

Prediction of late blight severity in a large panel of potato genotypes using low-altitude aerial images and machine learning methods

Hildo Loayza^{1,3*}, Johan Ninanya¹, Susan Palacios², Luis Silva¹, Fernando Pujaico Rivera³, Javier Rinza¹, Manuel Gastelo⁴, Mariela Aponte¹, Jan F. Kreuze¹, Hannele Lindqvist-Kreuze¹, Bettina Heider¹, Moctar Kante¹, David A. Ramírez¹

¹ International Potato Center (CIP), Headquarters P.O. Box 1558, Lima 15024, Peru

² Pontificia Universidad Católica del Perú, Lima 15088, Peru

³ Universidad de Huánuco, Huánuco, Peru

⁴ Programa Nacional de Investigación Científica y Estudios Avanzados - PROCIENCIA-CONCYTEC, <https://prociencia.gob.pe>, Lima, Peru

* Corresponding author

E-mail: h.loayza@cgiar.org

Abstract

Potato (*Solanum tuberosum* L.) is a staple crop crucial to global food security, yet its production is severely threatened by late blight (LB), caused by *Phytophthora infestans*, one of the most destructive plant diseases worldwide. Breeding programs for LB resistance have traditionally relied on labor-intensive and subjective visual assessments, which limit scalability and consistency, particularly in early-generation trials. Unmanned aerial vehicle (UAV)-based remote sensing combined with machine learning (ML) offers a promising alternative for objective, high-throughput disease phenotyping. This study evaluated the potential of UAV-derived multispectral imagery and ML techniques to estimate LB severity across large and genetically diverse potato breeding populations, comprising 2,745 clones in one trial and 492 accessions in another, conducted in Oxapampa, Pasco, Peru. We compared vegetation index-based approaches with a machine learning framework that integrates K-means clustering and Kernel Ridge Regression (KRR) and assessed their ability to capture genotypic variation and support selection decisions. NDVI consistently showed a strong correlation with visually assessed LB severity, particularly at advanced stages of disease development, enabling objective discrimination between healthy and diseased canopy tissues. However, the KRR-based approach outperformed linear NDVI-based models by capturing nonlinear relationships between spectral responses and disease progression. Estimates of LB severity derived from NDVI and KRR models, expressed as best linear unbiased estimates (BLUEs), showed strong and biologically consistent relationships with the area under the disease progress curve (AUDPC), particularly during later UAV acquisitions. Selection coincidence between UAV-derived estimates and AUDPC-based rankings was substantially higher at intermediate to advanced stages of disease progression, suggesting that UAV assessments at these stages may capture sufficient phenotypic variation to distinguish genotypes. These

findings indicate that UAV-based multispectral phenotyping, especially when integrated with ML, provides a practical and scalable approach for assessing LB severity in potato breeding programs while reducing the need for time-consuming field evaluations.

Introduction

Potato (*Solanum tuberosum* L.) is a staple crop of critical importance to global food security and is the most widely grown non-cereal crop worldwide [1,2]. Potatoes contribute significantly to caloric intake, nutrition, and rural income [3]. However, the constant threat of numerous diseases, particularly late blight (LB), poses a severe challenge to potato production [4,5]. The pathogen *Phytophthora infestans* (Mont.) de Bary is an oomycete that causes LB, a highly virulent disease that can devastate potato crops. Early-season infections can lead to significant yield losses of up to 80%. This results in substantial economic hardship for farmers who rely on this crucial crop, with annual losses estimated to be between 3 and 10 billion dollars [4,6,7]. The environmental, health, and financial implications of LB management further underscore the need for effective, sustainable solutions. LB spreads rapidly under optimum environmental conditions, especially at temperatures ranging between 10 and 23 °C and a relative humidity of more than 80% for 10 h daily [8–10]. It causes damage to foliage and tubers, which can lead to the loss of the entire crop within one week [11].

Potato breeding programs play an important role in LB integrated management by developing varieties with disease resistance, hence reducing reliance on chemical controls, lowering production costs, and contributing to sustainable agriculture [12–14]. The breeding process typically relies on visual disease assessments under controlled infection conditions, in which trained experts evaluate disease progression across several genotypes [15,16]. These

assessments involved estimating the percentage of infected leaf area per plant, assuming that LB was the only foliar disease present in the experiment. However, this method is inherently labor-intensive, time-consuming, and subject to observer bias, which limits its scalability and precision [17]. Furthermore, breeding programs targeting LB resistance face numerous challenges, including the need to test thousands of genotypes in early generation trials [18]. This large-scale phenotyping requires considerable human resources and incurs high costs due to the intensive, repeated manual assessments conducted throughout the growing season. Additionally, environmental variability and subjective assessment practices can introduce inconsistencies, complicating the accurate identification of resistant genotypes in different environments. These challenges underscore the need for alternative phenotyping approaches to increase consistency, scalability, and cost efficiency [19].

Advances in remote sensing offer potential solutions to these limitations in phenotyping. Remote sensing tools, including sensors mounted on unmanned aerial vehicles (UAVs), with improved spatial and temporal resolutions, now enable high-throughput, non-destructive data collection across extensive breeding plots [20,21]. These tools provide rapid and objective data on disease progression, enabling breeders to efficiently and at scale assess crop health. For example, RGB and multispectral images can be used to compute vegetation indices that are highly correlated with plant health indicators, providing an alternative approach to disease assessment [22]. The combination of multispectral imagery with machine learning (ML) and deep learning (DL) models enables the detection of complex disease-related patterns in large datasets, facilitating efficient disease detection and classification [17,23,24]. Thus, multispectral images acquired from UAVs, when integrated with ML algorithms, have become powerful tools for the classification, segmentation, and detection of plant diseases [17,23]. In potato disease phenotyping, particularly

for LB, UAV-based multispectral imaging and ML models have demonstrated promising results [17,25]. Despite these advancements, applying such tools to early-generation trials with several thousand genotypes remains challenging due to data-processing requirements and the need for consistent, reliable phenotyping.

This study aimed to evaluate the potential of UAV-based imagery, along with vegetation index-based approaches and ML techniques, to assess LB severity across a large panel of early-generation potato breeding trials. Specifically, the study's objectives are as follows: i) to quantify the extent to which aerial phenotyping captures genetic variation in LB resistance, by comparing image-derived indicators with conventional field assessments across diverse genotypes; ii) to analyze the potential reduction of temporal assessments when UAV-based imagery is used to monitor genotypes' LB severity. We hypothesized that ML modeling can effectively capture nonlinear relationships in multispectral reflectance data obtained from aerial imaging and disease severity. This capability may enhance the screening process for LB resistance across a large panel of genotypes. As a result, ML could reduce the time required for numerous field assessments typically conducted through conventional phenotyping, especially during critical developmental stages.

Materials and methods

Study area

Two potato trials were conducted in Oxapampa, Pasco, Peru (10.6039° S, 75.4153° W, 1,827 m a.s.l.), with the first trial (Trial A) from September 24, 2019, to January 7, 2020, and the second trial (Trial B) from October 21, 2019, to February 11, 2020 (Fig 1). Oxapampa is a site with high LB disease pressure [26], characterized by a temperate, rainy, and very humid climate

year-round, with minimum and maximum temperatures ranging from 11 to 17 °C and 25 to 29 °C, respectively, and annual precipitation between 1200 and 3000 mm [27]. In both trials, the minimum and maximum temperatures ranged from 8.6 to 16.2 °C and 18.9 to 27.0 °C, respectively, whereas the relative humidity ranged from 69.0 to 94.7% (S1 Fig). During the crop growing season, the total precipitation was 559.6 mm for Trial A and 894.7 mm for Trial B (S1 Fig). The soil had a loam texture (46, 37, and 17% of sand, silt, and clay, respectively) with phosphorus, potassium, and organic matter contents of 53.1 ppm, 235 ppm, and 4.55%, respectively, as determined by the Soil, Plant, Water, and Fertilizer Analysis Laboratory of the National Agrarian University-La Molina, Lima, Peru.

