

Qualification of a Plant Disease Simulation Model: Performance of the LATEBLIGHT Model Across a Broad Range of Environments

Jorge L. Andrade-Piedra, Gregory A. Forbes, Dani Shtienberg, Niklaus J. Grünwald, María G. Chacón, Marco V. Taïpe, Robert J. Hijmans, and William E. Fry

First and eighth authors: Department of Plant Pathology, Cornell University, Ithaca, NY 14853; second author: International Potato Center (CIP), P.O. Box 1558, Lima 12, Peru; third author: Department of Plant Pathology, ARO, The Volcani Center, Bet Dagan 50250, Israel; fourth author: Horticultural Crops Research Laboratory, USDA-ARS, 3420 NW Orchard Ave., Corvallis, OR 97330; fifth and sixth authors: International Potato Center (CIP), P.O. Box 17-21-1977, Quito, Ecuador; and seventh author: Museum of Vertebrate Zoology, University of California, 3101 Valley Life Sciences Building, Berkeley 94720.

Current address of J. L. Andrade-Piedra: International Potato Center (CIP), P.O. Box 17-21-1977, Quito, Ecuador.

Current address of M. V. Taïpe: Facultad de Ciencias Aplicadas, Escuela Superior Politécnica del Ejército (ESPE), Sangolquí, Ecuador.

Accepted for publication 5 August 2005.

ABSTRACT

Andrade-Piedra, J. L., Forbes, G. A., Shtienberg, D., Grünwald, N. J., Chacón, M. G., Taïpe, M. V., Hijmans, R. J., and Fry, W. E. 2005. Qualification of a plant disease simulation model: Performance of the LATEBLIGHT model across a broad range of environments. *Phytopathology* 95:1412-1422.

The concept of model qualification, i.e., discovering the domain over which a validated model may be properly used, was illustrated with LATEBLIGHT, a mathematical model that simulates the effect of weather, host growth and resistance, and fungicide use on asexual development and growth of *Phytophthora infestans* on potato foliage. Late blight epidemics from Ecuador, Mexico, Israel, and the United States involving 13 potato cultivars (32 epidemics in total) were compared with model predictions using graphical and statistical tests. Fungicides were not applied in any of the epidemics. For the simulations, a host resistance

level was assigned to each cultivar based on general categories reported by local investigators. For eight cultivars, the model predictions fit the observed data. For four cultivars, the model predictions overestimated disease, likely due to inaccurate estimates of host resistance. Model predictions were inconsistent for one cultivar and for one location. It was concluded that the domain of applicability of LATEBLIGHT can be extended from the range of conditions in Peru for which it has been previously validated to those observed in this study. A sensitivity analysis showed that, within the range of values observed empirically, LATEBLIGHT is more sensitive to changes in variables related to initial inoculum and to weather than to changes in variables relating to host resistance.

Additional keywords: confidence interval test, envelope of acceptance test, metamodeling, revalidation.

Plant disease simulation models are valuable tools in epidemiology. They organize available knowledge such that information gaps become clear. They also enable one to manipulate variables affecting disease in order to generate hypotheses. Finally, they can be used to illustrate epidemiological principles for educational purposes (8). However, many models have not achieved these objectives. Campbell and Madden (8) suggest that the primary reason is that model development has been considered as an end in itself, rather than as an approach to reach the objectives mentioned above. Even for models that have surpassed the development stage and have been used in research programs, it can be difficult for the creators of the model to achieve broad adoption and use. There are several obstacles to model use, including the need to establish confidence in the validity of the model. Models used in research programs are validated by necessity, but those models are sometimes only validated for very restricted conditions. Therefore, if a certain model is intended for extensive use, it is important to test its validity for a broad range of conditions to determine its domain of applicability (33,35). This is referred to as model qualification.

Model qualification has been defined as “discovering the domain over which a validated model may be properly used, i.e.,

whether the model is acceptable for use in a more general context, and amounts to revalidating a model for new cases” (33). If the model passes the revalidation tests, then its domain of applicability increases; if it fails, then its domain is qualified, i.e., restricted to those conditions for which it was originally validated (33).

One constraint for model qualification is the need to have the proper data for the expanded set of conditions. In that sense, the potato-*Phytophthora infestans* pathosystem offers a unique opportunity, because epidemics of the resulting disease, potato late blight, have been carefully recorded in many places of the world. In this study, we use the LB2004 version (3,4) of the LATEBLIGHT model (6,7) to illustrate the concept of model qualification for plant disease simulation models.

LATEBLIGHT simulates the effect of weather, host growth and resistance, and fungicide use on asexual development and growth of *P. infestans* on potato foliage. The LB2004 version of LATEBLIGHT includes improved equations for the effect of temperature on lesion growth rate (LGR) (Table 1 provides a list of acronyms and abbreviations), sporulation rate (SR), and latent period (LP) (3). It was also parameterized to reflect LGR, SR, and LP in three Peruvian potato cultivars infected with isolates of a new clonal lineage of *P. infestans* (EC-1, sensu Forbes et al. [15]) that is currently predominant in many regions of Peru (31). The model was then validated using data from field experiments located in the Peruvian Andes. Under these conditions, LB2004 accurately predicted epidemics of late blight (4).

We intend to use this model to illustrate epidemiological principles of disease management to diverse audiences and to improve

Corresponding author: J. L. Andrade-Piedra; E-mail address: j.andrade@cgiar.org

DOI: 10.1094/PHYTO-95-1412

© 2005 The American Phytopathological Society

MATERIALS AND METHODS

and focus field experiments by developing and testing hypotheses about the management of partially resistant cultivars in the Andes and other tropical highlands. Therefore, it is critical that the model be evaluated using conditions beyond those for which it has been validated. In the present study, we validated LB2004 in a number of environments, for additional potato cultivars, and for additional *P. infestans* populations in order to further qualify the model. We used observed data from a large sample of epidemics in which no fungicides were applied. These included 27 epidemics from the tropical highlands (Ecuador and Mexico), two epidemics from a subtropical region (Israel), and three epidemics from a temperate region (United States). Field data were first analyzed to determine their variance and to make certain that there were no obvious errors (4). Then, observed and simulated data were compared using graphical and statistical tests (4). The level of agreement between model output and reality was assessed by comparison with subjective and objective performance criteria (4). We also conducted a sensitivity analysis (34) to gain a better understanding of the behavior of LB2004 when various parameters and input variables are changed within the range of environments considered in this study.

Field experiments. Trials were conducted in four locations in Ecuador, two locations in Israel, one location in Mexico, and one location in the United States (Table 2). Thirteen potato cultivars with different levels of resistance against *P. infestans* (Table 3) were used in 15 experiments, for a total of 32 epidemics (Table 4). The number of cultivars evaluated in each experiment varied from one to three (Table 4). Mean air temperature at the coldest site was 9.5°C ($n = 145$ days, Pinantura, Ecuador, 1998) and 19.0°C in the warmest site ($n = 51$ days, Freeville, United States, 2002); and the mean period of high humidity (described below) varied from 12.9 to 18.8 h (Table 4). The locations ranged in latitude from 0°27'55''S in Pinantura (Ecuador) to 42°30'50''N in Freeville (United States), and in altitude from approximately 50 m above sea level in Mishmeret (Israel) to 3,543 m above sea level in Pinantura (Table 2).

Potato plants were grown in plots of diverse areas: between 7.0 m² (4.4 × 1.6 m) and 37.8 m² (4.5 × 8.4 m) in Ecuador; 26.6 m² (3.8 × 7.0 m) in Israel; 14.8 m² (3.7 × 4.0 m) in Mexico; and 16.6 m² (3.6 × 4.6 m) in the United States. Plant density (plants per square meter) ranged between 2.1 and 3.1 in Ecuador,

TABLE 1. Acronyms and abbreviations used in this study

| Acronym or abbreviation | Definition | Unit |
|-------------------------|---|---|
| AUDPC _{obs} | Observed area under the disease progress curve | Percent-days |
| AUDPC _{obs} | Mean AUDPC _{obs} | Percent-days |
| AUDPC _{sim} | Simulated area under the disease progress curve | Percent-days |
| AUDPC _{sim} | Mean AUDPC _{sim} | Percent-days |
| CIT | Confidence interval test | – |
| DAE | Days after emergence | – |
| DOI | Day of initiation of the simulated epidemic | Days after emergence |
| EAT | Envelope of acceptance test | – |
| EC-1 | Clonal lineage of <i>P. infestans</i> | – |
| H | Daily period when relative humidity was above 85% (sensor located above the potato canopy) or 90% (sensor located within the potato canopy) | h |
| H̄ | Mean H during the experiments | h |
| IE | Infection efficiency | Proportion (Table 5) |
| IL | Initial number of lesions for the simulated epidemic | Lesions m ⁻² of ground |
| LB2004 | Most recent version of LATEBLIGHT (3) | – |
| LGR | Lesion growth rate | m day ⁻¹ |
| LP | Latent period | Days |
| RAUDPC _{obs} | Observed relative area under the disease progress curve | % |
| RAUDPC _{obs} | Mean RAUDPC _{obs} | % |
| RH | Relative humidity | % |
| S | Severity of late blight | % |
| sAUDPC | Standard AUDPC | % |
| SD | Standard deviation | – |
| SR | Sporulation rate | Sporangia m ⁻² day ⁻¹ |
| T | Mean daily air temperature | °C |
| T̄ | Mean T during the experiments | °C |
| TH | Mean daily air temperature when relative humidity was above 85% (sensor located above the potato canopy) or 90% (sensor located within the potato canopy) | °C |
| t ₀ | Last evaluation on which no lesions of <i>Phytophthora infestans</i> were observed | Days after emergence |
| t ₁ | First evaluation on which lesions of <i>P. infestans</i> were observed | Days after emergence |

TABLE 2. Locations of experiments used to qualify the LB2004 version of LATEBLIGHT

| Country | Location ^a | Altitude ^b | Latitude | Longitude |
|---------------|-----------------------|-----------------------|-------------|-------------|
| Ecuador | Amaguaña | 2,756 | 0°22'18''S | 78°30'51''W |
| | Cutuglahua | 3,058 | 0°22'12''S | 78°33'23''W |
| | La Tola | 2,508 | 0°13'51''S | 78°22'10''W |
| | Pinantura | 3,543 | 0°27'55''S | 78°18'52''W |
| Israel | Mishmeret | ≈50 | 32°18'N | 34°56'E |
| | Nir-Eliyahu | ≈50 | 32°18'N | 34°56'E |
| Mexico | Toluca | 2,640 | 19°14'35''N | 99°35'23''W |
| United States | Freeville | 319 | 42°30'50''N | 76°20'49''W |

^a Amaguaña, Cutuglahua, La Tola, and Pinantura are located in the province of Pichincha; Mishmeret and Nir-Eliyahu are close to the city of Natanya; Toluca is in the state of Mexico; and Freeville is in the state of New York.

