Decision Trees to Forecast Risks of Strawberry Powdery Mildew Caused by *Podosphaera aphanis*

Odile Carisse * and Mamadou Lamine Fall

Saint-Jean-sur-Richelieu Research and Development Centre, Agriculture and Agri-Food Canada, Saint-Jean-sur-Richelieu, QC J3B 3E6, Canada; Mamadou.lamine.fall@canada.ca
* Correspondence: odile.carisse@canada.ca; Tel.: +1-579-224-3086

Abstract: Powdery mildew (*Podosphaera aphanis*) is a major disease in day-neutral strawberry. Up to 30% yield losses have been observed in Eastern Canada. Currently, management of powdery mildew is mostly based on fungicide applications without consideration of risk. The objective of this study is to use *P. aphanis* inoculum, host ontogenic resistance, and weather predictors to forecast the risk of strawberry powdery mildew using CART models (classification trees). The data used to build the trees were collected in 2006, 2007, and 2008 at one experimental farm and six commercial farms located in two main strawberry-production areas, while external validation data were collected at the same experimental farm in 2015, 2016, and 2018. Data on proportion of leaf area diseased (PLAD) were grouped into four severity classes (1: PLAD = 0; 2: PLAD > 0 and <5%; 3: >5% and <15%; and 4: PLAD > 15%) for a total of 681 and 136 cases for training and external validation, respectively. From the initial 92 weather variables, 21 were selected following clustering. The tree with the best balance between the number of predictors and highest accuracy was built with: airborne inoculum concentration and number of susceptible leaves on the day of sampling, and mean relative humidity, mean daily number of hours at temperature between 18 and 30 °C, and mean daily number of hours at saturation vapor pressure between 10 and 25 mmHg during the previous 6 days. For training, internal validation, and external validation datasets, the sensitivity, specificity, and accuracy ranged from 0.70 to 0.90, 0.87 to 0.98, and 0.82 to 0.97, respectively. The classification rules to estimate strawberry powdery mildew risk can be easily implemented into disease decision support systems and used to treat only when necessary and thus avoid preventable yield losses and unnecessary treatments.

Keywords: data mining; disease decision support system; disease risk; epidemiology; machine learning

1. Introduction

In Canada, the province of Quebec is the most important strawberry producer, with 14,117 MT produced from 1921 ha of strawberry plantings, representing 57% of the Canadian production in 2018 [1]. As a response to consumer demand for longer period of availability and for high-quality locally produced strawberries, the production has evolved. Until the end of 1980s, almost exclusively short-day (June bearing) varieties were grown. Strawberries were harvested during a 3 to 4-week period starting mid- to late June. This production system was gradually replaced by new production systems such as day-neutral cultivars, winter row covers, and production in tunnels. The combination of these techniques allows for a better distribution of the strawberry harvest throughout the growing season. However, with these new techniques came new challenges, one of them being the emergence of strawberry powdery mildew (SPM) in the early 2000s [2,3].

The disease, caused by the ascomycete *Podosphaera aphanis* (Wallr.), can affect all aerial parts of the strawberry plant, including leaves, stolons, flowers, and berries [4]. First signs of powdery mildew are usually white patches of mycelium and conidia on the abaxial leaf surface [5]. As the disease progresses, patches can cover entire leaves and lead to

defoliation [5]. Foliar infections cause indirect losses by reducing photosynthesis, but also act as a source of inoculum for flower and fruit infections, which are responsible for direct yield loss [3,6]. Infected flowers may produce less pollen, wilt, or die, whereas infected fruits may fail to ripen, remain soft or have a shortened shelf life [7–9]. Yield losses, albeit variable, can reach up to 30% [3].

Powdery mildew development is influenced by factors such as weather conditions, cultivar susceptibility, phenological stage of the crop, and amount of inoculum. Temperatures between 18 and 25 °C are optimal for conidial germination, lesion growth, and sporulation; high relative humidity is required for conidial germination, and rainfall has a detrimental effect on conidia germination [10–14]. Although no cultivar is completely resistant to powdery mildew, disease susceptibility varies among cultivars. Jewel and Seascape, two of the most popular cultivars in the province of Quebec, are susceptible to powdery mildew [15]. Susceptibility of leaves and berries to powdery mildew decreases exponentially with age [6,16]. Thus, flowers, green berries, and not-fully expanded leaves are highly susceptible, whereas pink and mature berries and fully expanded leaves are almost resistant [6,16]. In open fields, the disease is more problematic for day-neutral cultivars than for June-bearing ones. Indeed, day-neutral cultivars bear susceptible flowers and fruits from mid-July to September, when powdery mildew epidemics can be severe, whereas June-bearing cultivars bear susceptible flowers and fruits in late spring and beginning of summer, which is generally before the disease onset [2,3]. Powdery mildew is also known to be more problematic under tunnels because of favorable weather conditions for its development [17]. Actual management of SPM is mostly based on fungicide applications. Fungicides are sprayed regularly, either from the emergence of the first leaves or from the first symptoms to harvest [2,3]. However, fungicide resistance in *P. aphanis* populations has been reported [18,19].