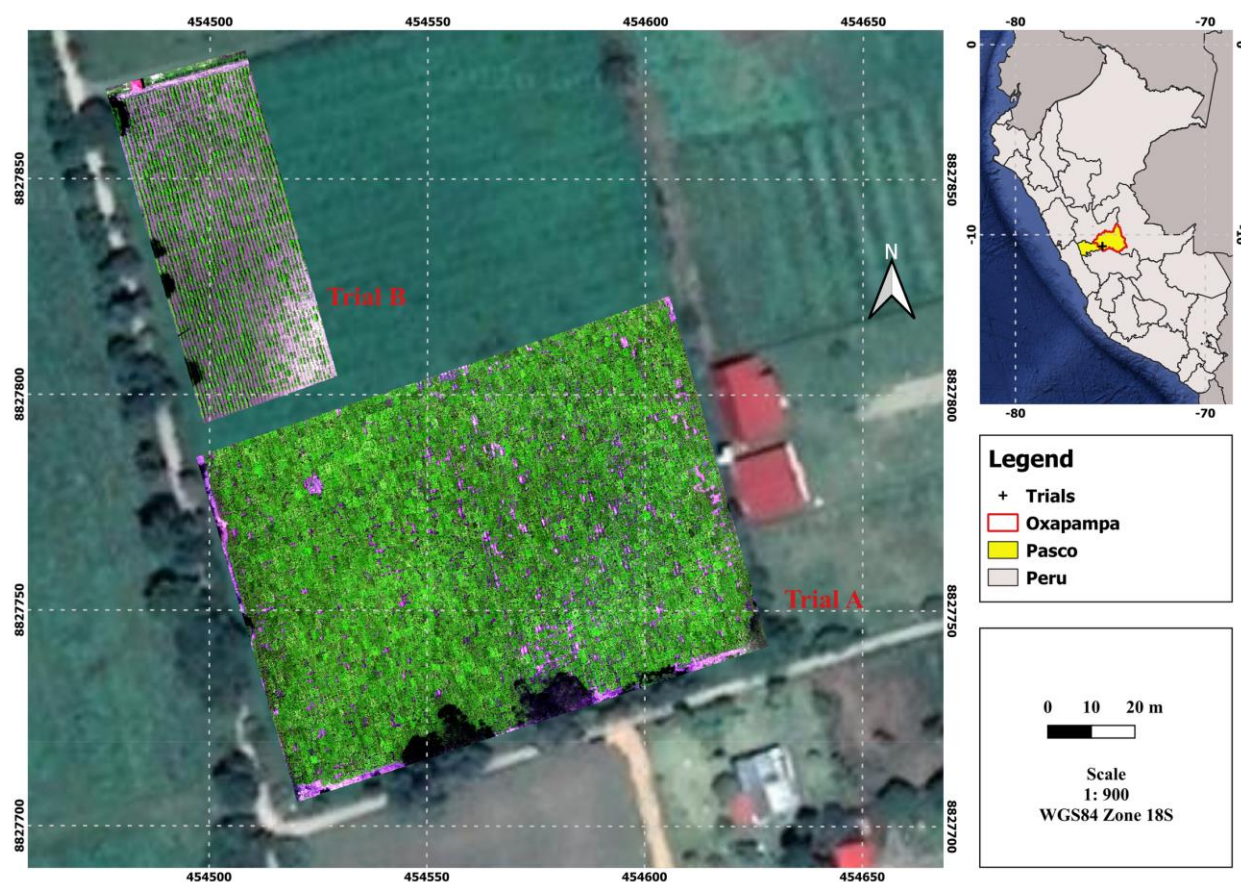


Fig 1. The study area. Location of the study area showing two potato trials in Oxapampa, Pasco, Peru: one including 2,745 clones with two replications over 0.9 ha (Trial A), and another involving 492 accessions with three replications covering 0.3 ha (Trial B).

Plant material and experimental design

In Trial A, 2,745 potato clones were evaluated as part of the preliminary yield trials of cycle 2 of the CIP LB-resistant and heat-tolerant (LBHT) breeding pipeline using a 43×65 resolvable row-column experimental design with two replications. In Trial B, a heterogeneous set of 492 potato accessions, comprising both diploid and tetraploid accessions from the CIP genebank core collection, was planted in a 12×41 row-column experimental design with three replications. In both trials, the plots consisted of single rows with five plants each, with a spacing of 0.3 m between plants and 0.9 m between plots. A common set of three varieties with known responses to LB—Amarilis (moderately resistant), Yungay (highly susceptible), and Kory (resistant)—was included in both trials as controls [15]. In Trial A, Amarilis and Yungay were replicated 34 times, whereas Kory was replicated 32 times. In Trial B, all three varieties were replicated 13 times each. Moreover, Yungay and Amarilis were planted around the field as “spreader rows” to allow for inoculum establishment and dissemination [28]. A description of the potato clones from Trials A and B is available in the CIP Dataverse repository [29,30].

Both trials were conducted according to standard agronomic practices [15] until all plants had emerged and grown sufficiently large for disease assessment, which was performed 40 days after planting (DAP). Consequently, to prevent early LB infection and before allowing natural LB infestation, fungicide Dithane M-45 WP NT (Mancozeb, BASF®, Peru) was applied preventively twice, at 80% and at full emergence. In both trials, fertilizers were applied at sowing at a rate of

200-180-160 kg ha⁻¹ of N-P₂O₅-K₂O, supplied as ammonium nitrate (238.3 kg), diammonium phosphate (373.5 kg), and potassium sulfate (305.4 kg). Additionally, 238.3 kg of ammonium nitrate was applied at 30 DAP (first hilling) to complete the remaining 50% of the required nitrogen. Potatoes were grown under rainfed conditions (S1 Fig), with total precipitation exceeding the 500 mm crop water requirement for 120-day potato crops [31].

Field late blight severity assessment, image acquisition, and preprocessing

LB severity was assessed as the percentage of LB-infected foliar area per plant, visually estimated by trained experts assuming LB was the only foliar disease present in the plant, and recorded on a 0–100% scale in 5% increments [15]. LB severity assessments were conducted weekly at the plot level, comprising seven assessments in Trial A (starting at 41 DAP) and five assessments in Trial B (starting at 38 DAP). From these sequential assessments, disease development over time was quantified by calculating the Area Under the Disease Progress Curve (AUDPC) using the trapezoidal integration method [15,32].

Two UAV flights were conducted to acquire multispectral aerial images corresponding to two LB severity assessments: 64 and 86 DAP for Trial A, and 38 and 58 DAP for Trial B, respectively. These flights generated four datasets by combining LB severity observations with corresponding multispectral imagery. A RedEdge M camera (MicaSense, Seattle, USA) was mounted on a DJI Inspire 2 quadcopter (Shenzhen, China) to capture 16-bit images in five spectral bands: blue (475 ± 20 nm), green (560 ± 20 nm), red (668 ± 10 nm), red edge (717 ± 10 nm), and near-infrared (NIR; 840 ± 40 nm). The UAV flight plan was designed to ensure 75% overlap between images in both the frontal and lateral directions. Images were acquired at altitudes of 50

m in Trial A and 40 m in Trial B, resulting in spatial resolutions of $3.4 \times 3.4 \text{ cm}^2$ and $2.9 \times 2.9 \text{ cm}^2$, respectively. Flights occurred between 11:00 and 14:00 h local time, when potato plants exhibit maximum non-photochemical quenching [33]. Images of a calibrated reference panel were acquired before and after each flight to convert the radiance measurements into reflectance values. The multispectral dataset was processed in Pix4Dmapper (Prilly, Switzerland) for radiometric correction, geometric alignment, and mosaicking, producing a five-band reflectance orthomosaic for each flight in both trials.

Late blight severity estimation

As a first step, reflectance orthomosaics from both trials were segmented using the Dzetsaka plug-in for QGIS, which implements a K-Nearest Neighbors (KNN) classifier [34]. The classifier was trained to label pixels and generate soil masks for each assessment date and trial; these masks were subsequently applied to produce orthomosaics containing only vegetation pixels from the potato plants. Subsequently, polygonal grids were created, labeled, and assigned to each plot according to the experimental design, resulting in 5,790 grids in Trial A and 1,476 grids in Trial B per dataset. Two methodologically distinct approaches were implemented to estimate LB severity (Fig 2): i) a vegetation-index-based linear regression approach applied at both the plot and pixel levels; and ii) a machine-learning-based approach that combines K-means clustering and Kernel Ridge Regression (KRR) using multispectral reflectance data.

