



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Ecological Modelling 180 (2004) 407–417

ECOLOGICAL  
MODELLING

[www.elsevier.com/locate/ecolmodel](http://www.elsevier.com/locate/ecolmodel)

# SIMBA-POP: a cohort population model for long-term simulation of banana crop harvest

P. Tixier<sup>a,1</sup>, E. Malezieux<sup>b,\*</sup>, M. Dorel<sup>a</sup>

<sup>a</sup> CIRAD-FIthor, UMR SYSTEM, 97130 Capesterre, Guadeloupe, French West Indies, France

<sup>b</sup> CIRAD-DS, UMR SYSTEM, TA179/01, 34398 Montpellier Cedex 5, France

Received 11 July 2003; received in revised form 27 February 2004; accepted 29 April 2004

## Abstract

Banana crops represent a collection of individual plants that vegetatively propagate at their own rhythm, with stabilised but unsynchronised production of inflorescences over time. Such agrosystems cannot be simulated with existing crop models due to the unsynchronized behavior of individual plants. A new simulation model (SIMBA-POP), based on the cohort population concept, was built to predict phenological patterns of the population and harvest dynamics in banana cropping systems. The model was calibrated and validated for *Musa* spp., AAA group, cv. Cavendish Grande Naine with field data from the French West Indies (Guadeloupe and Martinique). It can quite accurately predict temporally-varying banana harvesting dynamics (date and number of harvested bunches). The harvesting peak is predicted with a precision less than 3 weeks for the first 3 cropping cycles. The model structure is based on two linear chains of cohorts characterized by both physiological age (heat unit accumulation) and development-stage dispersion in the banana population due to flowering, harvesting and sucker selection. This model is a valuable tool for both farmers (field management) and crop scientists (to simulate and design cropping systems). This approach is a first step towards long-term simulation in non-synchronized agrosystems that cannot be simulated with existing crop models.

© 2004 Elsevier B.V. All rights reserved.

**Keywords:** Banana; Cohort population model; Harvest prediction; Simulation

## 1. Introduction

Crop models are now commonly used to simulate biomass and yields on a field scale for a variety of crops. They are implemented for a range of applications, e.g. research, crop management and policy support. Crop growth models of the CERES (Jones

et al., 2003), APSIM (Keating et al., 2003) or STICS (Brisson et al., 1998) family are well adapted for simulating development and growth in homogeneous annual crop populations. However, they are unsuitable for the simulation of complex, multispecific or multiannual cropping systems that require consideration and simulation of plant individual asynchronism that can occur and increase with time. Tree population structure, forest stages or the variation of the number of trees in a forest can be simulated using individual based models (Philipps et al., 2003), Markov chain models (Benabdella et al., 2003) or malthus extinction models (Karev, 2003). In complex conditions, which

\* Corresponding author. Tel.: +33-4-67-61-55-46; fax: +33-4-67-61-55-12.

E-mail addresses: [tixier@cirad.fr](mailto:tixier@cirad.fr) (P. Tixier), [malezieux@cirad.fr](mailto:malezieux@cirad.fr) (E. Malezieux), [dorel@cirad.fr](mailto:dorel@cirad.fr) (M. Dorel).

<sup>1</sup> Tel.: +33-5-90-41-36-87; fax: +33-5-90-86-80-77.

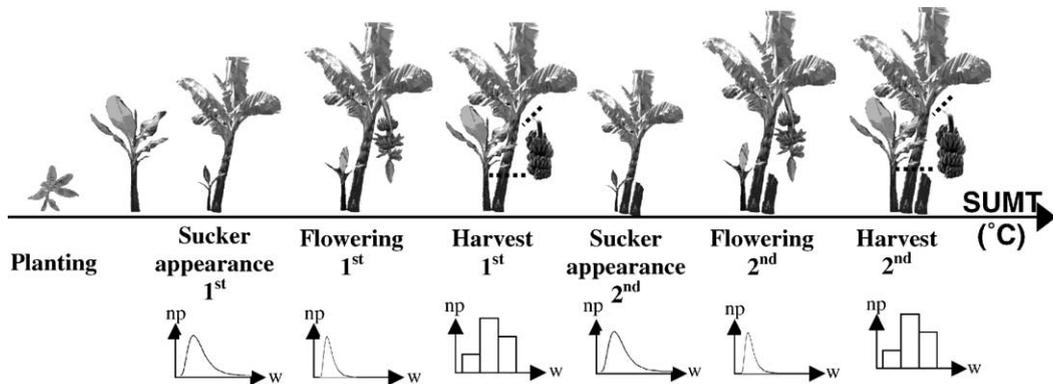


Fig. 1. Development stages of banana plant function of SUMT (heat unit accumulated by the crop in °C) and the six associated distributions expressed in the number of plants (np) per week (w) for two cropping cycles.

are common in perennial and/or multicrop systems, specific tools have to be built to predict yields and/or crop-environment relationships.

Bananas and plantains are rhizomatous herbs whose terminal bud produces the inflorescence. Each plant successively produces a series of bunches, each from a lateral shoot. The sequence can be repeated for 1–50 generations or more, which means that it can be considered as perennial (Turner, 1994). The main developmental stages of banana plants include sucker appearance, growth, flowering and harvest (Fig. 1). Banana crops represent a collection of individual plants derived from vegetative propagules. They develop at their own rhythm and do not follow a synchronous cycle. Hence, at any given time, a banana crop consists of a population of individual plants at various developmental stages.