Foliar severity was found to be a good predictor of subsequent crop losses for day-neutral strawberry in open field; 5% crop loss was reached when an average of 17% leaf area diseased was observed since the first symptoms [3]. However, assessing disease severity on leaves is difficult and time-consuming. To facilitate disease assessment, Carisse et al. [2] developed models allowing foliar severity to be estimated from disease incidence for different types of strawberry productions. Although disease severity and incidence are good predictors of crop losses, predicting disease development is of major interest. The use of airborne conidia concentration and weather conditions to predict disease development has been studied. A positive correlation between disease severity and daily airborne conidia concentration was reported for several diseases [20]. For these diseases, it is possible to use airborne inoculum concentration as disease risk indicator and to time fungicide applications [21]. The progression of SPM is also influenced by weather conditions, which can be used to estimate the disease risk. However, the difficulty of this approach lies in the wide range of temperatures and relative humidity under which SPM can develop [5,12,13,22]. Despite this difficulty, tools for estimating the risk of SPM were developed. In the United Kingdom, a prediction system that calculates the percentage of completion of a disease cycle based on temperature, relative humidity, and leaf wetness was proposed by Dodgson et al. [23]. Once a disease cycle is completed, the system predicts a high-risk day, which can guide fungicide applications [23]. In Italy, the SafeBerry diseasedecision-support system (DDSS) calculates a disease risk index which combines a basic risk index from cultivar susceptibility, presence of overhead irrigation, plant density, etc., and a daily risk index from phenological stage of the crop, disease incidence, time of disease onset, etc. [24]. The disease risk index is combined with the suitability of the temperature to make a recommendation of action (e.g., do not spray today) [24]. Unfortunately, the rules are not available and hence it would be difficult to adapt this DSS to specific conditions. In France, a predictive model calculates the risk of powdery mildew based on suitable weather conditions for the different phases of *P. aphanis* life cycle [25]. Finally, in the United States, a risk index developed at first for grape powdery mildew (*Erysiphe necator*) was adapted for SPM [26]. Based on the

index value, a fungicide and an interval between treatments are suggested [26]. With the exception of the DDSS developed by Hoffman and Gubler [26], the decision rules are not available and the DDSS were developed mostly for strawberry production in large tunnels. The DSS developed by Hoffman and Gubler [26] was evaluated under the conditions of eastern Canada in both June-bearing (Jewel) and day-neutral strawberry production (Seascape) [15]. For the two years of evaluation, the DDSS did not allow for a reduction in the number of fungicide treatments [15].

Risk assessment generally relies on variable-centered statistical techniques such as logistic, linear, or multiple regression to model the relationship between a group of independent variables, often weather variables, and disease risk or severity [27,28]. Unless specified within the model, these methods generally do not allow for consideration of complex interactions among the independent variables or for different assemblages of variables that may lead to disease development. Pattern-centered statistical techniques provide a different approach to determine disease risk by identifying the subgroups within a number of disease observations that share similar characteristics, allowing for the identification of patterns of risk factors. Methods of segmenting disease risk observations into subgroups, such as tree-based models, can be used to derive prediction rules for determining disease risk classes. As opposed to linear and additive models, pattern-centered approaches are nonparametric and relatively easy to implement and interpret. Also, algorithms such as CART models may reveal interactions among variables generally ignored by other modeling approaches [29,30]. The objective of this work was to use weather predictors (variables) alone, or combined with inoculum and with host susceptibility predictors, to forecast the risk of SPM expressed as classes (no, low, moderate and severe risk) using decision tree classification procedures. These classes can be used as early warning, warnings, and action thresholds for strawberry mildew management actions [3].