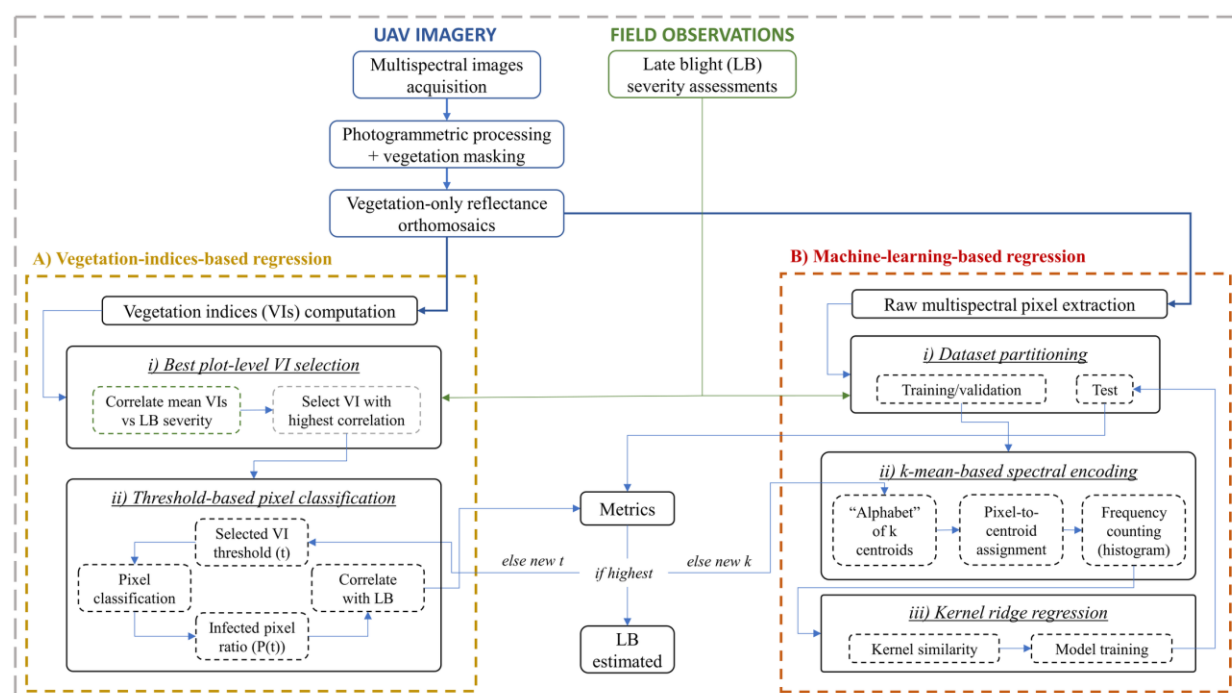


Fig 2. Workflow of NDVI- and KRR-based LB severity estimation. Methodology workflow of the two approaches used in this study: (A) Estimation of late blight (LB) severity using linear regression based on vegetation indices, and (B) a regression-based LB severity prediction system using Kernel Ridge Regression (KRR) combined with K-means clustering. In approach B, a dataset composed of five-dimensional multispectral pixel samples is first processed using K-means clustering with K centroids. This procedure generates a histogram vector with K elements, which is then used as input to the KRR model to estimate LB severity.

First approach: vegetation-index-based regression

This approach comprised two key steps (Fig 2A): i) identifying the vegetation index that best represents LB severity assessments, and ii) classifying potato plant cover as LB-infected or healthy using a threshold criterion. First, the plot-level mean values of four widely used vegetation indices—the Normalized Difference Vegetation Index (NDVI; [35]), Normalized Difference Red

Edge (NDRE; [38]), Red Edge reflected band (RE), and Green Normalized Difference Vegetation Index (GNDVI; [39])—were computed for each LB severity observation dataset. These mean values were correlated with their corresponding LB severity assessments, and the vegetation index with the highest Pearson correlation coefficient (R) was selected as the “best” index for representing LB severity. Second, polygonal grids were used to crop the selected vegetation index images for each dataset. The resulting grid-based images were processed using a custom object-detection algorithm developed in MATLAB [40]. This algorithm identifies clusters of pixels in each grid (plot), encloses them within rectangular regions, labels them, and extracts all the associated pixel values [41]. Subsequently, for each plot, the full range of vegetation index values (0–1) was evaluated in increments of 0.05. Each increment was used as a threshold to classify the pixels as either LB-infected (pixel value $\leq$ threshold) or healthy (pixel value $>$ threshold). For each threshold, the ratio of LB-infected pixels to the total number of pixels within each polygonal grid was determined. These ratios were compared with their corresponding LB severity assessments, and the threshold yielding the highest R was selected as the “optimal” threshold for distinguishing LB-infected from healthy plant cover in each dataset. Finally, the ratio of LB-infected pixels at the optimal threshold was regressed against plot-level LB severity observations to derive linear regression models for each trial. Although all four datasets were analyzed, only the results at 86 DAP for Trial A and 58 DAP for Trial B are presented in Fig 3C and Fig 3E, respectively, because only a limited number of LB-infected plants were identified during the first assessments in each trial. The same criterion was applied to the second approach as well.

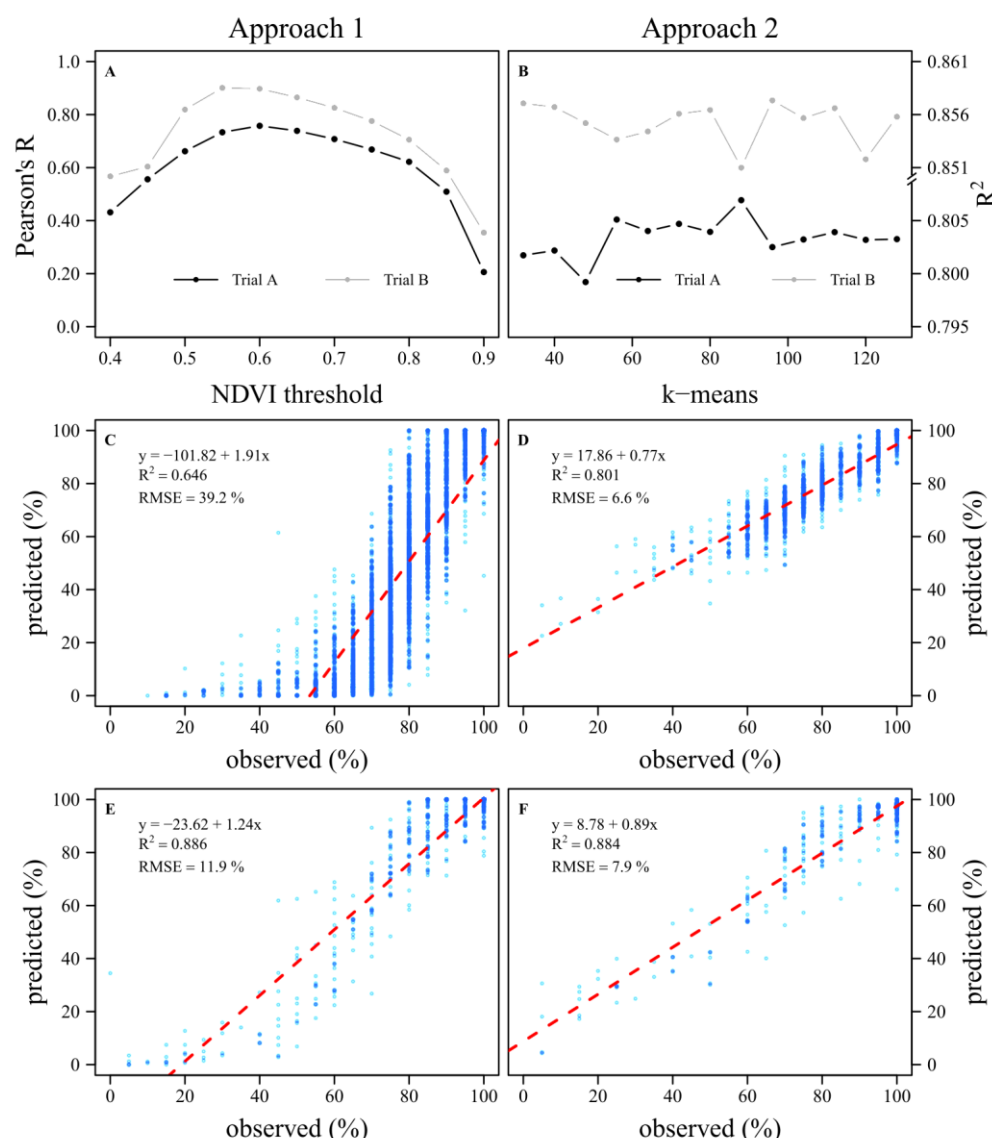


Fig 2. Optimization and validation of NDVI and KRR-based models for LB severity. (A) Pearson correlation coefficient (R) between observed and predicted Late Blight (LB) severity across a range of NDVI thresholds for Trials A and B. (B) Selection of the optimal number of clusters (K) in the K-means algorithm based on R^2 values from linear regressions between observed and predicted LB severity using Gaussian kernel ridge regression (KRR) models during the training and validation stages for Trials A and B. (C) Linear regression between observed and predicted LB severity using the optimal NDVI threshold for Trial A (2,745 potato clones). (D) Linear regression between observed and predicted LB severity using Gaussian kernel ridge regression (KRR) models during the training and validation stages for Trials A and B. (E) Linear regression between observed and predicted LB severity using the optimal NDVI threshold for Trial A (2,745 potato clones). (F) Linear regression between observed and predicted LB severity using Gaussian kernel ridge regression (KRR) models during the training and validation stages for Trials A and B.

Linear regression between observed and predicted LB severity using a Gaussian KRR model applied to the testing subset and the optimal K value for Trial A. (E) Same as panel C, but for Trial B (492 potato accessions). (F) Same as panel D but for Trial B. All analyses were conducted using data acquired at 86 Days After Planting (DAP) for Trial A and 58 DAP for Trial B.