^b Meters above sea level.

3.5 in Israel, 4.3 in Mexico, and 3.8 in the United States. Plots (experimental units) were randomly allocated to each cultivar in randomized complete block (Ecuador, Israel, and Mexico) or completely randomized designs (United States). To avoid interplot interference, plots were separated from each other by 3.0 to 4.0 m of a cover crop (barley or oats) in Ecuador, and 4.0 m of fallow soil in Mexico and the United States. In Israel, plots were located within a commercial potato crop; no space was left between plots; and the disease was evaluated in a central area of 9.5 m² (1.9 × 5.0 m). The number of repetitions varied from three to five (Table 4).

TABLE 3. Cultivars used to qualify the LB2004 version of LATEBLIGHT

| Cultivar ^a | CIP no. ^b | Reaction to <i>P. infestans</i> ^c | Country ^d | Reference |
|-----------------------|----------------------|--|----------------------|------------------|
| Alpha | 800050 | S | Mexico | (21,22) |
| Bolona | 703268 | S | Ecuador | (10) |
| INIAP-Cecilia | 720136 | S | Ecuador | (10) |
| INIAP-Gabriela | 720120 | S | Ecuador | (10) |
| Katahdin | 800102 | S | USA | (26,32) |
| LBR37 | 387132.2 | S | Ecuador | ... ^e |
| Nicola | ... | S | Israel | (5) |
| Uvilla | 703892 | S | Ecuador | (10) |
| Kennebec | 800098 | MR | USA | (26,32) |
| Ecopapa Shayari | ... | R | Ecuador | ... ^e |
| INIAP-Santa Catalina | 720075 | R | Ecuador | (10) |
| INIAP-Suprema | ... | R | Ecuador | (2) |
| Rosita | 720044 | R | Mexico | (21,22) |

^a 'INIAP-Cecilia' is referred to in the text as 'Cecilia'; 'INIAP-Gabriela' as 'Gabriela'; 'Ecopapa Shayari' as 'Ecopapa'; 'INIAP-Suprema' as 'Suprema'; and 'INIAP-Santa Catalina' as 'Catalina'. LBR37 is a genotype from the breeding program at the International Potato Center (CIP), Lima, Peru.

^b Identification given by CIP.

^c S, susceptible; MR, moderately resistant; and R, resistant.

^d Country where the cultivar was used.

^e M. V. Taipei (*unpublished data*).

Planting dates were between early November and mid-February in Ecuador, early January and early February in Israel, mid-June in Mexico, and mid- to late June in the United States. Application of fertilizers, insecticides, and herbicides as well as other cultural practices were done according to the recommendations of local investigators.

There were no applications of fungicides, except in four experiments carried out in Ecuador (Table 4, experiments A, C, F, and H), in which young plants were sprayed one to three times with fungicides against *P. infestans* in order to avoid infections immediately after emergence. The fungicides applied were cymoxanil (120 g ha⁻¹) and mancozeb (960 g ha⁻¹) as Curzate M8 (2.5 g liter⁻¹ at an application rate of approximately 600 liter ha⁻¹), or chlorothalonil (750 ml ha⁻¹) as Bravo 500 (2.5 g liter⁻¹ at the same application rate). In all experiments, except for those in the United States, fungicides were also applied on additional plots to estimate the growing period of a healthy crop. For the experiments in the United States, the growing period of a healthy crop was estimated from commercial plots adjacent to those of the experiment.

Late blight epidemics started from natural inoculum in most experiments, but in a few, the epidemics were initiated with artificial inoculations as described by Bruhn and Fry (7) and Doster et al. (12). Inoculations were done 61 days after emergence (DAE) in La Tola 1997, 28 DAE in Freeville 2001, and 44 DAE in Freeville 2002 (Table 4). In the United States, the isolates used in the experiments belonged to the US-11 and US-17 clonal lineages (Table 4). In Mexico, populations of *P. infestans* were genotypically diverse (20), and populations within a single plot consisted of multiple genotypes (N. J. Grünwald, *unpublished data*). Genotypes of *P. infestans* were not identified in Ecuador and Israel, although in Ecuador the experiments were located in areas where the EC-1 clonal lineage is predominant (15).

Severity of late blight (S, percent) was estimated visually on a plot basis every 3 to 14 days starting at 0 to 68 DAE following the method described by Fry (16). Epidemics were evaluated until

TABLE 4. Experiments used to qualify the LB2004 version of LATEBLIGHT

| Country ^a | Year | Location (experiment) ^b | n ^c | Cultivar (epidemic) ^d | T̄ ^e | H̄ ^f |
|----------------------|------|------------------------------------|----------------|---|-----------------|-----------------|
| Ecuador | 1997 | Cutuglhua (A) ^g | 3 | Bolona (1), Gabriela (2), Catalina (3) | 11.2 | 15.5 |
| | | La Tola (B) ^h | 3 | Bolona (4), Gabriela (5), Catalina (6) | 16.2 | 15.1 |
| | 1998 | Cutuglhua (C) ⁱ | 5 | Uvilla (7), Gabriela (8), Catalina (9) | 12.5 | 15.6 |
| | | La Tola (D) | 3 | Uvilla (10), Gabriela (11), Catalina (12) | 16.7 | 15.8 |
| | | Pinantura (E) | 4 | Uvilla (13), Gabriela (14), Catalina (15) | 9.5 | 17.4 |
| | 1999 | Cutuglhua (F) ^j | 4 | Cecilia (16), Ecopapa (17) | 10.6 | 18.8 |
| | | La Tola (G) | 3 | Cecilia (18), Ecopapa (19) | 15.1 | 15.2 |
| | 2000 | Amaguaña (H) ^k | 4 | LBR37 (20), Suprema (21) | 13.3 | 15.4 |
| Israel | 2000 | Mishmeret (I) | 5 | Nicola (22) | 16.8 | 14.3 |
| | | Nir-Eliyahu (J) | 4 | Nicola (23) | 16.6 | 15.9 |
| Mexico | 1998 | Toluca (K) | 4 | Alpha (24), Rosita (25) | 14.7 | 15.2 |
| | 1999 | Toluca (L) | 4 | Alpha (26), Rosita (27) | 14.3 | 13.5 |
| | 2000 | Toluca (M) | 4 | Alpha (28), Rosita (29) | 14.1 | 13.7 |
| USA | 2001 | Freeville (N) ^l | 3 | Katahdin (30), Kennebec (31) | 18.6 | 17.6 |
| | 2002 | Freeville (O) ^m | 3 | Katahdin (32) | 19.0 | 12.9 |

^a Experiments conducted in Mexico were previously reported by Grünwald et al. (21,22).

^b The letter in parenthesis identifies the experiment. The same code is used in Figure 2.

^c Number of repetitions.

^d The number in parenthesis identifies the epidemic. The same code is used in Figures 1 and 2.

^e Mean daily air temperature (°C).

^f Mean daily period (h) with relative humidity (RH) above 85 or 90%. The 85% RH threshold was used when the sensor was located above the potato canopy (experiments A to H) and the 90% threshold when it was located within (experiments I to O) (4).

^g Two applications of cymoxanil and mancozeb on young plants.

^h Plants were artificially inoculated with an isolate collected in an area where the EC-1 clonal lineage of *Phytophthora infestans* is predominant. Forty 20-μl drops of a sporangial suspension (1.5 × 10⁴ sporangia ml⁻¹) were applied to two plants located at opposite corners of each plot at 61 days after emergence (DAE). The resulting level of initial inoculum was estimated to be one lesion per 10 plants (0.23 lesions m⁻²).

ⁱ Three applications of cymoxanil and mancozeb on young plants.

^j Two applications of chlorothalonil on young plants.

^k Two applications of cymoxanil and mancozeb on young plants. Three repetitions on cv. Suprema.

^l Plants were artificially inoculated with isolate US990025 (US-11 clonal lineage) applied at 3,750 sporangia ml⁻¹ (10 ml) to the center of the 16.6-m² plot at 28 DAE. This isolate was obtained from the Cornell University *P. infestans* culture collection. Two repetitions on cv. Katahdin.

^m Plants were artificially inoculated with isolate US970001 (US-17 clonal lineage) applied at 2,000 sporangia ml⁻¹ (12 ml) to the center of the 16.6-m² plot at 44 DAE. This isolate was obtained from the Cornell University *P. infestans* culture collection.

healthy plants from plots with chemical control started to senesce (Ecuador and Israel), or until plants were killed either by application of a herbicide (Mexico) or by plowing the field (United States).