The heterogeneity in the field population structure is due to both physiological phenomena (inter-plant variations in planting to flowering duration, in sucker emergence, death of plants and farmers' practices (type of selected sucker, moment of selection and replacement of plants). In the literature, sucker emergence and flowering distribution in a field have been described as log-normal distributions (Fig. 1) (Ganry, 1978; Lassoudière, 1979a,b,c, 1980; Cottin, 1987). The asymmetry of the flowering distribution may be linked to different factors such as spatial heterogeneity in the field, farming practices or plant stresses (drought, root parasitism, plant nutrition, etc.). Systems under stress may have a temporally spread flowering distribution. The earliness of the

“follower” (sucker of the following cycle) may vary according to the plant physiology, but the choice of follower depends on the farmer's decision criteria. The follower is chosen among existing suckers and its growth pattern depends on whether or not other shoots have been eliminated by the farmer. Growers also often replace dead or fallen plants, due to wind and poor rooting, with new planting material. Death ratio and the strategy of replacement also determine the heterogeneity in the population structure.

At the field scale, harvest dynamics (number of bunches harvested per week over time) are highly dependent upon inter-plant asynchronism. Harvest dynamics are characterized by oscillations whose amplitude and frequency vary with time, climate, and farmers' practices. An example of harvest dynamics (field data and simulation) is shown for one field in Martinique (Fig. 2). Farmers may want to concentrate harvests within a short timespan in the year (requires synchronous cycles), or spread harvests throughout the year (requires asynchronous cycles). The long-term behavior of this system must be understood and effectively simulated to be able to tailor field management strategies to meet farmers' objectives. The structure of the population in a field (homogeneity versus heterogeneity, proportion of fallen plants) determines its harvest dynamics and is one of the factors that influence farmers' decisions with respect to replanting a field (destruction of the crop, tillage). It also has a marked effect concerning on-farm labor organization. A specific representation of the system is thus needed to ensure appropriate management.

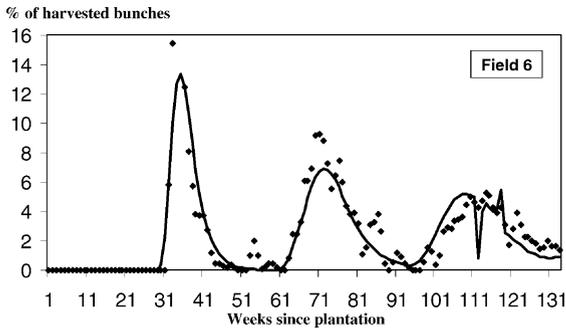


Fig. 2. Observed (◆) vs. simulated (—) harvest dynamics over time. Percentages (%) represent the ratio between the number of bunches harvested each week and the total banana crop population for field six (Martinique, F.W.I.; elevation: 315 m; cultivar: *Musa* spp., AAA group, cv. Cavendish Grande Naine).

In ecology, population models are often used to simulate the evolution of plant or animal population structures over time. Systems are also often represented as a collection of stocks linked by flows (Odum and Odum, 2000). These concepts were used to build SIMBA-POP, a model aimed at describing banana population structure over time by simulating the number of bunches harvested per week in a plot. The model takes into account the main agricultural practices, such as sucker selection, that influence plant management. Banana is an ideal system to study the linkage between a population structure and plant development because of the characteristics of banana crops, e.g. half-perennial, individual plant asynchronism, and short plant cycle.

## 2. Model structure

### 2.1. The cohort population dynamics concept

A cohort can be defined as a group of individuals characterized by the same phenological stage (Deaton and Winebrake, 2000). It is an appropriate unit for analyzing demographic processes in a population. A cohort chain is a group of cohorts linked to each other by flows controlled by various processes or laws (physiological, stochastic, etc.). The cohort population dynamics concept is commonly used in ecology to simulate animal population dynamics at different stages, e.g. insects

(Castañera et al., 2003; Cross and Crook, 1999; Hannon and Matthias, 1999). This concept was rarely used to represent vegetal systems (Goudriaan and van Roermund, 1989). We used it to simulate banana population dynamics. Cohort chains enable simulation of successive physiological stages of banana plant development, i.e. corm growth, inflorescence initiation, sucker emission, bunch harvest, and death of the main stem. These phenological stages can be assimilated by analogy with processes of birth, reproduction and death. Unlike existing banana models based on the mean plant concept (Mekwatanakarn, 1987), the cohort structure is well adapted for describing population dynamics in a banana field, where a cohort represents the number of individuals at a developmental stage for a given week. To be realistic, the model also has to take various environmental factors and agricultural practices that interact with the crop population and its development into account. The model presented here called SIMBA-POP, only includes the population dynamics and the main agricultural practices related to management of the population, such as sucker selection or plant replacement.

It simulates the dynamics of a banana plant population in a plot over numerous crop cycles. The main output is the number of bunches harvested per week over time. As for many plants, phenological stages are described by the accumulation of heat units that determine their durations (Bonhomme, 2000). The physiological development threshold of a banana plant was determined to be 14 °C (Ganry and Meyer, 1975; Ganry, 1980; Turner and Lahav, 1983; Turner and Hunt, 1983).

### 2.2. Mathematical structure

SIMBA-POP was developed using the STELLA® software environment from High Performance System® (Lebanon, NH, USA). It runs with a weekly step ( $t$ ) and is structured in two linked chains of cohorts that represent the plants before flowering (pre-flowering cohort chain) and plants after flowering (post-flowering cohort chain), respectively (Fig. 3). Variables used in SIMBA-POP are represented in Table 1. Variables refer to the whole population ( $X_t$ ) or to the ( $i$ ) cohort ( $X_{i,t}$ ). The only input data of the model is the mean daily temperatures ( $T_d$ ). Outputs include the number of flowering plants ( $FL_t$ ) and the