Second approach: machine-learning-based regression using multispectral reflectance data

The proposed system was implemented in Python as a modular and customizable ML framework, comprising two interconnected subsystems, as described by [42] (Fig 2B). The first subsystem receives as input the set u , which contains L samples $s_l \in \mathbb{R}^5$, for all $l \in \{1, 2, \dots, L\}$, where each s_l represents a five-dimensional multispectral pixel extracted from the polygonal grid of a plot (u) in which LB severity observations (y) were performed. The input set u is then processed through the prediction stage of a K-means clustering algorithm [43], using K centroids. Subsequently, a histogram of the input samples s_l is constructed based on the centroids $c_k \in \mathbb{R}^5$, for all $k \in \{1, 2, \dots, K\}$. This procedure yields a histogram vector $x \in \mathbb{R}^K$, whose elements sum to 1. The K-means model was fitted as described below. Finally, the vector x is used as input to the second subsystem, a Gaussian kernel ridge regression model [44,45], which outputs a prediction $y \in \mathbb{R}$, representing the estimated LB severity of the multispectral five-dimensional pixel set u .

K-means clustering and histogram-based analysis.

To apply the K-means algorithm, this must first be fitted to identify the K centroids $c_k \in \mathbb{R}^5$, for all $k \in \{1, 2, \dots, K\}$. For this purpose, the M sets u_m , for all $m \in \{1, 2, \dots, M\}$ in the dataset—where M denotes the number of plots—are combined into a single set of elements in $\mathbb{R}^5$.

The K-means algorithm is then applied to this unified set to determine all centroids c_k . These centroids are subsequently used in a manner like a bag-of-words system. Each set of pixels corresponding to plot m (u_m) is encoded using this “alphabet” of K centroids, such that each element of u_m is assigned to the nearest centroid to create a frequency histogram. The histogram is normalized so that its elements sum to 1, resulting in a vector $x_m \in \mathbb{R}^K$, which represents the coded version of the set u_m . This procedure is particularly important because each set u_m may contain a different number of samples ($L_m \in \mathbb{Z}$); however, the output of this subsystem always returns a fixed-size vector x_m of length K , independent of the number of pixels in u_m . For each dataset (64 and 86 DAP for Trial A, and 38 and 58 DAP for Trial B), an analysis was performed to determine the optimal value of K by evaluating values ranging from 32 to 128 in increments of 8. The optimal K value was selected as the value that maximizes the R^2 score during the validation stage of training a Gaussian kernel ridge regression model, as described below.

Gaussian kernel ridge regression model.

A kernel ridge regression (KRR) model was implemented using a Gaussian kernel (Radial Basis Function) [44,45]. The objective was to find the weight vector (w) that minimizes the cost function $L_{\alpha,\gamma}(w)$, given the hyperparameters α and γ :

$$L_{\alpha,\gamma}(w) = \|y - \hat{\Phi}_\gamma(X)^T w\|^2 + \alpha \|w\|^2 \quad (1)$$

where $\hat{\Phi}_\gamma(X)$ is the vectorial function that generates the kernel matrix $\hat{K} = \hat{\Phi}_\gamma(X)^T \hat{\Phi}_\gamma(X)$, constructed using a radial basis function with parameter γ . The vector y contains all target values y_m , for all $m \in \{1, 2, \dots, M\}$, while the matrix X includes all input samples x_m^T , with each sample represented as a row. Each dataset was split into two subsets: 80% for training and validation and

20% for testing. For the training and validation subsets, a 5-fold cross-validation procedure was applied, with four folds used for training and the fifth for validation, yielding five different combinations. After this process, the R^2 score was calculated for each validation fold, and the mean validation R^2 was computed for each model. This value was used to optimize α (regularization term), γ (kernel width), and K (K-means) through a grid search that maximized the mean validation R^2 .

Linear mixed model and rank-based selected coincidence analysis

To perform fair trial-level comparisons among genotypes, a linear mixed model was used to extract the best linear unbiased estimator (BLUE), which represents the adjusted genotype means accounting for trial design and environmental effects. The BLUEs were determined for the raw plot-level AUDPC and LB severity estimated using normalized difference vegetation index-based linear regression (LB_{LR}) and kernel ridge regression-based machine-learning models (LB_{KRR}). Thus, separate linear fixed models were fitted for Trials A and B, using the “statgenSTA” R package [46].

$$Y_{ijkl} = \mu + G_i + R_j + F_k + C_l + \varepsilon_{ijkl} \quad (2)$$

where Y_{ijkl} is the performance of the i^{th} genotype in the j^{th} replication located in the k^{th} row and l^{th} column. μ is the overall mean, G_i is the fixed effect of the i^{th} genotype, while R_j (j^{th} replication effect), F_k (k^{th} row effect), and C_l (l^{th} column effect) were treated as random effects to account for spatial variability. ε_{ijkl} is the error term associated with each Y_{ijkl} .

Later, Spearman’s rank correlation was used to assess the relationship between BLUEs of AUDPC and BLUEs of estimated LB severity (LB_{LR} and LB_{KRR}). Kernel density distributions of standardized BLUE values (z-scores) were computed using the “stats” package in R to visualize

and compare distribution patterns between AUDPC and model-estimated LB severity. Furthermore, the percentage of selected coincidence (SC) between AUDPC- and LB-based rankings within the top 10%, 20%, and 30% of clones ranked by AUDPC was calculated comparing the selections from both rankings. This analysis was conducted for both UAV assessment dates to determine whether a single time-point estimate of LB severity could provide comparable genotype rankings to those obtained with the integrative AUDPC indicator for early breeding selection.

Model performance evaluation

Model performance was evaluated by comparing observed and estimated LB severity values using R^2 and the Root Mean Square Error (RMSE). All statistical analyses were performed using the R programming language [47].

Results

Estimation of late blight severity using vegetation indices

NDVI was the most representative vegetation index for assessing plot-level LB severity, exhibiting the strongest correlation with LB severity across almost all datasets (-0.55 to -0.86, $p < 0.001$), except at 38 DAP in Trial B (Table 1). GNDVI and RE ranked second in Trials A and B, respectively, across both assessment dates (Table 1). Overall, plot-level NDVI showed stronger correlations with LB severity in the second assessment (-0.41 to -0.86) than in the first assessment (-0.18 to -0.55) (Table 1). Sweeping NDVI thresholds produced bell-shaped response curves relating NDVI threshold values to the correlations between threshold-based LB severity estimates and observed ones, with correlation values ranging from 0.21 to 0.76 in Trial A and from 0.35 to

0.90 in Trial B (Fig 3A). The strongest correlations were found at NDVI thresholds of 0.60 for Trial A and 0.55 for Trial B, both during the second assessment (Fig 3A). Linear regression between genotype-level observed LB severity and severity estimated using optimal NDVI thresholds showed contrasting performances across trials and assessment dates (Fig 3C, E). In Trial A, model performance was weak at 64 DAP ($R^2 = 0.19$, $p < 0.001$; RMSE = 15.5%) but improved at 86 DAP ($R^2 = 0.65$, $p < 0.001$; RMSE = 39.2%). In Trial B, the model performance showed a similar pattern, being weak at 38 DAP ($R^2 = 0.23$, $p < 0.001$; RMSE = 47.8) and substantially stronger at 58 DAP ($R^2 = 0.89$, $p < 0.001$; RMSE = 11.9%).

Table 1. Pearson correlation (R) values between mean vegetation indexes—Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation Index (GNDVI), Red Edge reflected (RE), and Normalized Difference Red Edge (NDRE)—and Late Blight severity observations from Trials A (1,476 plots) and B (5,490 plots). Asterisks (*) indicate statistically significant correlations at $p < 0.001$. DAP = Days after planting.

Vegetation Index	Trial A		Trial B	
	64 DAP	86 DAP	38 DAP	58 DAP
NDVI	-0.55 *	-0.86 *	-0.33 *	-0.60 *
GNDVI	-0.51 *	-0.64 *	-0.22 *	-0.41 *
RE	-0.40 *	-0.57 *	-0.37 *	-0.46 *
NDRE	-0.50 *	-0.54 *	-0.18 *	-0.45 *

Estimation of LB severity from multispectral reflectance data using clustering and KRR

The optimal number of clusters (K) varied across trials and assessment dates (Table 2, Fig 3B). In Trial A, the highest performance was achieved with $K = 96$ at 64 DAP and $K = 88$ at 86 DAP, whereas in Trial B, the highest performance was obtained with $K = 104$ at 38 DAP and $K = 96$ at 58 DAP (Table 2, Fig 3B). Overall, KRR models showed stronger predictive performance at the second assessment date, characterized by smaller values of the hyperparameters α and without pattern in the hyperparameter γ (Table 2). In Trial A, the model performed better at 86 DAP ($R^2 = 0.80$, $p < 0.001$; RMSE = 6.6%) than at 64 DAP ($R^2 = 0.46$, $p < 0.001$; RMSE = 4.8%). In Trial B, the model performance was lower at 38 DAP ($R^2 = 0.46$, $p < 0.001$; RMSE = 5.6%) and higher at 58 DAP ($R^2 = 0.88$, $p < 0.001$; RMSE = 7.9%) (Table 2, Fig 3D, F).