Temperature and relative humidity (RH) were measured using sensors connected to automated data loggers. The temperature/RH sensors were located above the potato canopy at approximately 1.20 m in experiments from Ecuador, or within the potato canopy at 0.15 to 0.50 m in Israel, Mexico, and the United States. The following equipment was used: Hobo (models RH, H8 Temperature, and Pro RH/Temp, Onset Computer Corporation, Pocasset, MA); Delta-T with standard sensors (Delta-T Devices Ltd., Cambridge, UK); Phyttech (stand-alone data logger with air temperature and humidity gauge model LSD-ATH-5L, Phyttech Ltd., Yad Mordechai, Israel); and Campbell (model CR10X with 207-probe, Campbell Scientific, Logan, UT). The sensors were logged every 1 to 60 min; and daily values of mean air temperature (T, degree Celsius), duration of the period of high humidity (H, hour), and mean air temperature during high humidity periods (TH, degree Celsius) were calculated. The period of high humidity was defined as the number of hours when RH was above 85%, if the sensor was located above the potato canopy, or 90%, if the sensor was located within the canopy (4).

Analysis of field data. Ninety-five percent confidence intervals of the mean of the observed relative area under the disease progress curve (\overline{RAUDPC}_{obs} , percent) (17) were calculated. This variable was estimated by dividing the mean observed AUDPC (\overline{AUDPC}_{obs} , percent-days) (8) by the mean length of the epidemic, which was estimated as the difference (days) between the end of the growing period of a healthy crop and the first appearance of lesions (4). Within an experiment, cultivars with non-overlapping confidence intervals of \overline{RAUDPC}_{obs} were statistically different among each other at $\alpha = 0.05$. In some epidemics, there were no S evaluations at the end of the growing period. In those cases, it was assumed that S was evaluated at harvest, and that its value was equal to that obtained in the last evaluation ($S = 100\%$ in most cases).

Simulations. Simulated blight severity was obtained with LB2004 (3). Environmental variables measured in each field experiment (T, H, and TH) were used as driving variables for the model. Other parameters taken from field experiments included: emergence date, date of the last evaluation on which no lesions were observed (t_0), date of the first evaluation on which lesions were observed (t_1), blight severity on t_1 (S_{t_1}), date of the end of the experiment (defined as the last date when both weather and severity data were available), and maturity level of the cultivar (early, mid-, or late). Cultivars with fewer than 100 days from planting to maturity in a given location were considered as early-season cultivars, with 100 to 120 days as mid-season, and with more than 120 days as late-season (32,36). Initial lesions (IL, lesions per square meter of ground) and day of inoculation (DOI, DAE) were calculated from field observations using t_0 , t_1 , and S_{t_1} (Table 1) (4). In some experiments (Table 4, experiments A, D, I, and J), the first measurement of S was done when the epidemic had already started and, therefore, t_0 was unknown. In those cases, t_0 was estimated as $t_1 - 7$ days. In several epidemics in which the first S evaluations repeatedly had values less than 1% (Table 4, experiments C, E, F, G, and H), the model predictions improved by using IL and DOI that were calculated by assuming that $S = 0\%$ for all evaluations before the first evaluation with $S \geq 1\%$.

The resistance categories (susceptible, moderately resistant, and resistant) to which the cultivars were assigned by local investigators (Table 3) were critical for parameterization of the model. Parameters of LGR, SR, and LP, as determined in a previous study (3) for susceptible cv. Tomasa, moderately resistant cv. Yungay, and resistant cv. Amarilis, were used for cultivars with equivalent resistance categories in this study. In this way, parameters of cv. Tomasa were used for cultivars classified as suscep-

tible, parameters of cv. Yungay were used for cultivars classified as moderately resistant, and parameters of cv. Amarilis were used for cultivars classified as resistant. Clearly, this categorization of the cultivars is a rough simplification of their actual resistance levels. Thus, it was expected that the model might over- or underestimate disease in cultivars in which their resistance levels were considerably different from those of the cultivars used for parameterization.

Model qualification. Graphical and statistical tests were used to compare disease progress curves and AUDPC values generated with LB2004 to those observed in the field experiments. The variable AUDPC, rather than RAUPDC, was used because observed and simulated epidemics had the same duration. The validity of predictions made by LB2004 with generalized parameters of LGR, SR, and LP was assessed using predefined subjective and objective performance criteria (4).

Graphical comparisons were made by plotting observed and simulated disease progress curves (blight severity versus time) to subjectively evaluate the goodness of fit (34) and to define whether LB2004 possesses sufficient accuracy to predict the effect of environmental conditions on late blight epidemics (4).

Simulated AUDPC values were analyzed statistically for how well they fit observed values using the confidence interval test (CIT) (1) on epidemics grouped by cultivar. The model predictions passed this test in a certain epidemic when the mean simulated AUDPC (\overline{AUDPC}_{sim}) fell within the 95% confidence interval of the mean \overline{AUDPC}_{obs} (\overline{AUDPC}_{obs}) (4). For each cultivar, the percentage of epidemics passing this test was referred to as CIT value. The performance criterion to consider model predictions acceptable for a cultivar was to obtain a CIT value of at least 75% (4,33). An equivalence test (19,37) used in a previous study (4) was not used here, because sample sizes were too small when epidemics were grouped by cultivar.

Deviations of AUDPC ($\overline{AUDPC}_{sim} - \overline{AUDPC}_{obs}$) were analyzed quantitatively and graphically using the envelope of acceptance test (EAT) (4,25,37) on epidemics grouped by cultivar. The predictions made by the model passed this test in a certain epidemic when the AUDPC deviation fell within an envelope of acceptance, which was calculated by multiplying \overline{AUDPC}_{sim} by a tolerance range of 0.15 based on the precision and accuracy of blight severity estimation in the field (4). For each cultivar, the percentage of epidemics passing this test was referred to as EAT value. The performance criterion to consider model predictions acceptable for a cultivar was to obtain an EAT value of at least 75% (4). For each cultivar, the envelope of acceptance and the AUDPC deviations were plotted on \overline{AUDPC}_{sim} in order to detect trends.

Estimating parameters by calibration. Since model predictions in some cultivars were consistently higher than observed epidemics (described below), it was hypothesized that these differences could be caused by erroneous estimates for LP, LGR, and/or SR. To test this hypothesis, the parameters for these fitness components in cv. Bolona were estimated by calibration. Bolona has been defined as susceptible by local investigators (Table 3) and, therefore, the parameters originally assigned to it were those obtained from susceptible cv. Tomasa (3). The parameters were decreased (increased for LP) until the deviations of AUDPC were minimized. This was done using weather data from one experiment (Table 4, experiment A). The limits for the parameter estimates were set according to the estimates obtained in resistant cv. Amarilis (3). Weather data from another experiment (Table 4, experiment B) were used for simulation, and not for calibration. Graphical comparisons between simulated and observed epidemics were made, and CIT and EAT values were calculated for both experiments.

Sensitivity analysis. Sensitivity of LB2004 to changes in parameters and input variables was evaluated through metamodelling (24,28). A metamodel is "an approximation of the input/output transformation that is implied by the simulation model"

(24). Parameters and input variables of the simulation model were considered to be factors, while an output variable of the simulation model was considered to be the response variable. Treatments, or scenarios, consisting of the combinations of different factor levels were fed into the simulation model. The scenarios and the model output were then analyzed to estimate the metamodel, which in turn was used to analyze the sensitivity of the simulation model to changes in parameters and input variables.

The parameters analyzed were those that define the host–pathogen interaction in LB2004: LP, LGR, SR, and infection efficiency (IE). The input variables analyzed were those related to initial inoculum (IL and DOI) and weather (\bar{T} and \bar{H}). Factor levels were selected according to the variability observed in this and other studies (described below), and defined the conditions for which the metamodel was considered valid. Three levels were selected for LP, LGR, SR, IE, IL, and DOI, and were coded as 0, 0.5, and 1 (Table 5). Level 0 represented the most favorable conditions for disease development (i.e., low values for LP and DOI, and high values for LGR, SR, IE, and IL); level 0.5 represented intermediate conditions for disease development; and level 1 represented the least favorable conditions for disease development (i.e., high values for LP and DOI, and low values for LGR, SR, IE, and IL). Maximum and minimum values for LP, LGR, and SR were experimentally obtained from Peruvian potato cultivars infected with isolates of the EC-1 clonal lineage (3). Intermediate values were calculated as the average of the maximum and minimum values. Values for IE were obtained from an earlier version of LATEBLIGHT (18). Values for IL and DOI were based on the variation observed in field experiments reported in this and a previous study (4).

The LB2004 version of the model was run for the resulting 729 (3⁶) scenarios. Weather data to run the model were obtained from 11 data sets from Ecuador, Mexico, and the United States (Table 4, all experiments except I, J, and N, which were excluded because of insufficient data; and E, which was excluded because model predictions were not acceptable, described below). Weather data to run the model were also obtained from four data sets from Peru (4): Comas (1999), Oxapampa (1999), Huancayo (2000), and Oxapampa (2000). The length of the epidemic was set to

50 DAE because of the availability of weather data. Values of \bar{T} and \bar{H} ($n = 50$ days) were calculated for each data set and were incorporated as factors in the analysis. Maximum values of \bar{T} (19.2°C in Freeville, United States, 2002) and \bar{H} (20.2 h in Comas, Peru, 1999) were coded as level 0, whereas minimum values for \bar{T} (11.0°C in Cutuglahua, Ecuador, 1999) and \bar{H} (10.6 h in Huancayo, Peru, 2000) were coded as level 1. Intermediate values were scaled accordingly. Therefore, 15 combinations of \bar{T} and \bar{H} were obtained, and the final number of scenarios available to construct the metamodel increased to 10,935 (729 × 15).