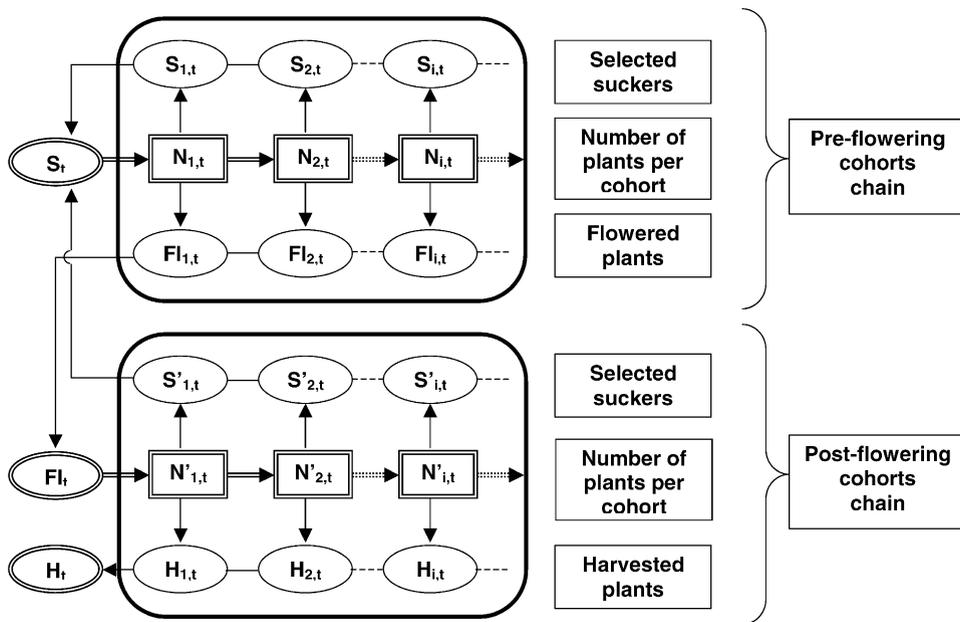


Fig. 3. The cohort structure of SIMBA-POP (refer Table 1 for variable meaning).

Table 1  
Variables used for SIMBA-POP simulation model

Input data	Definition
$T_d$	Daily average temperature of day, $d$ ( $^{\circ}\text{C}$ )
Output variables	Definition
$N_t$	Total number of plants in the simulated field at step ( $t$ )
$H_{i,t}$	Total number of harvested plants in cohort ( $i$ ) at step ( $t$ ) (in the post-flowering chain)
$H_t$	Total number of harvested plants in the simulated field at step ( $t$ )
$D_{i,t}$ and $D'_{i,t}$	Total number of dead plants in cohort ( $i$ ) at step ( $t$ )
$D_t$	Total number of dead plants in the simulated field at step ( $t$ )
$S_{i,t}$ and $S'_{i,t}$	Total number of suckers that appeared in cohort ( $i$ ) at step ( $t$ )
$S_t$	Total number of suckers that appeared in the simulated field at step ( $t$ )
$FL_{i,t}$	Total number of plants that reached the flower stage in cohort ( $i$ ) at step ( $t$ ) (in pre-flowering chain)
$FL_t$	Total number of plants that reached the flower stage in the simulated field at step ( $t$ )
$R_t$	Total number of plants to replace
$RR_t$	Total number of plants replaced at step ( $t$ )
$SUMT_{i,t}$	Heat units accumulated by cohort ( $i$ ) at step ( $t$ ) (day-degrees)
$SUMT_t$	Efficient temperature at step ( $t$ ) (day-degrees)
$s_{i,t}$	Sucker selection rate in cohort ( $i$ ) at step ( $t$ )
$WS_{i,t}$	Week since 1st week of sucker appearance when $SUMT_{i,t} >$ birth threshold
$fl_{i,t}$	Flowering rate of cohort ( $i$ ) at step ( $t$ )
$WF_{i,t}$	Week since 1st week of flowering when $SUMT_{i,t} >$ flowering threshold
$WH_{i,t}$	Week since 1st week of harvest when $SUMT_{i,t} >$ harvest threshold
$DTT_d$	Daily thermal temperature of day, $d$ (day-degrees)

$i$  is the number of the cohort in both pre-flowering and post-flowering chains;  $t$  the running step of the model in weeks;  $d$  the day of the simulated week.

Table 2  
Parameters used in SIMBA-POP

Parameters	Definition
$a_s, b_s, c_s$	Parameters of the sucker appearance rate log-normal curve
BT	Sucker appearance threshold (day-degrees)
$a_f, b_f, c_f$	Parameters of the flowering rate log-normal curve
FT	Flowering threshold (day-degrees)
$h_{i,t}$	Harvesting rate in cohort ( $i$ ) at step ( $t$ )
HT	Harvest threshold (day-degrees)
$d_{i,t}$	Death rate (falling) in cohort ( $i$ ) at step ( $t$ )
$r_t$	Replacement rate at step ( $t$ )
$T_{\text{banana}}$	Weekly physiologically active temperature threshold of banana ( $^{\circ}\text{C}$ )

$i$  is the number of the cohort in both pre-flowering and post-flowering chains;  $t$  is the running step of the model in weeks.

number of harvested plants ( $H_t$ ) for each step ( $t$ ). Parameters relative to plant physiology are given in Table 2.

Eqs. (1)–(4) calculate the number of plants in each cohort ( $i$ ) at each running step ( $t$ ) for both the pre-flowering chain ( $N_{i,t}$ ) and the post-flowering chain ( $N'_{i,t}$ ). In the pre-flowering chain, at each step, the pool of the cohort ( $N_{i,t}$ ) is transferred to the cohort ( $N_{i+1,t+1}$ ) minus the flowered plants ( $FL_{i,t}$ ), and the fallen and dead plants ( $D_{i,t}$ ) (Eq. (1)). In the post-flowering chain, at each step, the pool of the cohort ( $N'_{i,t}$ ) is transferred to ( $N'_{i+1,t+1}$ ) minus the harvested plants ( $H_{i,t}$ ) and dead plants ( $D'_{i,t}$ ) (Eq. (2)). New selected suckers in both chains ( $S_{i,t}$  and  $S'_{i,t}$ ) are transferred to the first cohort of the pre-flowering chain,  $N_{1,t}$  represents the pool of new planted and selected suckers in the plot (Eq. (3)). The flowered plants ( $FL_{i,t}$ ) are transferred to the first cohort of the post-flowering chain (Eq. (4)).