Table 2. Tuned hyperparameters—number of centroids (K), regularization term (α), and kernel width (γ)—of the combined K-means clustering and Kernel Ridge Regression model, and regression metrics—coefficient of determination (R^2) and root mean square error (RMSE)—comparing model-estimated and observed late blight severity in two large potato breeding trials. Asterisks (*) indicate statistically significant correlations at $p < 0.001$. DAP = Days after planting.

Parameters and metrics	Trial A		Trial B	
	64 DAP	86 DAP	38 DAP	58 DAP
K	96	88	104	96
α	0.0608	0.0025	0.1155	0.0608
γ	2.6567	0.6199	5.3210	8.9340
R^2	0.46 (*)	0.80 (*)	0.46 (*)	0.88 (*)
RMSE	4.8 %	6.6 %	5.6 %	7.9 %

Relationship between AUDPC and estimated LB severity

A wide range of disease susceptibility levels was observed across genotypes in both trials, with BLUEs of AUDPC ranging from 24.1 to 2007.0%-days in Trial A and from 32.2 to 2322.8%-

days in Trial B (Fig 4). The kernel density distributions of BLUEs of LB_{LR} and LB_{KRR} generally followed the AUDPC distribution across trials, although model-based estimates showed narrower spreads and slight shifts relative to AUDPC, particularly in the second assessment of Trial B (Fig S3). During the first UAV assessment, BLUEs of model-based LB severity estimates were relatively low for most genotypes, particularly in Trial A, resulting in a weaker association with AUDPC (Spearman's coefficient ranging from 0.311 to 0.707) (Fig S2). In the second UAV assessment, BLUEs of LB severity showed a more homogeneous distribution across low, intermediate, and high values, resulting in stronger correlations with AUDPC (Spearman's coefficient ranging from 0.774 to 0.895) (Fig 4). The percentage of SC between AUDPC- and LB-based rankings was lower during the first UAV assessment, ranging from 17.1% to 48.1% in Trial A and from 46.8% to 70.5% in Trial B, depending on the selection intensity and LB estimation method (Fig S2). In contrast, substantially higher SC values were observed in the second UAV assessment, ranging from 47.6% to 77.4% in Trial A and from 80.9% to 91.5% in Trial B (Fig 4). Overall, LB estimates obtained with KRR showed better performance in selecting top-ranked genotypes across assessment dates and selection intensities (Figs S2 and 4).

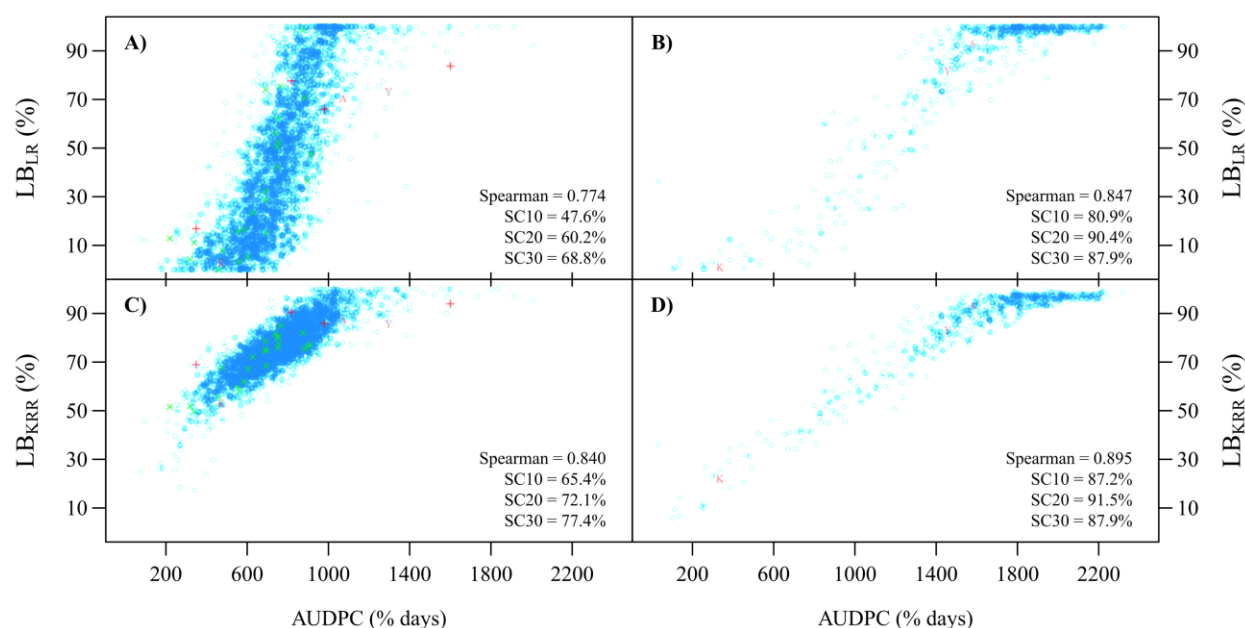


Fig 4. BLUEs of AUDPC vs. UAV-based LB severity estimates. Best linear unbiased estimators (BLUEs) of the area under the disease progress curve (AUDPC) versus BLUEs of estimated late blight (LB) severity in large potato breeding trials. LB severity was estimated by two approaches: normalized difference vegetation index-based linear regression (LB_{LR}; panels (A) and (B)) and kernel ridge regression-based machine-learning models (LB_{KRR}; panels (C) and (D)). Panels (A) and (C) show results for 2,745 clones (Trial A) evaluated at 86 days after planting (DAP), whereas panels (B) and (D) show 492 accessions (Trial B) evaluated at 58 DAP. Red symbols (with A=Amarilis, K=Kori, and Y=Yungay common to both trials) and green symbols represent check and parent varieties, respectively. SC10, SC20, and SC30 denote the percentage of selected coincidence between AUDPC- and LB-based rankings within the top 10%, 20%, and 30% of clones ranked by AUDPC.

Among the checks cultivars, Amarilis and Yungay consistently exhibited high BLUEs of AUDPC across both trials, with values ranging from 1,071.7 to 1,296.1 % -days in Trial A and from 1,456.1 to 1,578.2 % -days in Trial B. Consistent with these high disease progress values,

both cultivars also showed elevated BLUEs of LB severity, as estimated from NDVI (70.5 to 73.2 % in Trial A and 81.7 to 92.8 % in Trial B) and KRR-based models (85.7-87.4 % in Trial A and 83.1-90.3 % in Trial B), in agreement with their known susceptibility profiles. In contrast, the resistant cultivar Kory consistently displayed low BLUEs of AUDPC in both trials, with values of 473.7 %-days in Trial A and 336.0 %-days in Trial B. Correspondingly, BLUEs of LB severity were markedly lower, ranging from 2.8% (NDVI-based) and 53.0% (KRR-based) in Trial A to 1.0% (NDVI-based) and 22.0% (KRR-based) in Trial B (Fig 4), reinforcing its comparatively resistant response.

Discussion

UAV-based remote sensing combined with K-means-KRR improves late blight severity estimation over NDVI-based models

Late blight progresses through two distinct phases: an initial biotrophic phase involving physiological changes, mainly chlorophyll reduction [48–50], without visible symptoms [51], and a subsequent necrotrophic phase characterized by structural tissue damage, including cell wall degradation [52,53] and membrane disruption [54–57]. These biochemical and structural modifications directly alter the canopy's spectral reflectance properties, making remote sensing a viable approach for detecting and quantifying disease severity. Spectral reflectance in the near-infrared (NIR) region is primarily governed by the internal leaf structure and dry matter content, including proteins, lignin, and cellulose [58,59], whereas chlorophyll degradation predominantly affects reflectance in the blue and red absorption bands [38,60]. Consequently, vegetation indices that combine NIR with blue or red spectral bands, such as NDVI, enhance the remote detection of

LB symptoms across both biotrophic (early) and necrotrophic (advanced) stages. Consistent with this principle, NDVI showed strong associations with LB severity in this study. The identification of threshold values (0.6 and 0.55) enabled objective classification of healthy and diseased canopy tissues across assessment periods. Similar threshold-based approaches have proven effective for disease severity detection in several crops [61–64], supporting the robustness of vegetation index-derived metrics for plant disease monitoring.

Beyond vegetation indices, integrating K-means clustering with Kernel Ridge Regression provided a scalable framework for estimating LB severity across large breeding trials. Clustering allowed heterogeneous pixel distributions to be summarized into fixed-length representations, facilitating high-throughput phenotyping across variable canopy structures [65]. The KRR-based model for estimating LB severity (LB-KRR) further captured nonlinear relationships within multispectral reflectance data, leading to consistently improved predictive performance relative to NDVI-based linear models [66,67]. Fine-tuning key hyperparameters enhanced the model's sensitivity to both early physiological changes and severe late-stage damage. As a result, the LB-KRR approach achieved superior predictive performance during early assessments in both trials, whereas at later stages, improvements were primarily reflected in reduced RMSE, indicating enhanced plot-level prediction precision.