Standard AUDPC (sAUDPC, %) was used as the response variable for the metamodel. sAUDPC was calculated by dividing each AUDPC value obtained with LB2004 by the maximum AUDPC value obtained with each weather data set multiplied by 100.

The metamodel was estimated using ordinary least squares regression (27,28). First, assumptions for regression analysis were tested (27). Then, data were fit to a first-order polynomial metamodel with eight factors, after which nonsignificant factors ($P > 0.05$) were excluded. Next, quadratic effects and two-factor interactions were included in the metamodel. Finally, nonsignificant quadratic effects and interactions, as well as those that did not increase the adjusted coefficient of multiple determination (R^2) by at least 0.01, were eliminated. Only two-factor interactions were considered.

The validity of the metamodel was estimated through cross-validation (28). Scenarios were eliminated individually, and then parameters of the metamodel were re-estimated. The resulting metamodel was used to predict the simulation result for the eliminated scenario. Performance of the metamodel was estimated by calculating the Pearson product-moment correlation coefficient (27) between predictions made with the metamodel after the cross-validation and those obtained with LB2004. Analyses were done using SAS software, release 9.1 (Statistical Analysis System, SAS Institute, Cary, NC).

RESULTS

Field experiments. The observed late blight epidemics were highly variable, ranging from epidemics that reached 100% blight severity in a few days (Fig. 1, epidemic 30) to epidemics that never exceeded 5% blight severity (Fig. 1, epidemic 19). In epidemics that started from natural inoculum, the first symptoms of disease were observed on average at 40 DAE (standard deviation [SD] = 20) and ranged between 8 and 89 DAE. Initial blight severity averaged 0.5% (SD = 1) and varied from 0.001 to 8%. The mean period to reach 50% blight severity (in those epidemics that reached that level) was 68 DAE (SD = 25) and varied from 22 to 108 DAE. Final severity ranged between 2 and 100%.

Analysis of field data. Within-treatment variability of field data expressed as 95% confidence intervals for RAUDPC_{obs} was erratic (Fig. 2). There were experiments in which all cultivars had relatively low variability (Fig. 2, experiment M), while in others all cultivars had relatively high variability (Fig. 2, experiment E). In most experiments, however, within-treatment variability was dependent on the cultivar: RAUDPC_{obs} in some cultivars was highly variable, but not in others. Furthermore, the effect of cultivar was not constant across locations. For example, in experiment C (Fig. 2), cv. Gabriela (epidemic 8) showed relatively high variability and cv. Catalina (epidemic 9) showed relatively low variability, while in experiment D the opposite occurred. Five epidemics had wide confidence intervals, four in Ecuador (Fig. 2, epidemics 4, 6, 19, and 21) and one in Mexico (epidemic 24). Two of the erratic epidemics from Ecuador occurred in the experiment in which the pathogen was artificially inoculated (Fig. 2, experiment B).

Reactions of potato cultivars to *P. infestans* in the field (Fig. 2) generally agreed with those described by local investigators (Table 3). The previously reported susceptible cultivars (Alpha,

TABLE 5. Factors and levels used in the sensitivity analysis for the LB2004 version of LATEBLIGHT

| Factor ^a | Level ^b | | |
|---------------------|--------------------|------|------|
| | 0 | 0.5 | 1 |
| LP | 2.80 | 3.10 | 3.41 |
| LGR | 4.10 | 3.75 | 3.40 |
| SR | 2.92 | 2.07 | 1.22 |
| IE | 1.0 | 0.9 | 0.8 |
| IL | 500 | 50 | 5 |
| DOI | 1 | 11 | 21 |

^a Factors represent parameters (LP, LGR, SR, and IE) and input variables (IL and DOI) of LB2004. LP, latent period (days); LGR, lesion growth rate ($m\ day^{-1}$, $\times 10^{-3}$); SR, sporulation rate (sporangia $m^{-2}\ day^{-1}$, $\times 10^8$); IE, infection efficiency (proportion of sporangia and zoospores that infect the host over a maximum of 0.1 infections per germinated sporangium per day, and 0.01 infections per germinated zoospore per day [18]); IL, initial lesions (lesions m^{-2} of ground); and DOI, day of inoculation (days after emergence).

^b Level 0 represents the most favorable conditions for the disease, level 0.5 represents intermediate conditions for the disease, and level 1 represents the least favorable conditions for the disease. Minimum and maximum values for LP, LGR, and SR were obtained from Peruvian potato cultivars infected with isolates of the EC-1 clonal lineage of *Phytophthora infestans* (3). For IE, minimum and maximum values were obtained from an earlier version of LATEBLIGHT (18). Intermediate values for LP, LGR, SR, and IE were calculated as the average of minimum and maximum values. For IL and DOI, minimum, intermediate, and maximum values were obtained based on the variation observed in field experiments from this and a previous study (4).

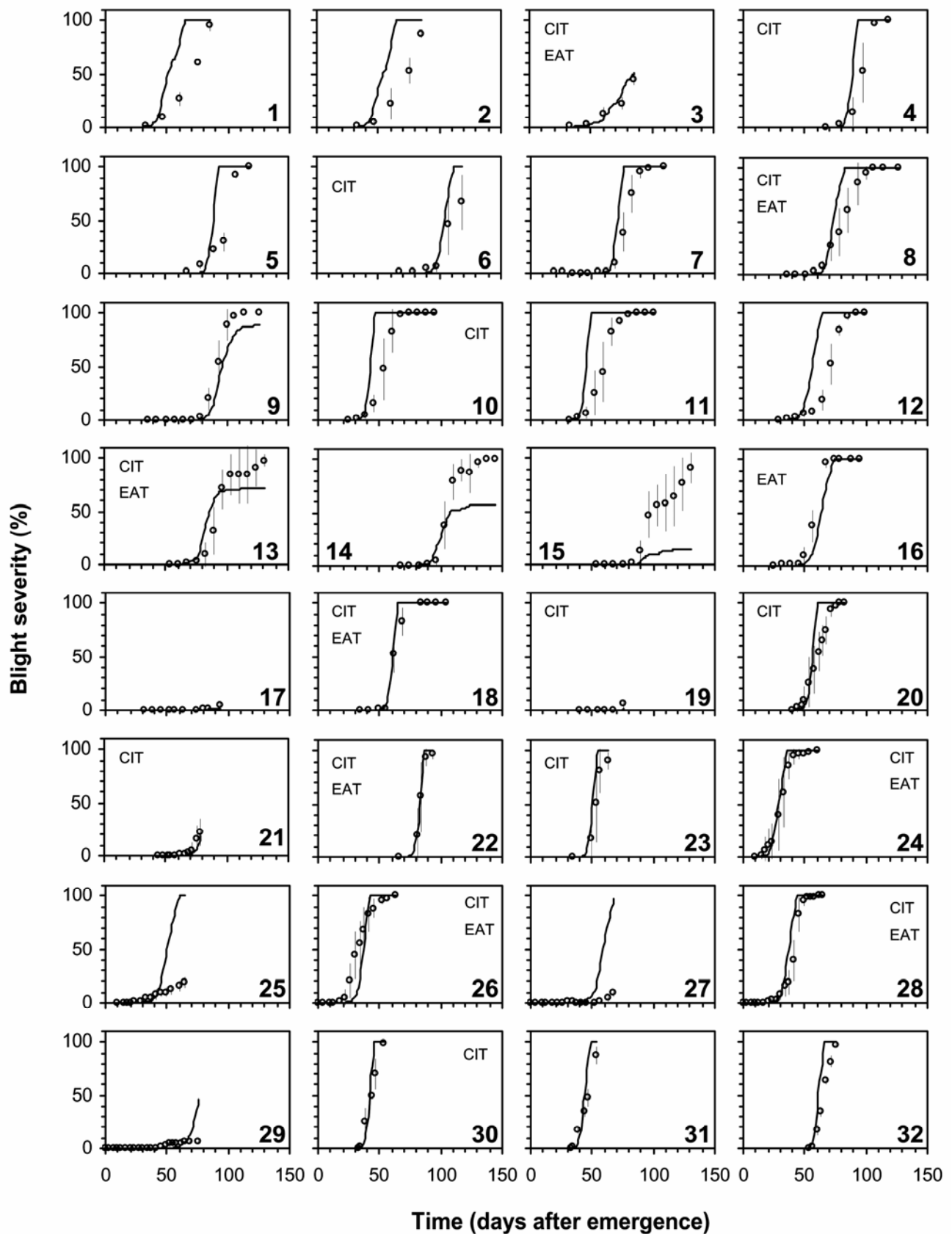


Fig. 1. Observed (circles) and simulated (continuous line) disease progress curves of 32 potato late blight epidemics under field conditions in Ecuador (epidemics 1 to 21, Table 4), Israel (epidemics 22 and 23), Mexico (epidemics 24 to 29), and the United States (epidemics 30 to 32). The simulated progress curves were obtained with the LB2004 version of LATEBLIGHT. Vertical lines represent the standard deviation of the observed mean blight severity. In epidemics coded as CIT, the model predictions passed the confidence interval test (CIT), i.e., the mean of the simulated area under the disease progress curve ($AUDPC_{sim}$) fell within the 95% confidence interval of the mean of the observed area under the disease progress curve ($AUDPC_{obs}$). In those coded as EAT, the model predictions passed the envelope of acceptance test (EAT), i.e., $AUDPC$ deviations ($AUDPC_{sim} - AUDPC_{obs}$) fell within the envelope of acceptance ($AUDPC_{sim} \times 0.15$).