$$N_{i,t} = N_{i-1,t-1} - FL_{i-1,t-1} - D_{i-1,t-1} \quad (\text{pre-flowering chain}) \quad (1)$$

$$N'_{i,t} = N'_{i-1,t-1} - H_{i-1,t-1} - D'_{i-1,t-1} \quad (\text{post-flowering chain}) \quad (2)$$

$$N_{1,t} = \sum^n S_{i,t} + \sum^n S'_{i,t} + RR_t \quad (\text{pre-flowering chain}) \quad (3)$$

$$N'_{1,t} = \sum^n FL_{i,t} \quad (\text{post-flowering chain}) \quad (4)$$

Flowered plants ( $FL_{i,t}$ ), harvested plants ( $H_{i,t}$ ) and new selected suckers ( $S_t$  and  $S'_t$ ) at step  $t$  are calculated in Eqs. (5)–(7) and (7'). The flowering rate ( $fl_{i,t}$ ), harvesting rate ( $h_{i,t}$ ) and new sucker selection rate ( $s_{i,t}$ ) are calculated for each cohort in accordance with a probabilistic law whose parameters relate to plant physiology and farmers' practices. The parameters were determined from experimental field data (see calibration).

$$FL_{i,t} = N_{i,t} \times fl_{i,t} \quad (\text{pre-flowering chain}) \quad (5)$$

$$H_{i,t} = N_{i,t} \times h_{i,t} \quad (\text{post-flowering chain}) \quad (6)$$

$$S_t = S_{i,t} \times s_{i,t} \quad (\text{pre-flowering chain}) \quad (7)$$

$$S'_t = S'_{i,t} \times s_{i,t} \quad (\text{post-flowering chain}) \quad (7')$$

Fallen and dead plants for both chains ( $D_{i,t}$  and  $D'_{i,t}$ ) are calculated in Eqs. (8) and (8'). Falling depends on plant mass, parasitism that affect roots and wind strength. These factors are integrated in the death rate ( $d_{i,t}$ ). In the model presented here, links with other sub-models that simulate wind and root parasitism are not taken into account.

$$D_{i,t} = N_{i,t} \times d_{i,t} \quad (\text{pre-flowering chain}) \quad (8)$$

$$D'_{i,t} = N'_{i,t} \times d_{i,t} \quad (\text{post-flowering chain}) \quad (8')$$

The total number of plants on the plot ( $N_t$ ), the total number of flowered plants ( $FL_t$ ), new selected suckers ( $S_t$ ), the total number of harvested plants ( $H_t$ ) and fallen and dead plants ( $D_t$ ) are calculated for each step  $t$  in Eqs. (9)–(13), respectively.

$$N_t = \sum^n N_{i,t} + \sum^n N'_{i,t} \quad (\text{whole population}) \quad (9)$$

$$FL_t = \sum^n FL_{i,t} \quad (\text{whole population}) \quad (10)$$

$$S_t = \sum^n S_{i,t} + \sum^n S'_{i,t} \quad (\text{whole population}) \quad (11)$$

$$H_t = \sum^n H_{i,t} \quad (\text{whole population}) \quad (12)$$

$$D_t = \sum^n D_{i,t} + \sum^n D'_{i,t} \quad (\text{whole population}) \quad (13)$$

The number of plants to replace at step  $t$  ( $R_t$ ) is the sum of plants to replace at step  $t - 1$  and dead plants at step  $t$ , minus the replaced ones at step  $t - 1$  (Eq. (14)). The number of replaced plants ( $RR_t$ ) is calculated in Eq. (15) using a replacement factor ( $r_t$ ) in accordance with farmers' practices (Eq. (15)).

$$R_t = R_{t-1} + D_t - RR_{t-1} \quad (\text{whole population}) \quad (14)$$

$$RR_t = R_t \times r_t \quad (\text{whole population}) \quad (15)$$

Heat units accumulated at each step by each cohort ( $SUMT_{i,t}$  in 'degree-days') are calculated in Eq. (16). Heat units accumulated for each week ( $SUMT_t$  in °C) is calculated in Eqs. (16') and (16''), where daily thermal time ( $DTT_d$ ) is the number of heat unit accumulated during one day,  $T_{\text{banana}}$  the temperature below which development is assumed to cease, and  $T_d$  the mean day temperature.

$$SUMT_{i,t} = SUMT_{i-1,t-1} + SUMT_t \quad (\text{whole population}) \quad (16)$$

$$SUMT_t = \sum_{(d=1 \text{ to } 7)} DTT_d \quad (\text{whole population}) \quad (16')$$

$$DTT_d = (T_d - T_{\text{banana}}) \quad (\text{whole population}) \quad (16'')$$

### 2.3. Stochastic laws

Phenological stages were considered as being related to these heat units accumulated by each cohort. Flowering, sucker selection and harvest were assumed to follow a log-normal curve according to results reported by Ganry (1978), Lassoudière (1979c), Cottin (1987). Hence, stochastic laws were used to establish the rates of flowering ( $fl_{i,t}$ ), sucker selection ( $s_{i,t}$ ) and harvest ( $h_{i,t}$ ). We assumed they apply when the corresponding threshold for heat units accumulated is reached. When applied, these laws follow a log-normal curve function.