Robust ML performance under breeding-scale conditions

A key contribution of this study lies in the scale at which ML-based LB severity estimation was evaluated. While previous studies using UAV imagery and advanced ML or deep learning models have reported high predictive performance, they have generally been limited to small experimental populations (typically fewer than 15 genotypes) [17,25,68,69] constraining their

transferability to breeding-scale applications. In contrast, the present study evaluated ML-based disease severity estimation across large, genetically diverse populations comprising 2,745 clones in Trial A and 492 accessions in Trial B. This scale introduces substantial variability—such as heterogeneous canopy architectures, differential disease progression, and variable spectral responses across genotypes—that is rarely addressed in literature. Despite these challenges, the K-means-KRR framework maintained robust predictive performance, highlighting its scalability and suitability for early-generation breeding pipelines characterized by large populations and high phenotypic diversity.

UAV-based methods can complement and, in specific contexts, partially replace repeated visual assessments

Conducting UAV-based phenotyping across multi-environment breeding trials introduces logistical constraints that can affect both data availability and model performance. Coordinating flights across locations requires suitable weather conditions, trained personnel, and available equipment, which often limit the number and timing of image acquisitions during the growing season. Similar challenges have been reported in large-scale UAV phenotyping studies [70–72]. Nevertheless, in breeding programs screening hundreds or thousands of genotypes, the primary objective is often to achieve high selection accuracy through reliable ranking of genotypes to prioritize promising clones for subsequent evaluation stages [73], rather than to obtain a precise quantification of LB severity. In this context, defining the optimal number and timing of UAV acquisitions becomes critical. Disease progression, canopy development, and environmental conditions strongly influence the spectral expression of plant stress and, in turn, the detectability of disease using optical sensors [62,74]. Evaluating the contribution of individual UAV

acquisitions across environments and disease stages is therefore essential to identify the most informative phenotyping windows. In this study, UAV flights conducted during intermediate to advanced stages of LB disease progression (e.g., the second flight) provided the most relevant information for genotype ranking and selection decisions. The high coincidence observed for the top 10% of genotypes (SC10 = 65.4% in Trial A and 87.2% in Trial B with KRR) indicates that a limited number of acquisitions—and potentially even a single well-timed flight—can approximate AUDPC-based selections while reducing the need for repeated field assessments.

Early-generation trials typically involve large populations, limited replication, and high phenotypic variability, making precise manual assessments challenging [73]. Given these conditions, the predictive performance observed in this study—although variable across environments and assessment dates—appears sufficient to support early-stage selection by enabling the identification of highly susceptible genotypes and prioritizing promising clones for further evaluation. Consistent with previous high-throughput phenotyping studies, even moderate prediction accuracies can contribute to genetic gain by enabling the screening of much larger breeding populations [75]. Furthermore, because visual assessments are often subject to observer bias and inconsistency across environments, UAV-derived estimates—even when moderately correlated—can improve the objectivity and efficiency of selection. When integrated with efficient data-processing pipelines, these UAV-derived phenotypes can provide timely and objective information that complements genomic-assisted breeding strategies and enhances selection efficiency in large-scale breeding programs [76].

Conclusion

UAV-based multispectral imaging provides a viable approach for estimating late blight severity under breeding field conditions, capturing biologically meaningful variation across large and genetically diverse populations, including 2,745 clones and 492 accessions. While NDVI offers a simple and effective baseline for disease detection, the integration of K-means clustering and kernel-based regression improves the representation of spectral–disease relationships by capturing nonlinear patterns, resulting in more accurate severity estimation and genotype ranking. When expressed as BLUEs, UAV-derived estimates showed strong and biologically consistent relationships with AUDPC and achieved higher selection coincidence, particularly at intermediate to advanced stages of disease development. These findings indicate that image-based phenotyping can reliably support genotype ranking for selection, especially when acquisitions are conducted at appropriate disease stages. Overall, UAV-based phenotyping represents a scalable and practical complement to conventional late blight assessments in early-generation breeding trials, where large population sizes limit precise manual evaluation. Future work should focus on optimizing acquisition strategies and evaluating their impact on selection decisions across diverse environments.

References

1. Devaux A, Goffart JP, Kromann P, Andrade-Piedra J, Polar V, Hareau G. The Potato of the Future: Opportunities and Challenges in Sustainable Agri-food Systems. *Potato Res.* 2021;64:681–720. doi:10.1007/s11540-021-09501-4
2. Zhang H, Xu F, Wu Y, Hu H hai, Dai X feng. Progress of potato staple food research and industry development in China. *Journal of Integrative Agriculture.* 2017;16:2924–32. doi:10.1016/S2095-3119(17)61736-2
3. Wijesinha-Bettoni R, Mouillé B. The Contribution of Potatoes to Global Food Security, Nutrition and Healthy Diets. *Am J Potato Res.* 2019;96:139–49. doi:10.1007/s12230-018-09697-1

- 511 4. Dong S meng, Zhou S qun. Potato late blight caused by *Phytophthora infestans*: From molecular
512 interactions to integrated management strategies. Journal of Integrative Agriculture. 2022;21:3456–
513 66. doi:10.1016/j.jia.2022.08.060
- 514 5. Gastelo M, Perez W, Quispe K, Pozo V. Phenotypic Stability and Correlation for Late Blight
515 Resistance in Advanced Potato Clones Under Field and Controlled Conditions. Am J Potato Res.
516 2022;99:150–9. doi:10.1007/s12230-022-09861-8
- 517 6. Guenthner JF, Michael KC, Nolte P. The economic impact of potato late blight on US growers. Potato
518 Res. 2001;44:121–5. doi:10.1007/BF02410098
- 519 7. Haverkort AJ, Boonekamp PM, Hutten R, Jacobsen E, Lotz LAP, Kessel GJT, et al. Societal Costs of
520 Late Blight in Potato and Prospects of Durable Resistance Through Cisgenic Modification. Potato
521 Res. 2008;51:47–57. doi:10.1007/s11540-008-9089-y
- 522 8. Andrade-Piedra JL, Hijmans RJ, Forbes GA, Fry WE, Nelson RJ. Simulation of Potato Late Blight in the
523 Andes. I: Modification and Parameterization of the LATEBLIGHT Model. Phytopathology®.
524 2005;95:1191–9. doi:10.1094/PHYTO-95-1191
- 525 9. Baker KM, Kirk WW, Stein JM, Andresen JA. Climatic Trends and Potato Late Blight Risk in the Upper
526 Great Lakes Region [Internet]. 2005. doi:10.21273/HORTTECH.15.3.0510
- 527 10. Günaçtı H, Ay T. Forecasting of Potato Late Blight Disease Using Alternative Sets of Meteorological
528 Data and Disease Epidemiology. KSU J Agric Nat. 2021;24:6. doi:10.18016/ksutarimdogan.vi.863578
- 529 11. Ray SS, Jain N, Arora RK, Chavan S, Panigrahy S. Utility of Hyperspectral Data for Potato Late Blight
530 Disease Detection. J Indian Soc Remote Sens. 2011;39:161–9. doi:10.1007/s12524-011-0094-2
- 531 12. Kankwatsa P, Adipala E, Hakiza JJ, Olanya M, Kidanemariam HM. Effect of Integrating Planting Time,
532 Fungicide Application and Host Resistance on Potato Late Blight Development in South-western
533 Uganda. Journal of Phytopathology. 2002;150:248–57. doi:10.1046/j.1439-0434.2002.00750.x
- 534 13. Nærstad R, Hermansen A, Bjør T. Exploiting host resistance to reduce the use of fungicides to
535 control potato late blight. Plant Pathology. 2007;56:156–66. doi:10.1111/j.1365-3059.2006.01491.x
- 536 14. Ritchie F, Bain RA, Lees AK, Boor TRW, Paveley ND. Integrated control of potato late blight:
537 predicting the combined efficacy of host resistance and fungicides. Plant Pathology. 2018;67:1784–
538 91. doi:10.1111/ppa.12887
- 539 15. Forbes GA, Pérez, W, Andrade-Piedra J. Field assessment of resistance in potato to *Phytophthora*
540 *infestans*: International cooperators guide [Internet]. International Potato Center; 2014 [cited 2022
541 Nov 19]. Available from: <https://hdl.handle.net/10568/65180> doi:10.4160/9789290604402
- 542 16. Yuen JE, Forbes GA. Estimating the Level of Susceptibility to *Phytophthora infestans* in Potato
543 Genotypes. Phytopathology®. 2009;99:782–6. doi:10.1094/PHYTO-99-6-0782
- 544 17. Duarte-Carvajalino JM, Alzate DF, Ramirez AA, Santa-Sepulveda JD, Fajardo-Rojas AE, Soto-Suárez
545 M. Evaluating Late Blight Severity in Potato Crops Using Unmanned Aerial Vehicles and Machine
546 Learning Algorithms. Remote Sensing. 2018;10:10. doi:10.3390/rs10101513