Bolona, Cecilia, Gabriela, Katahdin, LBR37, Nicola, and Uvilla) had the highest levels of disease. There were no statistically significant differences ($\alpha = 0.05$) among susceptible cultivars within an experiment, with the exception of experiment C (Fig. 2) in which cv. Uvilla (epidemic 7) had less disease than cv. Gabriela (epidemic 8). The \overline{RAUDPC}_{obs} of susceptible cultivars averaged 66% (SD = 15) with a maximum of 98% in epidemic 32 and a minimum of 36% in epidemic 5 (Fig. 2). Moderately resistant cv. Kennebec (Fig. 2, epidemic 31) had slightly less disease than susceptible cv. Katahdin (epidemic 30), but the difference was not statistically significant ($\alpha = 0.05$). Resistant cvs. Ecopapa (Fig. 2, epidemics 17 and 19), Suprema (epidemic 21), and Rosita (epidemics 25, 27, and 29) had low to very low levels of disease and, within experiments, the differences from the susceptible cultivars were significant ($\alpha = 0.05$). Catalina (Fig. 2, epidemics 3, 6, 9, 12, and 15), reported to be resistant by local investigators (Table 3), had generally less disease than the susceptible cultivars, but the differences within experiments were statistically significant only in one experiment (Fig. 2, experiment A). The \overline{RAUDPC}_{obs} in the resistant cultivars (excluding cv. Catalina) averaged 6% (SD = 4) with a maximum of 13% in epidemic 21 and a minimum of 3% in epidemic 29 (Fig. 2).

Model qualification. Using subjective and objective performance criteria, the results were grouped into three categories. In the first category, the relationship between model prediction and observation was consistent and some or all of the predefined performance criteria were satisfied. For cv. Alpha in Mexico, the simulated disease progress curves fit the observed data (Fig. 1, epidemics 24, 26, and 28); the CIT and EAT values were higher than the predefined performance criteria (Table 6); and the AUDPC deviations scattered randomly around zero and within the envelope of acceptance (Fig. 3A). For cvs. Cecilia, LBR37, and Suprema in Ecuador, and cv. Nicola in Israel, simulated progress curves were close enough to the observed data to allow useful inferences (Fig. 1, epidemics 16, 18, 20, 21, 22, and 23); either the CIT or the EAT values were higher than the predefined performance criteria (Table 6); and the AUDPC deviations scattered randomly around zero and within or close to the envelope of acceptance (Fig. 3B). Finally, for cv. Ecopapa in Ecuador and for cvs. Katahdin and Kennebec in the United States, the simulated progress curves, when evaluated subjectively for goodness of fit, adequately fit the observed data (Fig. 1, epidemics 17, 19, 30, 31, and 32). However, neither the CIT nor the EAT values were

higher than the predefined performance criteria (Table 6), and the AUDPC deviations showed a consistent trend of underestimation (cv. Ecopapa) or overestimation (cvs. Katahdin and Kennebec) of disease (Fig. 3C).

In the second category, the model predictions satisfied neither the subjective nor the objective criteria (with one exception), but the AUDPC deviations showed again a consistent trend of overestimation (Fig. 3D). This category included cv. Bolona in Ecuador (Fig. 1, epidemics 1 and 4) and cv. Rosita in Mexico (Fig. 1, epidemics 25, 27, and 29) (Table 6). The exception was cv. Bolona in La Tola (1997), in which the model predictions passed the CIT criterion because of the high variability in the observed epidemic (Figs. 1 and 2, epidemic 4).

In the third category, there was no meaningful relationship between model predictions and observed data. In some locations, and depending on a particular cultivar, the model predictions passed the subjective and objective performance criteria, while in other locations, the predictions were significantly higher or lower than the observed epidemics. Thus, AUDPC deviations scattered around zero showing no clear trend (Fig. 3E and F). Three cultivars evaluated in Ecuador fit in this category: Gabriela (Fig. 1, epidemics 2, 5, 8, 11, and 14), Catalina (Fig. 1, epidemics 3, 6, 9, 12, and 15), and Uvilla (Fig. 1, epidemics 7, 10, and 13). The relationship between model predictions and observations was particularly poor in the location of Pinantura with cv. Gabriela (Fig. 1, epidemic 14) and cv. Catalina (Fig. 1, epidemic 15).

For those cultivars belonging to the first category, the model accurately reflected the effect of weather on epidemic development. For example, epidemics in cv. Cecilia progressed slower in Cutuglahua ($\bar{T} = 10.6^{\circ}\text{C}$ and $\bar{H} = 18.8$ h; Fig. 1, epidemic 16) than in La Tola ($\bar{T} = 15.1^{\circ}\text{C}$ and $\bar{H} = 15.2$ h; Fig. 1, epidemic 18), and that was reasonably well predicted by LB2004.

Estimating parameters by calibration. Because the consistent overestimation of disease for some cultivars could have been caused by inaccurate estimates of the parameters for LP, LGR, and SR, other parameter estimates were obtained through calibration, and new simulations were run on independent data. For cv. Bolona, the estimates originally assigned (from cv. Tomasa [3]) were 2.82 days for LP, 4.10×10^{-3} m day⁻¹ for LGR, and

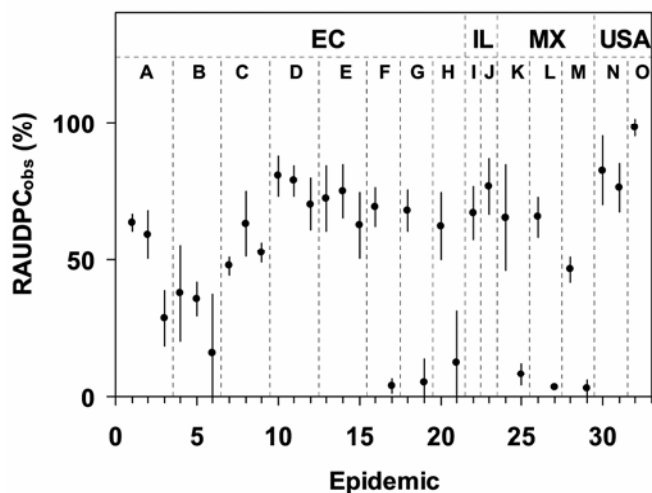


Fig. 2. Within-treatment variability of the observed relative area under the disease progress curve (\overline{RAUDPC}_{obs}) of 32 potato late blight epidemics (epidemics 1 to 32, Table 4) in 15 experiments (experiments A to O, Table 4) conducted in Ecuador (EC), Israel (IL), Mexico (MX), and the United States under field conditions. Circles represent \overline{RAUDPC}_{obs} and vertical lines indicate their 95% confidence intervals.

TABLE 6. Values of the confidence interval test (CIT) and the envelope of acceptance test (EAT) obtained with the LB2004 version of LATEBLIGHT for epidemics of late blight in 13 potato cultivars from Ecuador, Israel, Mexico, and the United States

| Cultivar | n^a | CIT (%) ^b | EAT (%) ^c |
|----------|-------|----------------------|----------------------|
| Alpha | 3 | 100 | 100 |
| Cecilia | 2 | 50 | 100 |
| LBR37 | 1 | 100 | 0 |
| Suprema | 1 | 100 | 0 |
| Nicola | 2 | 100 | 50 |
| Ecopapa | 2 | 50 | 0 |
| Katahdin | 2 | 50 | 0 |
| Kennebec | 1 | 0 | 0 |
| Bolona | 2 | 50 | 0 |
| Rosita | 3 | 0 | 0 |
| Catalina | 5 | 40 | 20 |
| Uvilla | 3 | 67 | 33 |
| Gabriela | 5 | 20 | 20 |

^a Number of epidemics.

^b Values of CIT represent the percentage of epidemics in which the mean of the simulated area under the disease progress curve (\overline{AUDPC}_{sim}) fell within the 95% confidence interval of the mean of the observed AUDPC (\overline{AUDPC}_{obs}). The performance criterion to consider the model valid was that \overline{AUDPC}_{sim} had to fall within the 95% confidence interval of \overline{AUDPC}_{obs} in at least 75% of the epidemics (4,33).

^c Values of EAT represent the percentage of epidemics in which AUDPC deviations ($\overline{AUDPC}_{sim} - \overline{AUDPC}_{obs}$) fell within an envelope of acceptance ($\overline{AUDPC}_{sim} \times 0.15$). The performance criterion to consider the model valid was that AUDPC deviations had to fall within the envelope of acceptance in at least 75% of the epidemics (4).

2.92×10^8 sporangia $m^{-2} day^{-1}$ for SR, while the parameters obtained by calibration were 3.41 days for LP, 3.40×10^{-3} m day^{-1} for LGR, and 1.67×10^8 sporangia $m^{-2} day^{-1}$ for SR. Using these new parameters, the simulated progress curves fit the observed data (Fig. 4A and B), and the model predictions passed both the CIT and EAT criteria when using weather data from Cutuglahua, 1997 (Table 4, experiment A), and La Tola, 1997 (Table 4, experiment B). Similar results were obtained when other combinations of calibrated estimates were used.

Sensitivity analysis. All main effects of the factors analyzed were highly significant ($P < 0.01$), and the adjusted R^2 for the first-order polynomial metamodel was 0.909. The addition of quadratic effects increased the adjusted R^2 to 0.927, but this was caused mainly by the quadratic effect of \bar{T} , which increased it to 0.923. The effects of several two-factor interactions were significant ($P < 0.05$) but none of those were biologically meaningful, nor did they increase the adjusted R^2 by more than 0.010. Final adjusted R^2 of the metamodel was 0.923. The Pearson product-moment correlation coefficient between predictions made with the metamodel after cross-validation and those obtained with LB2004 was 0.961.