Flowering occurs in a cohort when the heat units accumulated is over the FT threshold. The flowering rate ( $fl_{i,t}$ ) of the cohort ( $i$ ) at step ( $t$ ) follows a log-normal curve function according to the number of weeks since first flowering ( $WF_{i,t}$ ) with  $a_f$ ,  $b_f$  and  $c_f$  parameters (Eqs. (17), (17') and (17'')).

$$fl_{i,t} = a_f \times \exp \left( -0.5 \left( \frac{\ln(WF_{i,t}/b_f)}{c_f} \right)^2 \right) \quad (17)$$

$$\text{if } SUMT_{i,t} < FT \text{ then } (WF_{i,t} = 0) \quad (17')$$

$$\text{if } SUMT_{i,t} \geq FT \text{ then } (WF_{i,t} = WF_{i-1,t-1} + 1) \quad (17'')$$

Sucker selection occurs in a cohort when the heat units accumulated is above the BT threshold. The sucker selection rate ( $s_{i,t}$ ) of the cohort ( $i$ ) at step ( $t$ ) follows a log-normal curve function according to the number of weeks since first sucker selection ( $WS_{i,t}$ ), with  $a_s$ ,  $b_s$  and  $c_s$  parameters (Eqs. (18), (18') and (18'')).

$$s_{i,t} = a_s \times \exp \left( -0.5 \left( \frac{\ln(WF_{i,t}/b_s)}{c_s} \right)^2 \right) \quad (18)$$

$$\text{if } SUMT_{i,t} < BT \text{ then } (WS_{i,t} = 0) \quad (18')$$

$$\text{if } SUMT_{i,t} \geq BT \text{ then } (WS_{i,t} = WS_{i-1,t-1} + 1) \quad (18'')$$

Harvest occurs in a cohort when the sum of accumulated heat units is above the HT threshold. The harvesting rate ( $h_{i,t}$ ) of the cohort ( $i$ ) at step ( $t$ ) follows a log-normal curve function according to the number of weeks since first flowering ( $WH_{i,t}$ ) (Eqs. (19), (19') and (19'')). According to farmers' practices, this curve is summarized by a 3-week histogram, as described in Eq. (19).

$$\begin{aligned} &\text{if } (WH_{i,t} = 1 \text{ then } h_{i,t} = 0.2) \text{ else} \\ &(\text{if } WH_{i,t} = 2 \text{ then } h_{i,t} = 0.625) \text{ else} \\ &(\text{if } WH_{i,t} = 3 \text{ then } h_{i,t} = 1) \text{ else } (h_{i,t} = 0) \end{aligned} \quad (19)$$

$$\text{if } SUMT_{i,t} < HT \text{ then } (WH_{i,t} = 0) \quad (19')$$

$$\text{if } SUMT_{i,t} \geq HT \text{ then } (WH_{i,t} = WH_{i-1,t-1} + 1) \quad (19'')$$

### 3. Parameter calibration and model validation

Calibration and validation data were collected in banana fields in Guadeloupe and Martinique (French West Indies). All data are presented in Table 3. Plants were grown in field conditions with Cavendish cv. Grande Naine (*Musa* spp., AAA group, cv. Cavendish Grande Naine) with no major stress (no hurricanes, drought or parasitism damage).

Table 3  
Fields used for calibration and validation of SIMBA-POP

Number	Island	Field name	Elevation (m)	Observation dates	Number of observed weeks
1	Martinique	HANGA	325	8 July, 1997–19 August, 2001	215
2	Martinique	MC07	350	16 June 1998–3 March, 2002	194
3	Martinique	MJO1	325	13 May, 1998–3 March, 2002	199
4	Martinique	FPAIN	300	27 June, 1996–5 November, 2000	227
5	Martinique	LIS 2	315	20 June, 1997–3 March, 2003	297
6	Martinique	MC04	315	19 January, 1995–27 July, 1997	131
7	Guadeloupe	JARDIN	45	1 October, 1996–19 December, 2001	272
8	Martinique	DACHI	260	16 November, 1995–17 February, 2002	326
9	Martinique	FSAVA	160	30 May, 1997–24 February, 2002	247
10	Martinique	GOYAVE	290	10 June, 1996–16 December, 2001	288
11	Martinique	MC06	315	19 May, 1998–17 February, 2002	196
12	Martinique	GCHEM	340	18 October, 1997–3 March, 2002	228
13	Martinique	COMPL	335	30 October, 1996–3 December, 2000	214
14	Martinique	CVERT	295	24 April, 1997–21 October, 2001	234

Number of flowered and harvested plants (*Musa* spp., AAA group, cv. Cavendish Grande Naine) and temperature were recorded weekly. Fields were located in Guadeloupe (16°15'N, 61°32'W) and Martinique (14°36'N, 61°5'W), French West Indies.

### 3.1. Calibration

The flowering rate ( $fl_{i,t}$ ) was fitted to a log-normal curve using flowering data for the first cycle (fields 1–6). The parameters are:  $a_f = 0,12$ ,  $b_f = 6,35$  and  $c_f = 0,47$  (Fig. 4). The FT threshold was calibrated using the sum of heat units accumulated between planting and first flowering for fields 1–7. FT is 2350 degree-days (S.D. = 223; 7 fields).

The sucker selection rate ( $s_{i,t}$ ) was calibrated by adjusting the sucker selection parameters to minimize the error between simulation and field data

for all harvesting peaks for each calibration field. The log-normal curve parameters were calibrated for fields 1–7 ( $a_s = 1,00$ ,  $b_s = 3,25$  and  $c_s = 0,31$ , respectively). The BT threshold was calibrated using the sum of heat units accumulated between planting and first sucker appearance. BT was 2180 degree-day (S.D. = 156; 7 fields). The harvest threshold HT was fixed to 900 degree-days according to data from the literature (Ganry, 1978).