18. Duan Y, Duan S, Xu J, Zheng J, Hu J, Li X, et al. Late Blight Resistance Evaluation and Genome-Wide Assessment of Genetic Diversity in Wild and Cultivated Potato Species. *Front Plant Sci.* 2021;12. doi:10.3389/fpls.2021.710468
19. Sugiura R, Tsuda S, Tamiya S, Itoh A, Nishiwaki K, Murakami N, et al. Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle. *Biosystems Engineering.* 2016;148:1–10. doi:10.1016/j.biosystemseng.2016.04.010
20. Khanal S, Kc K, Fulton JP, Shearer S, Ozkan E. Remote Sensing in Agriculture—Accomplishments, Limitations, and Opportunities. *Remote Sensing.* 2020;12:22. doi:10.3390/rs12223783
21. Sishodia RP, Ray RL, Singh SK. Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sensing.* 2020;12:19. doi:10.3390/rs12193136
22. Bannari A, Morin D, Bonn F, Huete AR. A review of vegetation indices. *Remote Sensing Reviews.* 1995. Located at: world. doi:10.1080/02757259509532298
23. Madiwalar SC, Wyawahare MV. Plant disease identification: A comparative study. In: 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI). 2017. p. 13–8. doi:10.1109/ICDMAI.2017.8073478
24. Ramanjot, Mittal U, Wadhawan A, Singla J, Jhanjhi NZ, Ghoniem RM, et al. Plant Disease Detection and Classification: A Systematic Literature Review. *Sensors.* 2023;23:10. doi:10.3390/s23104769
25. Rodríguez J, Lizarazo I, Prieto F, Angulo-Morales V. Assessment of potato late blight from UAV-based multispectral imagery. *Computers and Electronics in Agriculture.* 2021;184:106061. doi:10.1016/j.compag.2021.106061
26. Perez WG, Gamboa JS, Falcon YV, Coca M, Raymundo RM, Nelson RJ. Genetic Structure of Peruvian Populations of *Phytophthora infestans*. *Phytopathology®.* 2001;91:956–65. doi:10.1094/PHYTO.2001.91.10.956
27. Gastelo M, Pérez W, Eyzaguirre R, Quispe K, Sanabria K, Bastos C, et al. CIP-PODEROSA CROCANTE, CIP-PODEROSA POLLERA, and CIP-PODEROSA WATIA: New Potato Varieties for Family Farming with Resistance to Late Blight and High Quality for the Frying Industry. *Am J Potato Res.* 2023;100:288–303. doi:10.1007/s12230-023-09917-3
28. Bonierbale M, Haan S, Forbes A, Crissman L, Zschocke T, Puccini A. Procedures for standard evaluation trials of advanced potato clones: An International Cooperators' guide. In. 2007 [cited 2024 Dec 10]. Available from: <https://www.semanticscholar.org/paper/Procedures-for-standard-evaluation-trials-of-potato-Bonierbale-Haan/20692a454e517ccc633ef9c5b0330f68ef1ded53>
29. Gastelo M, Quispe K, Bastos M, Pozo J, Erquinio E, Vega R. Dataset for: Breeding for resistance to late blight - LBHTC2 population- Phenotypic and Genomic Selection [Internet]. International Potato Center; 2020 [cited 2024 Feb 21]. Available from: <https://data.cipotato.org/dataset.xhtml?persistentId=doi:10.21223/GCJRH6> doi:10.21223/GCJRH6
30. Lindqvist-Kreuzer H, Aponte M. Dataset for: Genome wide association study (GWAS) to identify favourable alleles for late blight resistance in core collection [Internet]. International Potato Center;

2021 [cited 2024 Feb 27]. Available from:
<https://data.cipotato.org/dataset.xhtml?persistentId=doi:10.21223/YT4UZU> doi:10.21223/YT4UZU

31. Doorenbos J, Kassam AH, Bentvelsen C, Uittenbogaard G. Yield Response to Water1. In: Johl SS, editor. Irrigation and Agricultural Development [Internet]. Pergamon; 1980 [cited 2025 Aug 11]. p. 257–80. Available from: <https://www.sciencedirect.com/science/article/pii/B9780080256757500212> doi:10.1016/B978-0-08-025675-7.50021-2

32. Campbell CL, Madden LV. Introduction to Plant Disease Epidemiology. New York; 1990. 532 p.

33. Loayza H, Moya I, Quiroz R, Ounis A, Goulas Y. Active and passive chlorophyll fluorescence measurements at canopy level on potato crops. Evidence of similitude of diurnal cycles of apparent fluorescence yields. Photosynth Res. 2022. doi:10.1007/s11120-022-00995-8

34. Karasiak N, Perbet P. Remote Sensing of Distinctive Vegetation in Guiana Amazonian Park. In: QGIS and Applications in Agriculture and Forest [Internet]. John Wiley & Sons, Ltd; 2018 [cited 2023 Jan 9]. p. 215–45. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119457107.ch7> doi:10.1002/9781119457107.ch7

35. Rouse JW, Haas RH, Well JA, Deering, DW. Monitoring vegetation systems in the Great Plains with ERTS - NASA Technical Reports Server (NTRS). In. 1974 [cited 2022 Nov 19]. Available from: <https://ntrs.nasa.gov/citations/19740022614>

36. Barnes EM, Clarke TR, Richards SE, Colaizzi PD, Haberland J, Kostrzewski M, et al. Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. Proceedings of the 5th International Conference on Precision Agriculture and other resource management July 16-19, 2000, Bloomington, MN USA [Internet]. 2000 [cited 2022 Nov 19]. Available from: <http://hdl.handle.net/10113/4190>

37. Gitelson AA, Kaufman YJ, Merzlyak MN. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment. 1996;58:289–98. doi:10.1016/S0034-4257(96)00072-7

38. Gitelson A, Merzlyak MN. Quantitative estimation of chlorophyll-*a* using reflectance spectra: Experiments with autumn chestnut and maple leaves. Journal of Photochemistry and Photobiology B: Biology. 1994;22:247–52. doi:10.1016/1011-1344(93)06963-4

39. Gitelson AA, Kaufman YJ, Merzlyak MN. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment. 1996;58:289–98. doi:10.1016/S0034-4257(96)00072-7

40. The MathWorks Inc. MATLAB version: 9.2 (R2017a) [Internet]. Natick, Massachusetts, United States; 2017. Available from: <https://www.mathworks.com>

41. Loayza H. Automatic detection of potato plots to retrieve just the vegetation data of multispectral images [Internet]. International Potato Center; 2024 [cited 2024 Mar 10]. Available from: <https://data.cipotato.org/dataset.xhtml?persistentId=doi:10.21223/DOWPZH> doi:10.21223/DOWPZH

42. Loayza H, Pujaico F. Late blight severity estimation using a Kernel Ridge Regression algorithm with K-means clustering [Python] [Internet]. 2025. Available from: <https://gitlab.com/PrecisionAgriculture/RegressionSeverityEstimation>
43. Steinley Douglas. K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*. 2006;59:1–34. doi:10.1348/000711005X48266
44. Murphy KP. *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press; 2012. 1104 p. (Bach F, editor. *Adaptive Computation and Machine Learning series*).
45. Shawe-Taylor J, Cristianini N. *Kernel Methods for Pattern Analysis*. Cambridge University Press; 2004. 520 p.
46. Van Rossum BJ. Biometris/statgenSTA [R] [Internet]. Wageningen Universitij & Research, Biometris; 2024 [cited 2024 Dec 18]. Available from: <https://github.com/Biometris/statgenSTA>
47. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing [Software]. [Internet]. 2022. Available from: <http://www.R-project.org/>
48. Fernández CI, Leblon B, Haddadi A, Wang J, Wang K. Potato Late Blight Detection at the Leaf and Canopy Level Using Hyperspectral Data. *Canadian Journal of Remote Sensing*. 2020;46:390–413. doi:10.1080/07038992.2020.1769471
49. Zhang M, Liu X, O’Neill M. Spectral discrimination of *Phytophthora infestans* infection on tomatoes based on principal component and cluster analyses. *International Journal of Remote Sensing*. 2002;23:1095–107. doi:10.1080/01431160110106078
50. Zhang M, Qin Z, Liu X, Ustin SL. Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing. *International Journal of Applied Earth Observation and Geoinformation*. 2003;4:295–310. doi:10.1016/S0303-2434(03)00008-4
51. Hipsch M, Michael Y, Lampl N, Sapir O, Cohen Y, Helman D, et al. Early detection of late blight in potato by whole-plant redox imaging. *The Plant Journal*. 2023;113:649–64. doi:10.1111/tpj.16071
52. Huitema E, Bos JIB, Tian M, Win J, Waugh ME, Kamoun S. Linking sequence to phenotype in *Phytophthora*–plant interactions. *Trends in Microbiology*. 2004;12:193–200. doi:10.1016/j.tim.2004.02.008
53. Kamoun S, Smart CD. Late Blight of Potato and Tomato in the Genomics Era. *Plant Dis*. 2005;89:692–9. doi:10.1094/PD-89-0692 PubMed PMID: 30791237.
54. Abramovitch RB, Martin GB. Strategies used by bacterial pathogens to suppress plant defenses. *Current Opinion in Plant Biology*. 2004;7:356–64. doi:10.1016/j.pbi.2004.05.002
55. Kelley BS, Lee SJ, Damasceno CMB, Chakravarthy S, Kim BD, Martin GB, et al. A secreted effector protein (SNE1) from *Phytophthora infestans* is a broadly acting suppressor of programmed cell death. *The Plant Journal*. 2010;62:357–66. doi:10.1111/j.1365-313X.2010.04160.x