The parameter estimates in Table 7 represented the effect on sAUDPC when changing a factor from 0 to 1 (i.e., from very favorable conditions to less favorable conditions for the disease), while maintaining the remaining factors constant. For example, by increasing LP from 2.80 to 3.41 days, sAUDPC decreased linearly by 9.7%, assuming that all the other factors did not change. In the case of \bar{T} , sAUDPC decreased quadratically by 37.3% when \bar{T} changed from 19.2 to 11.0°C. The factors DOI, \bar{T} , and IL had the largest effects on sAUDPC (reduction $\geq 29.0\%$); \bar{H} and SR had a moderate impact on sAUDPC (reduction between 15.1 and 18.7%); and LP, LGR, and IE had a low impact on sAUDPC (reduction $< 9.7\%$).

DISCUSSION

Model qualification is a relatively new term in modeling. It was proposed by Rykiel (33) and, as described earlier, it amounts to revalidating a model for new cases, allowing the identification of new conditions for which it can be used and, therefore, increasing the domain of applicability of the model. Thus, qualification is a critical step when a validated model is intended to be used in a

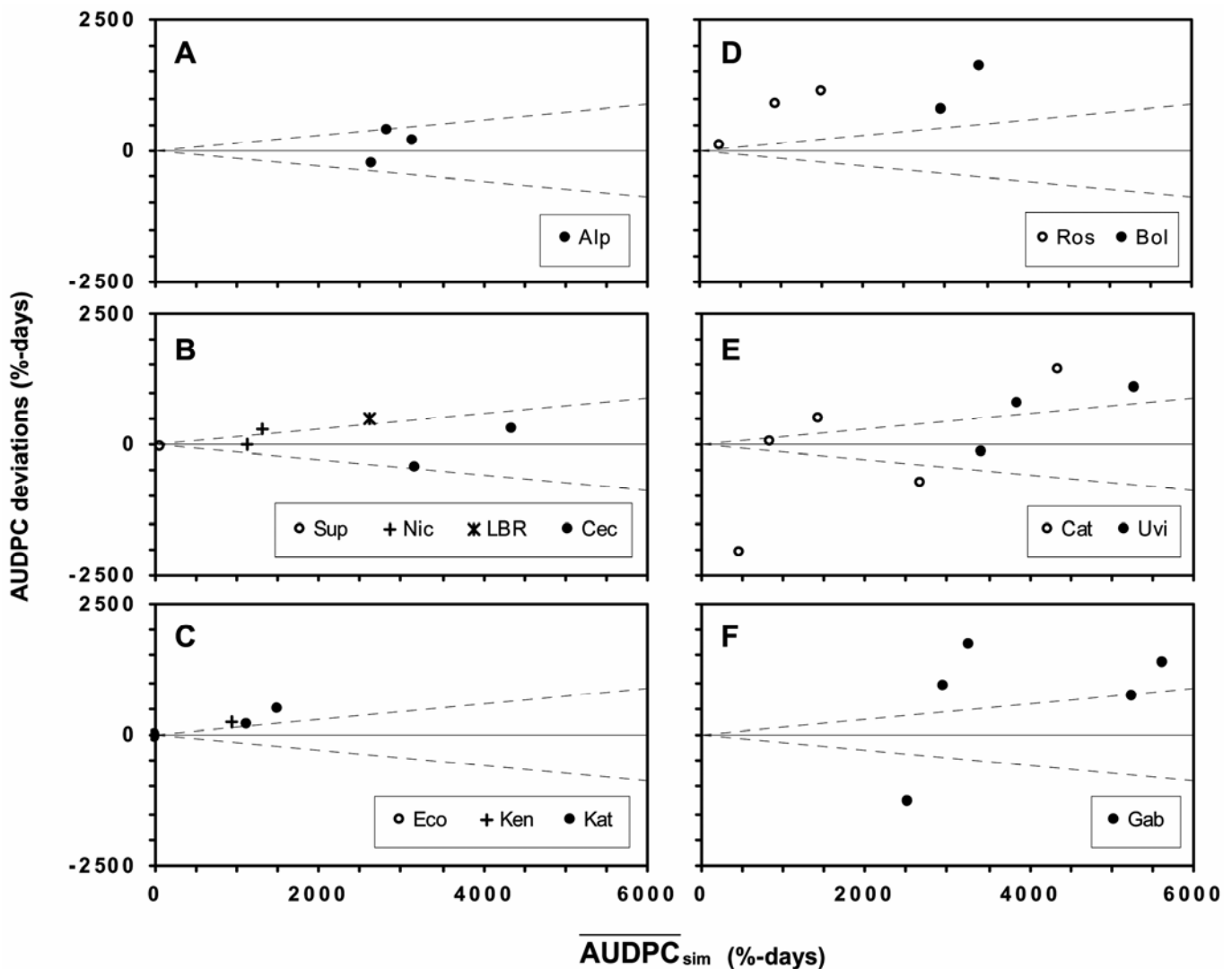


Fig. 3. Envelope of acceptance test for predictions made with the LB2004 version of LATEBLIGHT in epidemics in 13 potato cultivars from Ecuador, Israel, Mexico, and the United States. Deviations of area under the disease progress curve (AUDPC) (difference between the mean simulated AUDPC [$AUDPC_{sim}$] and the mean observed AUDPC [$AUDPC_{obs}$]) and an envelope of acceptance ($AUDPC_{sim} \times 0.15$, dotted lines) were plotted on $AUDPC_{sim}$. Each data point represents a late blight epidemic on a potato cultivar in a certain location. The model predictions were considered accurate when the AUDPC deviations fell within the envelope of acceptance. The graphs represent different categories of model predictions, from highly accurate (A) to erratic (F). Alp = 'Alpha'; Sup = 'Suprema'; Nic = 'Nicola'; LBR = 'LBR37'; Cec = 'Cecilia'; Eco = 'Ecopapa'; Ken = 'Kennebec'; Kat = 'Katahdin'; Ros = 'Rosita'; Bol = 'Bolona'; Cat = 'Catalina'; Uvi = 'Uvilla'; and Gab = 'Gabriela'. In most cases, epidemics started from natural inoculum.

range of conditions that might be outside the domain for which it was originally validated. In addition, model credibility depends on validation and qualification tests (33). There is another definition for model qualification (“determination of adequacy of the conceptual model to provide an acceptable level of agreement for the domain of intended application” [35]), which was not used in this study.

We illustrated the qualification concept with the LB2004 version of the LATEBLIGHT model. LB2004 is intended to be used for the study of potato late blight in tropical highlands, where there is large variation among locations in weather, cultivars, and socio-economic conditions (9,29). Since LB2004 was validated for a narrow set of Andean conditions (4), qualification was crucial to determine whether it can be used in a broader range of conditions and to build model credibility in the user community. To the best of our knowledge, this is one of the few examples in the field of plant pathology in which a plant disease model is tested under such a range of host, pathogen, and weather conditions. Another model that has been validated under a wide range of conditions is RICEPEST (37). This model was validated with epidemics from the Philippines, China, and India.

Since qualification is to revalidate the model for new cases, the same tests used in the validation of LB2004 (4) were used in this study. However, the CIT and EAT criteria showed some limitations that were not detected before. In some epidemics, model predictions were not accurate from a subjective point of view (Fig. 1, epidemic 4), but they passed the CIT criterion, because the observed variability in blight severity was high and, therefore, the confidence interval of \overline{AUDPC}_{obs} was wide (Fig. 2, epidemic

4). In other cases, model predictions were accurate from a subjective point of view (Fig. 1, epidemic 17), but they did not pass the CIT criterion, because of the low observed variability in blight severity that caused a narrow confidence interval of \overline{AUDPC}_{obs} . The width of the confidence interval of \overline{AUDPC}_{obs} can be manipulated by changing the number of repetitions in the field experiments: for fewer repetitions, there are wider confidence intervals and vice versa. This represents a limitation of the CIT criterion. We therefore recommend defining in advance the number of repetitions in field experiments for model validation or qualification. From our experience, it seems that four repetitions are adequate for those purposes.

In other epidemics, model predictions were accurate from a subjective point of view (Fig. 1, epidemics 19 and 21), but they did not pass the EAT criterion. This was due to the small value of \overline{AUDPC}_{sim} that generated an even lower value for the envelope of acceptance. The envelope of acceptance is defined by multiplying \overline{AUDPC}_{sim} by a tolerance range of ± 0.15 that was defined based on the precision and accuracy of blight severity estimation in the field (4). Therefore, in epidemics with low values of \overline{AUDPC}_{sim} (2 and 69 percent-days for epidemics 19 and 21, respectively), EAT might not be a reliable criterion to test the performance of a model.

It is necessary to clearly define how the epidemics are pooled for the CIT and EAT criteria. In our case, we decided to pool the epidemics by cultivar, because if all the epidemics were pooled, the results would be misleading. For example, the mean CIT value would be 50% and the mean EAT value would be 28%. Thus, the conclusion would be that the predictions made by LB2004 are not accurate, because the performance criterion in both cases was 75%. We consider that pooling all epidemics is not appropriate, because model performance varied widely, and this variation was mainly caused by inaccurate estimates of cultivar resistance. Generic classification of cultivar resistance provided by local investigators was used to run the model, but such estimates are only first approximations. Fortunately, these estimates were apparently accurate for eight cultivars (Alpha, Cecilia, LBR37, Nicola, Katahdin, Kennebec, Ecopapa, and Suprema) as model predictions for these cultivars were acceptable.