The replacement rate ( $r_t$ ) and death rate ( $d_t$ ), which are, respectively related to farmers' practices and the plant environment (wind and para-

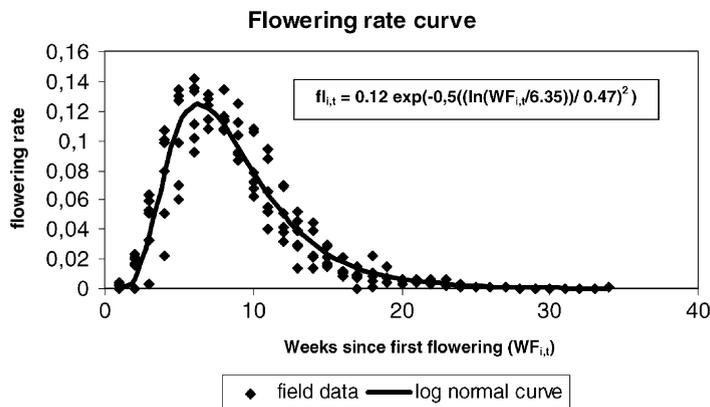


Fig. 4. Calibration of the flowering rate curve ( $fl_{i,t}$ ) using first cycle data from fields 1 to 6 (Martinique F.W.I.), as a function of the number of weeks since the first flowering of cohort ( $i$ ) at step ( $t$ ) ( $WF_{i,t}$ ).

sitism), were not considered or calibrated in this version of the model. Sub-models necessary to take falling and replacement into account are discussed below.

### 3.2. Validation

Banana harvest dynamics at the crop scale are characterized by oscillations over time (Fig. 2). The dynamic aspect of these outputs requires specific statistic tools for model validation. The predictive accuracy of the model depends on reproducing the harvest dynamics and magnitude of harvesting peaks through successive cycles. In this regard, the cumulated number of harvested plants and the capacity of the model to predict successive harvesting peak dates are suitable for long-term validation.

Simulations were run for seven plots (field 8–14, Table 3). Both the relation between the observed and simulated cumulated number of bunches harvested over time and the ability of the model to predict harvesting peaks (date of the maximal number of bunches harvested for each cropping cycle) were used to validate the model.

Fig. 5 shows that the model can quite accurately predict ( $r^2 = 0.9858, n = 1058$ ) the cumulated number of bunches harvested after many cycles for fields 8–14.

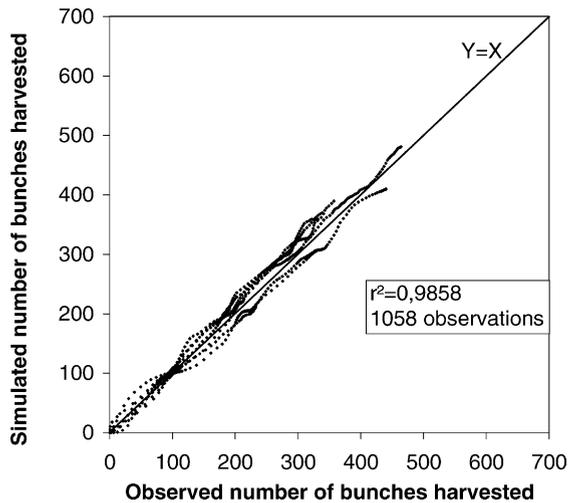


Fig. 5. Simulated vs. observed cumulated number of harvested bunches per week (%) for 7 fields (field 8–14; Martinique F.W.I.).  $R^2$  was calculated using all data for the 7 fields.

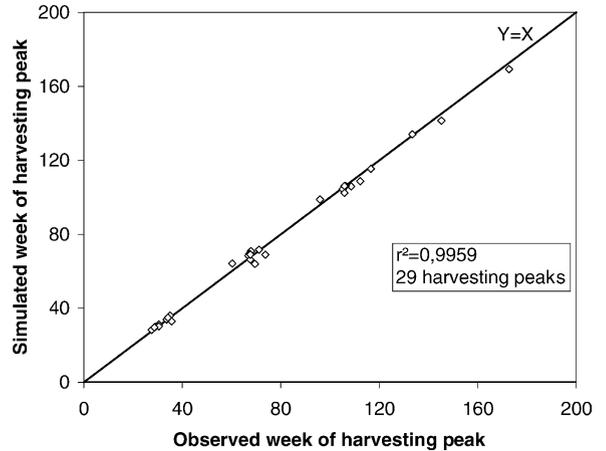


Fig. 6. Simulated vs. observed week of harvesting peaks for fields 8–14 (Martinique F.W.I.).

The fact that the model error is distributed equally around the median curve ( $Y = X$ ) indicates that the model does not systematically lead to underestimation or overestimation.

Harvesting peak dates (29 harvesting peaks) were estimated for field data (8–14) using a log-normal curve fitting. In the log-normal equation:  $F(X) = a \times \exp(-0.5((\ln(X/b))/c)^2)$ , the  $b$  parameter represents the abscissa of the curve peak. Fig. 6 shows successive harvesting peaks for simulated versus observed data.

Complementary statistic methods were used to evaluate the predictive accuracy of the model to determine harvesting peak dates. The mean deviation (MD) between observed and simulated values is a measurement of the tendency of the model to over or underestimate the observed value. MD is defined as:

$$MD = \left(\frac{1}{n}\right) \left(\sum ERR_i\right)$$

where  $ERR_i = Y_i - X_i$  and  $Y_i$  and  $X_i$  are the  $i$ th predicted and observed value, respectively, where  $n$  is the sample size. The mean predicted error (MPE) is defined as:

$$MPE = \left(\frac{1}{n}\right) \left(\sum |ERR_i|\right)$$

MPE is an indicator of the accuracy of the model. The mean square error of the prediction (MSEP) is a

Table 4  
Statistical results relative to the simulation of the harvesting peak date with SIMBA-POP

Statistic method	All peaks	1st cycle peak	2nd cycle peak	3rd cycle peak
MD	−0.38	0.29	−0.25	−1.08
MPE	1.92	0.94	2.30	1.88
MSEP	5.4	1.35	8.20	5.30
RMSEP	2.42	1.16	2.86	2.30

MD: mean deviation; MPE: mean square error; MSEP: mean square error of the prediction; RMSEP: root mean square error predicted.

good criterion for comparing the predictive accuracy of alternative models (Wallach and Goffinet, 1989). MSEP is the mean squared deviation of the predicted values from the observed values:

$$\text{MSEP} = \left(\frac{1}{n}\right) \sum (\text{ERR}_i)$$

The root mean square error predicted (RMSEP), expressed in the same unit as the observed values (weeks), is defined as:

$$\text{RMSEP} = \sqrt{\text{MSEP}}$$

These statistics were applied to the simulated and observed harvesting peak dates. The first three harvesting peaks were taken into consideration in fields 8–14. Table 4 shows these validation statistic results (26 cropping cycles analyzed).