- 656 56. Oh SK, Young C, Lee M, Oliva R, Bozkurt TO, Cano LM, et al. In Planta Expression Screens of
657 Phytophthora infestans RXLR Effectors Reveal Diverse Phenotypes, Including Activation of the
658 Solanum bulbocastanum Disease Resistance Protein Rpi-blb2. The Plant Cell. 2009;21:2928–47.
659 doi:10.1105/tpc.109.068247
- 660 57. Van Damme M, Bozkurt TO, Cakir C, Schornack S, Sklenar J, Jones AME, et al. The Irish potato famine
661 pathogen Phytophthora infestans translocates the CRN8 kinase into host plant cells. PLoS Pathog.
662 2012;8:e1002875. doi:10.1371/journal.ppat.1002875 PubMed PMID: 22927814; PubMed Central
663 PMCID: PMC3426532.
- 664 58. Jacquemoud S, Ustin SL, Verdebout J, Schmuck G, Andreoli G, Hosgood B. Estimating leaf
665 biochemistry using the PROSPECT leaf optical properties model. Remote Sensing of Environment.
666 1996;56:3. doi:10.1016/0034-4257(95)00238-3
- 667 59. Liu L yun, Huang W jiang, Pu R liang, Wang J hua. Detection of Internal Leaf Structure Deterioration
668 Using a New Spectral Ratio Index in the Near-Infrared Shoulder Region. Journal of Integrative
669 Agriculture. 2014;13:760–9. doi:10.1016/S2095-3119(13)60385-8
- 670 60. Datt B. A New Reflectance Index for Remote Sensing of Chlorophyll Content in Higher Plants: Tests
671 using Eucalyptus Leaves. Journal of Plant Physiology. 1999;154:30–6. doi:10.1016/S0176-
672 1617(99)80314-9
- 673 61. Fuente CPD la, Valdés-Gómez H, Roudet J, Verdugo-Vásquez N, Mirabal Y, Laurie VF, et al. Vigor
674 thresholded NDVI is a key early risk indicator of Botrytis bunch rot in vineyards. OENO One.
675 2020;54:2. doi:10.20870/oeno-one.2020.54.2.2954
- 676 62. Mahlein AK, Rumpf T, Welke P, Dehne HW, Plümer L, Steiner U, et al. Development of spectral
677 indices for detecting and identifying plant diseases. Remote Sensing of Environment. 2013;128:21–
678 30. doi:10.1016/j.rse.2012.09.019
- 679 63. Zhao H, Yang C, Guo W, Zhang L, Zhang D. Automatic Estimation of Crop Disease Severity Levels
680 Based on Vegetation Index Normalization. Remote Sensing 2020, 12, 1930. Remote Sensing.
681 2020;12:22. doi:10.3390/rs12223761
- 682 64. Shi Y, Huang W, Ye H, Ruan C, Xing N, Geng Y, et al. Partial Least Square Discriminant Analysis Based
683 on Normalized Two-Stage Vegetation Indices for Mapping Damage from Rice Diseases Using
684 PlanetScope Datasets. Sensors. 2018;18:1901. doi:10.3390/s18061901
- 685 65. Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. Machine Learning in Agriculture: A Review.
686 Sensors. 2018;18. doi:10.3390/s18082674
- 687 66. Bishop CM. Pattern Recognition and Machine Learning [Internet]. New York: Springer; 2006 [cited
688 2026 Mar 9]. Available from: <https://link.springer.com/book/9780387310732>
- 689 67. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning [Internet]. New York, NY:
690 Springer; 2009 [cited 2026 Mar 9]. (Springer Series in Statistics). Available from:
691 <http://link.springer.com/10.1007/978-0-387-84858-7> doi:10.1007/978-0-387-84858-7

68. Shi Y, Han L, Kleerekoper A, Chang S, Hu T. Novel CropdocNet Model for Automated Potato Late Blight Disease Detection from Unmanned Aerial Vehicle-Based Hyperspectral Imagery. *Remote Sensing*. 2022;14. doi:10.3390/rs14020396
69. Sun H, Song X, Guo W, Guo M, Mao Y, Yang G, et al. Potato late blight severity monitoring based on the relief-mRmR algorithm with dual-drone cooperation. *Computers and Electronics in Agriculture*. 2023;215:108438. doi:10.1016/j.compag.2023.108438
70. Araus JL, Cairns JE. Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*. 2014;19:52–61. doi:10.1016/j.tplants.2013.09.008 PubMed PMID: 24139902.
71. Sankaran S, Khot LR, Espinoza CZ, Jarolmasjed S, Sathuvalli VR, Vandemark GJ, et al. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. *European Journal of Agronomy*. 2015;70:112–23. doi:10.1016/j.eja.2015.07.004
72. Zhang C, Kovacs JM. The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agric*. 2012;13:693–712. doi:10.1007/s11119-012-9274-5
73. Araus JL, Kefauver SC, Zaman-Allah M, Olsen MS, Cairns JE. Translating High-Throughput Phenotyping into Genetic Gain. *Trends in Plant Science*. 2018;23:451–66. doi:10.1016/j.tplants.2018.02.001
74. Mahlein AK. Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. *Plant Disease*. 2016;100:241–51. doi:10.1094/PDIS-03-15-0340-FE
75. Crain J, Mondal S, Rutkoski J, Singh RP, Poland J. Combining High-Throughput Phenotyping and Genomic Information to Increase Prediction and Selection Accuracy in Wheat Breeding. *The Plant Genome*. 2018;11:170043. doi:10.3835/plantgenome2017.05.0043
76. Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de los Campos G, et al. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends in Plant Science*. 2017;22:961–75. doi:10.1016/j.tplants.2017.08.011

Supporting information

S1 Fig. Weather conditions during potato trials and UAV acquisitions. Weather conditions in the study area of Oxapampa, Peru (September 2021–February 2022). Minimum (blue lines) and maximum (red lines) temperatures are on the left y-axis, while precipitation (PP, cyan bars) and relative humidity (RH, black dotted lines) are on the right. Vertical green and orange dashed lines

mark planting and harvest for the first (A) and second (B) trials. Red arrows indicate the two unmanned aerial vehicle flight dates for multispectral imagery acquisition.

S2 Fig. AUDPC vs. NDVI- and KRR-based LB severity: implications for selection. Best linear unbiased estimators (BLUEs) of the area under the disease progress curve (AUDPC) versus BLUEs of estimated late blight (LB) severity in large potato breeding trials. LB severity was estimated by two approaches: normalized difference vegetation index-based linear regression (LB_{LR} ; panels (A) and (B)) and kernel ridge regression-based machine-learning models (LB_{KRR} ; panels (C) and (D)). Panels (A) and (C) show results for 2,745 clones (Trial A) evaluated at 64 days after planting (DAP), whereas panels (B) and (D) show 492 accessions (Trial B) evaluated at 38 DAP. Red symbols (with A=Amarilis, K=Kori, and Y=Yungay common to both trials) and green symbols represent check and parent varieties, respectively. SC10, SC20, and SC30 denote the percentage of selected coincidence between AUDPC- and LB-based rankings within the top 10%, 20%, and 30% of clones ranked by AUDPC.

S3 Fig. Density distributions of AUDPC and LB severity estimates. Kernel density distributions of standardized values (z-scores) of best linear unbiased estimators (BLUEs) of the area under the disease progress curve of late blight (AUDPC; light blue) and the BLUEs of late blight severity estimated using normalized difference vegetation index-based linear regression (LB_{LR} ; green) and kernel ridge regression-based machine-learning models (LB_{KRR} ; pink). Panels (A) and (C) show results for 2,745 clones (Trial A) evaluated at 64 and 86 Days After Planting (DAP), and panels (B) and (D) for 492 accessions (Trial B) evaluated at 38 and 58 DAP.