Inaccurate fitness parameters seem to be the most likely cause of consistent over- or underestimation of disease by LB2004. For example, the resistance levels of cvs. Rosita and Bolona were apparently higher than those of cvs. Amarilis and Tomasa, respectively, from which fitness parameters for the simulations were obtained. Therefore, model predictions overestimated disease in cvs. Rosita and Bolona. We confirmed that the overestimation in cv. Bolona was caused by inaccurate parameter settings by estimating new parameters through calibration in one location (Cutuglahua, 1997; $\bar{T} = 11.2^{\circ}\text{C}$ and $\bar{H} = 15.5$ h) and independ-

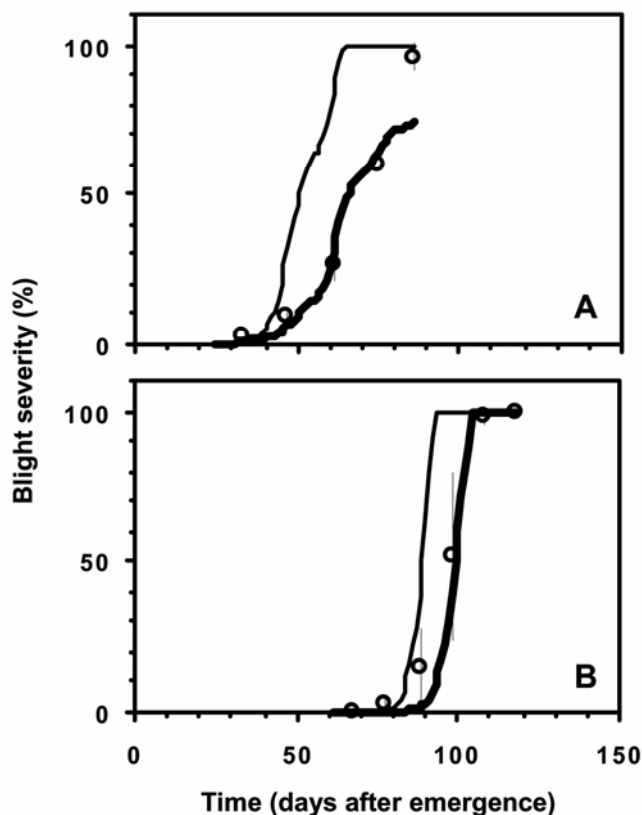


Fig. 4. Observed (circles) and simulated disease progress curves of potato cv. Bolona obtained with the LB2004 version of LATEBLIGHT using parameters for latent period (LP), lesion growth rate (LGR), and sporulation rate (SR) estimated experimentally from the interaction between potato cv. Tomasa and isolates of the EC-1 clonal lineage (thin continuous line), or using parameters for LP, LGR, and SR estimated by calibration (thick continuous line) in **A**, Cutuglahua (Ecuador, 1997) and applied in **B**, La Tola (Ecuador, 1997). Vertical lines represent the standard deviation of the observed mean blight severity.

TABLE 7. Significant factor effects included in the metamodel used in the sensitivity analysis for the LB2004 version of LATEBLIGHT

| Factor ^a | Parameter estimate | Standard error | $P > t $ ^b |
|---------------------|--------------------|----------------|------------------------|
| Intercept | 121.2 | 0.38 | <0.01 |
| LP | -9.7 | 0.18 | <0.01 |
| LGR | -7.2 | 0.18 | <0.01 |
| SR | -15.1 | 0.18 | <0.01 |
| IE | -3.8 | 0.18 | <0.01 |
| IL | -29.0 | 0.18 | <0.01 |
| DOI | -43.8 | 0.18 | <0.01 |
| \bar{H} | -18.7 | 0.31 | <0.01 |
| \bar{T}^2 | -37.3 | 0.87 | <0.01 |

^a LP, latent period; LGR, lesion growth rate; SR, sporulation rate; IE, infection efficiency; IL, initial lesions; DOI, day of inoculation; \bar{T} , mean air temperature; and \bar{H} , mean period with relative humidity above 85 or 90%. The response variable for the metamodel was the standard area under the disease progress curve (sAUDPC, %).

^b Values of P for partial t tests (27). $H_0, \beta_k = 0$ and $H_a, \beta_k \neq 0$.

ently validating them in another location (La Tola, 1997; \bar{T} = 16.2°C and \bar{H} = 15.1 h).

The poor relationship between model predictions and observations in the location of Pinantura (Ecuador, 1998) with cv. Gabriela (Fig. 1, epidemic 14) and cv. Catalina (Fig. 1, epidemic 15) apparently was caused by unidentified factors that affect epidemic development, but are not considered in the model. These unidentified factors appeared to interact with the host in complex ways, because model performance was poor for some cultivars (Gabriela and Catalina) but acceptable for others (cv. Uvilla, Fig. 1, epidemic 13). Pinantura is a high-altitude location (3,543 m above sea level), and during the experiment, \bar{T} was 9.5°C and \bar{H} was 17.4 h. It seems unlikely that \bar{T} and \bar{H} alone explain the poor relationship between prediction and observation in Pinantura, because model predictions were acceptable at Cutuglahua (Ecuador, 1999) under weather conditions similar to those of Pinantura (\bar{T} = 10.6°C and \bar{H} = 18.8 h). Further field and growth chamber experiments are necessary, first, to confirm the results obtained in this location and, second, to determine which factors are responsible for poor model performance.

The cause of inaccurate model predictions in cvs. Catalina, Gabriela, and Uvilla could be due to any of several factors. In the case of cvs. Gabriela and Uvilla, the model overestimated the disease in all the locations (Fig. 1, epidemics 2, 5, 8, and 11 for cv. Gabriela; and epidemics 7 and 10 for cv. Uvilla), except in Pinantura (Fig. 1, epidemic 14 for cv. Gabriela, and epidemic 13 for cv. Uvilla). Therefore, if we exclude the location of Pinantura because of the reasons explained above, the inaccurate model predictions for cvs. Gabriela and Uvilla may be explained by the use of inaccurate fitness parameters, similar to the case of cvs. Rosita and Bolona. In the case of cv. Catalina, excluding the location of Pinantura, we found that model predictions were accurate in one experiment (Fig. 1, epidemic 3), while in three other experiments model predictions either over- (Fig. 1, epidemics 6 and 12) or underestimated disease (Fig. 1, epidemic 9). The erratic model predictions could be pointing out that the relatively low level of resistance of cv. Catalina (Fig. 2, epidemics 3, 6, 9, 12, and 15) is unstable across sites. However, this cultivar has quantitative resistance against *P. infestans* (10), and there is little evidence for phenotypic instability in this type of resistance (14). Fungicide applications on certain experiments immediately after plant emergence (Table 4, experiments A, C, F, and H) could also have contributed to inaccurate model predictions by delaying the observed epidemics, especially in the susceptible cultivars (Fig. 1, epidemics 1, 2, and 3).

The sensitivity analysis showed that, within the range of values observed empirically, LB2004 is more sensitive to changes in variables related to initial inoculum (DOI and IL) and to weather (\bar{T} and \bar{H}) than to changes in the fitness parameters that define the host-pathogen interaction in the model (LP, LGR, SR, and IE). Among the fitness parameters, LB2004 was more sensitive to changes in SR, followed by LP, LGR, and IE. However, these conclusions are strongly influenced by the levels of each factor. Factors that have a wide range of observed values will appear more important in the analysis than factors that have a narrow range of observed values (28). This is a weakness of sensitivity analysis and, therefore, selection of factor levels should be done carefully because they define the domain of applicability of the metamodel. The criterion that we followed for selection of factor levels was to consider the variability in the real system observed in this and other studies (3,4). Thus, our conclusions are valid only within the limits of the factor levels. Importantly, the estimates for IE were obtained from an earlier study (18) in which the values were obtained by calibration. Therefore, the inferences regarding IE should be taken cautiously.

The strong sensitivity of the model to the weather and initial inoculum parameters relative to the fitness components may also explain why our rough classification of cultivars in resistance

groups was generally successful. Nevertheless, it would have been preferable to measure the fitness components of each cultivar, but this was not feasible in the current study. The sensitivity analysis also provided insight into the design of improved methods for management of potato late blight. Under the conditions in which our metamodel is valid, the disease is predicted to be best controlled by delaying the appearance of initial inoculum (defined as DOI in the metamodel) and decreasing its level (defined as IL), or by reducing the rates of disease increase by making the environment less favorable to the disease (i.e., reducing \bar{T} and \bar{H}), and to a lesser degree, by increasing the resistance level of the plant. This insight can help explain patterns in regional distribution of potato production. For example, farmers can shift production to different environments to escape late blight, e.g., through early planting (later DOI and lower IL), growing the crop partly outside of the rainy season (lower \bar{H}), and planting in higher locations (lower \bar{T}) (11,13,23,30). However, in many circumstances, such options are not available to farmers and, apart from fungicides, host resistance is the main component in disease management. To further explore the relative contributions of LP, LGR, SR, and IE in host resistance, our metamodeling approach could be used. However, factor levels for DOI, IL, \bar{T} , and \bar{H} should be selected based on the variability observed in the specific region where the inferences are to be applied, and factor levels for LP, LGR, SR, and IE should be selected based on the available genetic variation.

The LB2004 version of LATEBLIGHT accurately predicted the effect of weather on epidemics of potato late blight under a wide range of environmental conditions, potato cultivars, and *P. infestans* populations. This conclusion was based on the fact that the level of agreement between model output and reality was acceptable according to predefined objective and subjective performance criteria. We suggest that the formal domain of applicability of LB2004 can be extended from the range of field conditions for which it has been validated (three locations in Peru [4]) to the range of conditions observed in the field experiments of this study, with the exception of the experiment in Pinantura (Ecuador). With respect to \bar{T} , we propose that the domain of applicability can be extended from a range of 12.0 to 16.5°C to a range of 10.6 to 19.0°C. Our experiments did not enable us to extend the domain of applicability of \bar{H} ; it remains as previously defined (4), between 10.6 and 19.5 h. We also conclude that LB2004 can be used both as an educational tool, to demonstrate epidemiological principles to diverse audiences, and as a research tool, to generate hypotheses about disease management. However, two key submodels still need to be improved. The first is the fungicide submodel, which will allow the design and evaluation of improved strategies for chemical control, and the second is the host growth submodel, which will allow estimation of yield losses due to late blight and yield gains due to disease control.

