graines du fruit et le rapport feuilles:fruits du plant. Les erreurs statistiques sont associées au modèle de prédiction pour estimer la variabilité des sorties. La distribution des dates de récolte à l'échelle du plant est estimée à partir du rapport feuilles:fruits du plant et du nombre moyen de graines dans ses fruits.

La démarche générale de cette étude est discutée, ainsi que les résultats obtenus. Des pistes de recherches complémentaires sont proposées. Les retombées pratiques de ces travaux sont soulignées. Enfin, la généralisation de l'approche agronomique de domestication et des modèles proposés est discutée.

Mots-clés: domestication, modèle, composantes du rendement, floraison, fécondation, croissance du fruit, *Psidium cattleianum*, Ile de la Réunion

Laboratoires d'accueil:

CIRAD-FLHOR, BP 180, 97455 Saint Pierre cedex, Ile de la Réunion

INRA-PSH, Bat A, Domaine St Paul, Site Agroparc, 84914 Avignon cedex 9

Summary

From flower to fruit: study and modelling of flowering, fecundation-fruit set, and fruit growth in strawberry guava (*Psidium cattleianum*).

Strawberry guava (*Psidium cattleianum*) domestication is studied within a project of crop diversification in humid highlands of Réunion Island. The number of fruit per plant and their individual weight determine the commercial yield. These characteristics are highly variable and do not seem to be linked to the genotype. An agronomic approach of domestication is proposed for this species. It consists in identifying how are built-up these characteristics and their variability through the different steps of the production cycle and under the effect of environmental, physiological and genetic factors. The objective is to act on some of these factors using appropriate cultivation techniques. Main features of this approach are the genetic diversity and the lack of references on the species, as confirmed by a bibliographical analysis.

Strawberry guava phenological cycle was studied on feral stands in Réunion Island. In cultivation, nitrogen modifies the natural phenological cycle by triggering a new production cycle. The setting of the study is a natural production cycle in an orchard, without interaction with triggered cycles. A model representing the process of fruit number and individual weight build-up is proposed, structured around five steps: number of flowers build-up, flowering, pollination, fecundation-fruit set, and fruit growth. The model outputs are the individual fruit weight and harvest dates distributions at the orchard level. Its functions are simulation and prediction in order to assist farmers in their technical choices. This work deals with the flowering, fecundation-fruit set and fruit growth steps.

The plant genetic diversity induces a flowering asynchrony which is not taken into account by flowering model conceived for a monovarietal orchard. A general stochastic flowering model is then built. It describes flowering at different levels and considers the flowering asynchrony both within and between plants. A controlled pollination experiment showed that strawberry guava is partially self-compatible, and that self-compatibility level varies among plants. The distribution of seed number per fruit depends on the individual, on the pollen genetic origin and on the pollination conditions. The modelling of pollination and fecundation in a genetically heterogeneous population appears complex and further research is needed. A general approach to model the effect of growth factors on individual fruit growth curve is presented. It is then applied to strawberry guava fruit growth, with the seed number per fruit and the plant leaf:fruit ratio as growth factors. Statistical errors are associated to the prediction model in order to estimate the output variability. The distribution of harvest dates at the plant level is estimated from the plant leaf:fruit ratio and the mean number of seeds per fruit.

The general approach used in this study and the results are discussed, and further research is outlined. Practical consequences of this work are underlined. Finally, the generalisation of the agronomic approach of domestication and of the models is discussed.

Key words: plant domestication, modelling, yield components, flowering, fecundation, fruit growth, *Psidium cattleianum*, Réunion Island.