These results show that the model is able to predict the first three cycles with a relatively good accuracy. RMSEP is always less than 3 weeks and is stable over time. This confirms the ability of the model to predict the long-term harvesting peak dates. It is important to note that even though the accuracy of the model decreases with time, less accuracy is necessary in the long term because harvesting peaks become more spread out.

#### 4. Application of SIMBA-POP and perspectives

SIMBA-POP potentially has various applications for scientists and farmers. It may be used by scientists to evaluate different agricultural practices relative to the management of a banana popula-

tion in a field. It can also be linked with other sub-models to make realistic long-term simulations of cropping system performance. SIMBA-POP can also be used as a predictive tool for farmers to optimize practices and be a powerful tool for long-term management (labour per cycle, harvesting periods, etc.).

##### 4.1. SIMBA-POP: a tool to test and optimize population management decision criteria

SIMBA-POP can be linked with specific sub-models that simulate farmers' practices and then be used to test new population management decision criteria. In banana cropping systems, main decision criteria relative to the crop population management are sucker selection and dead (fallen) plant replacement. The sucker selection criterion is defined by a set of cropping actions geared to removing unwanted suckers and to keeping only one follower (with expected characteristics that include spatial location, shape and size). Testing a new set of parameters (sucker selection parameters) is an interesting use of the model, with the aim of potentially designing new cropping systems. Fig. 7 shows harvest dynamics for two sucker selection strategies: "synchronous cycle strategy" with narrow harvesting peaks, and "asynchronous cycle strategy" with broad and earlier harvesting peaks. The synchronous cycle strategy is based on late selection (the sucker selection BT threshold is high) and selection occurs over a short period of time (the sucker selection rate curve is narrow). In contrast, the asynchronous strategy is based on early selection (BT is small), but the selection period is temporally extended (the sucker selection rate curve is broad). In these two simulations, the sucker selection strategy affects the harvesting dynamics beyond the first cycle. SIMBA-POP can help farmers to select crop management strategies to meet their specific objectives and conduct virtual tests. For example, in choosing a planting date and sucker selection criteria to optimise benefits in relation with seasonal variations in banana market prices or labour availability. The model can help to optimise these choices throughout many cropping cycles, not only for the first cycle and not only on the basis of empirical knowledge, as is currently done. With a link to a sub-model that controls the replacement of fallen plants criteria, SIMBA-POP can

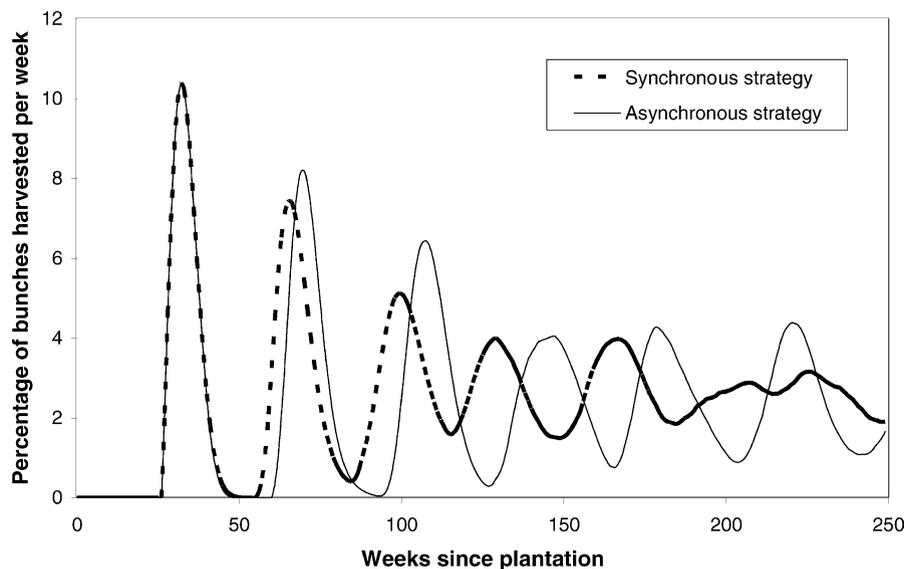


Fig. 7. Percentage of bunches harvested per week relative to the initial number of plants:  $(H_t/N_{1,1}) \times 100$ ; with  $H_t$  the total number of bunches harvested at step  $t$  and  $N_{1,1}$  the number of planted plants in cohort 1 at step 1; simulated for two sucker selection strategies: “synchronous strategy” with late pruning (sucker appearance threshold BT is high) and narrow sucker selection rate curve of cohort ‘ $i$ ’ at step ‘ $t$ ’ ( $s_{i,t}$  is low) and “asynchronous strategy” with early pruning (‘BT’ is low) and broad selection curve ( $s_{i,t}$  is high).

be used to test new criteria to control heterogeneity in a plot.

#### 4.2. Perspectives: linkage with other sub-models to simulate the whole cropping system

The population model SIMBA-POP is the first step of a larger model that aims at simulating the whole cropping system over the long-term. Simulation of the plant population structure in a field can provide very important information for assessing some environmental and agronomic variables of cropping systems. Long-term simulations of cropping systems requires linkage of population models with crop growth model. The development of such models involves some structural issues. In the case of SIMBA-POP, the double cohort chain structure allows easy linkage with growth sub-models to each sub-population (each cohort). In contrast with a model based on linear equations that describe the population structure, the cohort chain structure is flexible because all variables of each sub-population are accessible. First SIMBA-POP can be improved by taking into account the falling rate, linked to wind, parasitism and population manage-

ment criteria, as described earlier. It also can be linked to plant growth, biomass and yield sub-models.