ACKNOWLEDGMENTS

This study was supported by US-AID Project "Computer Simulation Models: A Tool for Transferring IPM Technologies from Developed to Developing Countries" (C15-073-TA-MOU-99). We thank the following persons for technical assistance in the field experiments: O. Bastidas, C. Cañizares, R. Jaramillo, J. Molineros, and J. Montalvo from the International Potato Center (CIP) in Quito, Ecuador; scientists from the Instituto Nacional de Meteorología e Hidrología in Quito, Ecuador; and J. Cianchetti and H. Mayton from Cornell University.

LITERATURE CITED

1. Alewell, C., and Manderscheid, B. 1998. Use of objective criteria for the assessment of biogeochemical ecosystem models. *Ecol. Model.* 107:213-224.
2. Andrade, H., Montesdeoca, F., and Cuesta, X. 2000. Información técnica de la variedad de papa INIAP-Suprema. Instituto Nacional Autónomo de Investigaciones Agropecuarias (INIAP), Quito, Ecuador.

3. Andrade-Piedra, J. L., Hijmans, R. J., Forbes, G. A., Fry, W. E., and Nelson, R. J. 2005. Simulation of potato late blight in the Andes. I: Modification and parameterization of the LATEBLIGHT model. *Phytopathology* 95:1191-1199.
4. Andrade-Piedra, J. L., Hijmans, R. J., Juárez, H. S., Forbes, G. A., Shtienberg, D., and Fry, W. E. 2005. Simulation of potato late blight in the Andes. II: Validation of the LATEBLIGHT model. *Phytopathology* 95:1200-1208.
5. Anonymous. 2002. Potato variety handbook. National Institute of Agricultural Botany, British Potato Council, Cambridge, UK.
6. Bruhn, J. A., Bruck, R. I., Fry, W. E., Arneson, P. A., and Keokosky, E. V. 1980. User's manual for LATEBLIGHT: A plant disease management game. Cornell Univ. Dept. Plant Pathol. Mimeo. 80-1.
7. Bruhn, J. A., and Fry, W. E. 1981. Analysis of potato late blight epidemiology by simulation modeling. *Phytopathology* 71:612-616.
8. Campbell, C. L., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley & Sons, New York.
9. Crissman, C. C., Espinosa, P., Ducrot, C. E. H., Cole, D. C., and Carpio, F. 1998. The case study site: Physical, health, and potato farming systems in Carchi province. Pages 85-120 in: Economic, Environmental, and Health Tradeoffs in Agriculture: Pesticides and the Sustainability of Andean Potato Production. C. C. Crissman, J. M. Antle, and S. M. Capalbo, eds. Kluwer Academic Publishers and International Potato Center (CIP), Dordrecht, The Netherlands.
10. Cuesta, X., Andrade, H., Bastidas, O., Quevedo, R., and Sherwood, S. 2002. Botánica y mejoramiento genético. Pages 33-50 in: El Cultivo de la Papa en Ecuador. M. Pumisacho and S. Sherwood, eds. Instituto Nacional Autónomo de Investigaciones Agropecuarias (INIAP) y Centro Internacional de la Papa (CIP), Quito, Ecuador.
11. Devaux, A., and Haverkort, A. J. 1987. The effects of shifting planting dates and mulching on late blight (*Phytophthora infestans*) and drought stress of potato crops grown under tropical highland conditions. *Exp. Agric.* 23:325-333.
12. Doster, M. A., Milgroom, M. G., and Fry, W. E. 1990. Quantification of factors influencing potato late blight suppression and selection for metalaxyl resistance in *Phytophthora infestans*: A simulation approach. *Phytopathology* 80:1190-1198.
13. Fernández-Northcote, E. N., Navia, O., and Gandarillas, A. 2000. Basis of strategies for chemical control of potato late blight developed by PROINPA in Bolivia. *Fitopatología* 35:137-149.
14. Forbes, G. A. 1999. Genotype by environment reaction to the late blight pathogen. Pages 57-66 in: Impact on a Changing World—Program Report 1997-98. International Potato Center (CIP), Lima, Peru.
15. Forbes, G. A., Escobar, X. C., Ayala, C. C., Revelo, J., Ordoñez, M. E., Fry, B. A., Doucett, K., and Fry, W. E. 1997. Population genetic structure of *Phytophthora infestans* in Ecuador. *Phytopathology* 87:375-380.
16. Fry, W. E. 1977. Integrated control of potato late blight: Effects of polygenic resistance and techniques of timing fungicide applications. *Phytopathology* 67:415-420.
17. Fry, W. E. 1978. Quantification of general resistance of potato cultivars and fungicide effects for integrated control of potato late blight. *Phytopathology* 68:1650-1655.
18. Fry, W. E., Milgroom, M. G., Doster, M. A., Bruhn, J. A., and Bruck, R. I. 1991. LATEBLIGHT: A Plant Disease Management Game—User's Manual. Version 3.1. Microsoft Windows adaptation by B. E. Ticknor and P. A. Arneson. Cornell Univ. Dept. Plant Pathol., New York.
19. Garrett, K. A. 1997. Use of statistical tests of equivalence (bioequivalence tests) in plant pathology. *Phytopathology* 87:372-374.
20. Grünwald, N. J., Flier, W. G., Sturbaum, A. K., Garay-Serrano, E., van den Bosch, T. B. M., Smart, C. D., Matuszak, J. M., Lozoya-Saldaña, H., Turkensteen, L. J., and Fry, W. E. 2001. Population structure of *Phytophthora infestans* in the Toluca valley region of central Mexico. *Phytopathology* 91:882-890.
21. Grünwald, N. J., Romero Montes, G., Lozoya Saldaña, H., Rubio Covarrubias, O. A., and Fry, W. E. 2002. Potato late blight management in the Toluca Valley: Field validation of SimCast modified for cultivars with high field resistance. *Plant Dis.* 86:1163-1168.
22. Grünwald, N. J., Rubio-Covarrubias, O. A., and Fry, W. E. 2000. Potato late-blight management in the Toluca Valley: Forecasts and resistant cultivars. *Plant Dis.* 84:410-416.
23. Haverkort, A. J., and Bicamumpaka, M. 1986. Correlation between intercepted radiation and yield of potato crops infested by *Phytophthora infestans* in central Africa. *Neth. J. Plant Pathol.* 92:239-247.
24. Kleijnen, J. P. C., and Sargent, R. G. 2000. A methodology for fitting and validating metamodels in simulation. *Eur. J. Oper. Res.* 120:14-29.
25. Mitchell, P. L. 1997. Misuse of regression for empirical validation of models. *Agric. Syst.* 54:313-326.
26. Murphy, H. J. 1986. Potato cultivars for the Northeastern United States. Pages 13-15 in: Potato Production in the Northeast: A Guide to Integrated Pest Management. C. S. Hollingsworth, D. N. Ferro, and W. M. Coli, eds. Cooperative Extension Service, University of Massachusetts, Amherst, MA.
27. Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. 1996. Applied Linear Statistical Models. 4th ed. WCB McGraw-Hill, Boston, MA.
28. Noordegraaf, A. V., Nielen, M., and Kleijnen, J. P. C. 2003. Sensitivity analysis by experimental design and metamodelling: Case study on simulation in national animal disease control. *Eur. J. Oper. Res.* 146:433-443.
29. Ortiz, O., Garrett, K. A., Heath, J. J., Orrego, R., and Nelson, R. J. 2004. Management of potato late blight in the Peruvian highlands: Evaluating the benefits of farmer field schools and farmer participatory research. *Plant Dis.* 88:565-571.
30. Oyarzún, P. J., Taipe, J. A., and Forbes, G. A. 2001. *Phytophthora infestans* characteristics and activity in Ecuador—Country profile. Pages 15-25 in: Proc. Int. Workshop Complementing Resistance to Late Blight (*Phytophthora infestans*) in the Andes. E. N. Fernández-Northcote, ed. Cochabamba, Bolivia. International Potato Center (CIP), Lima, Peru.
31. Pérez, W. G., Gamboa, J. S., Falcón, Y. V., Coca, M., Raymundo, R. M., and Nelson, R. J. 2001. Genetic structure of Peruvian populations of *Phytophthora infestans*. *Phytopathology* 91:956-965.
32. Precheur, R. J. (ed.) 2003. Ohio Vegetable Production Guide. The Ohio State Univ. Bull. 672-03.
33. Rykiel, E. J. 1996. Testing ecological models: The meaning of validation. *Ecol. Model.* 90:229-244.
34. Sargent, R. G. 1984. A tutorial on verification and validation of simulation models. Pages 114-121 in: Proc. 1984 Winter Simulation Conference. S. Sheppard, U. W. Pooch, and C. D. Pedgen, eds. IEEE Press, Piscataway, NJ.
35. Schlesinger, D., Crosbie, R. E., Gagné, R. E., Innis, G. S., Lalwani, C. S., Loch, J., Sylvester, R. J., Wright, R. D., Kheir, N., and Bartos, D. 1979. Terminology for model credibility. *Simulation* 32:103-104.
36. Swiader, J. M., and Ware, G. W. 2002. Producing Vegetable Crops. 5th ed. Interstate Publishers, Danville, IL.
37. Willocquet, L., Savary, S., Fernandez, L., Elazegui, F. A., Castilla, N., Zhu, D., Tang, Q., Huang, S., Lin, X., Singh, H. M., and Srivastava, R. K. 2002. Structure and validation of RICEPEST, a production situation-driven, crop growth model simulating rice yield response to multiple pest injuries for tropical Asia. *Ecol. Model.* 153:247-268.