## 5. Conclusion

This study highlights several aspects of SIMBA-POP, a population model based on the cohort concept. The model was successfully calibrated and validated for banana with data from Guadeloupe and Martinique. Although there are some limits under stressed situations or high falling rates, the model can be used to predict harvesting periods over cropping cycles. It also can be used to simulate or compare new population management decision criteria. This study also shows that population models are useful for long-term simulation of agrosystems with unsynchronised plants. In this case, linkage with sub-models that simulate water balance, plant growth and soil parasitism will enable realistic long-term simulation of cropping systems in the near future. This approach could be extended to systems with asynchronous population dynamics behavior, especially for perennial or half-perennial systems based on vegetative propagation.

## Acknowledgements

We are particularly grateful to the farmers of Guadeloupe and Martinique that helped us to build the plant database. We would also like to thank our colleagues from CIRAD-FIhor who helped us to contact farmers and provided us with temperature data (especially André Lassoudière).

## References

- Benabdella, B., Albrecht, K.F., Pomaz, V.L., Denisenko, E.A., Logofet, D.O., 2003. Markov chain models for forest successions in the Erzgebirge, Germany. *Ecol. Model* 159, 145–160.
- Bonhomme, R., 2000. Bases and limits to using 'degree-day' unit. *Eur. J. Agron.* 12, 1–10.
- Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonielli, R., Durr, C., Richard, G., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J.M., Meynard, J.M., Delécolle, R., 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie* 18, 311–346.
- Castañera, M.B., Aparicio, J.P., Gürtler, R.E., 2003. A stage-structured stochastic model of the population dynamics of *Triatoma infestans*, the main vector of Chagas disease. *Ecol. Model* 162, 33–53.
- Cottin, R., 1987. Modélisation de la production bananière. Influence de quelques paramètres en Martinique. *Fruits* 42 (12), 691–701.
- Cross, J.V., Crook, D.J., 1999. Predicting spring emergence of blackcurrant leaf midge (*Dasineura tetensi*) from air temperatures. *Entomologia Experimentalis et Applicata* 91, 421–430.
- Deaton, M.L., Winebrake, J.J., 2000. Modeling mobile source air pollution inventories. In: M.L. Deaton, J.J. Winebrake (Eds.), *Dynamic Modeling of Environmental Systems*. Springer-Verlag, New York, pp. 142–157.
- Ganry, J., 1978. Recherche d'une méthode d'estimation de la date de récolte du bananier à partir de données climatiques dans les conditions des Antilles. *Fruits* 33 (10), 669–679.
- Ganry, J., 1980. Note de synthèse: Le développement du bananier en relation avec les facteurs du milieu: Action de la température et du rayonnement d'origine solaire sur la vitesse de croissance des feuilles. Etude du rythme de développement de la plante. *Fruits* 35 (12), 727–744.
- Ganry, J., Meyer, J.P., 1975. Recherche d'une loi d'action de la température sur la croissance des fruits du bananier. *Fruits* 30 (6), 375–392.
- Goudriaan, J., van Roermund, H.J.W., 1989. Modelling of ageing, development, delays and dispersion. In: Rabbinge, R., Ward, S.A., Van Laar H.H. (Eds.), *Simulation and Systems Management in Crop Protection*. Pudoc, Wageningen, pp. 47–79.
- Hannon, B., Matthias, R., 1999. Multistage insect model. In: Hannon, B., Ruth, M. (Eds.), *Modelling Dynamic Biological Systems*. Springer-Verlag, New York, pp. 170–180.
- Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. *Eur. J. Agron.* 18, 235–265.
- Karev, G.P., 2003. Inhomogeneous models of tree stand self-thinning. *Ecol. Model* 160, 23–37.
- Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* 18, 267–288.
- Lassoudière, A., 1979a. Comportement du bananier Poyo au second cycle. I: Rejetonnage et multiplication végétative. *Fruits* 34 (11), 645–727.
- Lassoudière, A., 1979b. Comportement du bananier Poyo au second cycle. II: Aspects généraux. *Fruits* 34 (12), 713–727.
- Lassoudière, A., 1979c. Comportement du bananier Poyo au second cycle. IV: Mise en évidence d'interactions entre rejet et pied mère et entre rejets frères. *Fruits* 35 (2), 69–93.
- Lassoudière, A., 1980. Comportement du bananier Poyo au second cycle. III: Etude d'une population. *Fruits* 35 (1), 3–17.
- Mekwatanakarn, W., 1987. A simulation model of production and flowering in bananas. Thesis of the University of Western Australia, p. 139.
- Odum, H.T., Odum, E.C., (Eds.), 2000. Energy systems diagramming. In: *Modelling for All Scales: An Introduction to System Simulation*. Academic Press, San Diego. pp. 19–38.
- Philipps, P.D., Brash, T.E., Yasman, I., Subagyo, P., van Gardingen, P.R., 2003. An individual-based spatially explicit tree growth model for forests in East Kalimantan (Indonesian Borneo). *Ecol. Model* 159, 1–26.
- Turner, D.W., 1994. Banana and plantain. In: *Environmental Physiology of Fruit Crop*. Vol. 2, 37–64, CRC Press. pp. 206–229.
- Turner, D.W., Hunt, N., 1983. The relationship between temperature and the rate of appearance of new leaves on thirty banana varieties grown in the subtropics. *Garcia de Orta, Sér. Est. Agron.* 10 (1–2), 91–94.
- Turner, D.W., Lahav, E., 1983. The growth of banana plants in relation with temperature. *Aust. J. Plant Physiol.* 10, 43–53.
- Wallach, D., Goffinet, B., 1989. Mean squared error of prediction as a criterion for evaluating and comparing system models. *Ecol. Model* 44, 299–306.