2. Materials and Methods

2.1. Description of the Sampling Sites

The data used in this study were described in Carisse et al. [3]. Briefly, data were collected at the Agriculture and Agri-Food Canada experimental farm located in Frelighsburg, Quebec (latitude 45°03′12″ N; longitude 72°51′42″ W), and in commercial strawberry plantings located in Saint-Paul-d'Abbotsford (latitude 45°25′60″ N; longitude 72°52′60″ W) and in Île d'Orléans (latitude 46°55′06″ N; longitude 70°58′35″ W). For the purpose of the study, only data from non-sprayed plots were used. The data were collected at three, five, and four sites in 2006, 2007, and 2008, respectively, for a total of 12 epidemics. At the site of Frelighsburg (experimental farm), plots were also established in 2015, 2016, and 2018. At all sites, data were collected in plots planted with the day-neutral strawberry cultivar Seascape, which was planted during the last 2 weeks of May in raised beds set 1.4 m apart, with two rows of strawberry plants per bed covered with black polyethylene mulch. On each bed, strawberry plants were spaced 30 cm apart. At the sites in Frelighsburg, Saint-Paul-d'Abbotsford and Île d'Orléans, plots comprised of 13 raised beds of 15 m long, 15 raised beds of 30 m long, and 13 raised beds of 12 m long, respectively. At all sites, flower trusses were removed until mid-June.

2.2. Disease, Host, Inoculum, and Weather Monitoring

At each site, the same 25 plants, randomly selected at the first sampling date, were assessed for powdery mildew severity. Severity was assessed every 2 days from the first week of June to the first week of October as percent leaf area diseased (PLAD) on the three youngest fully expanded leaves. In 2015, 2016, and 2018 disease severity was assessed twice weekly. Severity was estimated using a diagrammatic scale with 5% steps (0%, 5%, 10%, 15% . . . 100%). Host susceptibility was assessed based on the number of susceptible leaves per 1 m of row. A leaf was considered susceptible if leaflets were not completely unfolded [6]. Airborne conidia concentration was assessed every 2 days using two rotating-arm impaction spore samplers placed in the central row, with one at 5, 10,

and 4 m upward and the other one at 5, 10, and 4 m downward from the middle of the plots at the Frelighsburg, Saint-Paul-d'Abbotsford, and Île d'Orléans sites, respectively. In 2015, 2016, and 2018 airborne conidia concentration was assessed twice weekly. At all sites, the samplers ran for 20 min every hour (10 min on and 20 min off) from 8:00 a.m. to 8:00 p.m. The number of P. aphanis conidia on the sampling surface was counted under a microscope at ×250 magnification. Conidia of P. aphanis were identified based on their size (20–23 × 13–20 μm), their barrel shape when turgid, and the presence of granules inside the conidia [31]. The concentration of airborne conidia (ACC) was expressed as the number of conidia per cubic meter of air (mean over the two samplers). Weather data were monitored using automatic weather stations (CR-21X; Campbell Scientific Inc., Edmonton, AB, Canada) placed in an unobstructed area 3 m from the plot edge. Data were measured every 15 min and saved as hourly averages or totals (rain). Temperature and relative-humidity probes were positioned in a white shelter at 1.5 m above the ground. Rainfall was recorded using a tipping bucket rain gauge (Geneq, Montreal, QC, Canada) at 50 cm above the ground.

2.3. Description of the Response Variable and Classification Trees Predictors

The severity of strawberry powdery mildew, expressed as percent leaf area diseased (PLAD, mean over the 25 plants assessed, three leaves per plant), was transformed into severity classes to generate ordinal response variable as: Class 1: PLAD = 0; Class 2: PLAD > 0 and <5%; Class 3: PLAD > 5% and <15%; and Class 4: PLAD > 15% and considered as the dependent variable. The interval used to create the severity classes was selected based on the reported relationship between leaf disease severity and yield losses [3,32] and distribution of cases within each class. The classification tree predictor related to airborne inoculum was airborne conidia concentration (ACC), which was log-transformed and expressed as a continuous variable ($log10(ACC + 1)$). The predictor related to strawberry leaf receptivity was the number of susceptible leaves per 1 m row, expressed as a continuous variable (LVS). Weather variables were summarized over a 1- or 6-day period. Air temperature (°C) and relative humidity (%) were expressed as averages, minima, and maxima. Rainfall was expressed as accumulated values (millimeters), or number of hours with rain (>2 mm). For the 6-day period, temperature and humidity were expressed as averages, duration (hours), or sum within pre-set thresholds [5,12,13,22]. The description of the 94 predictors is provided in Table 1. A total of 681 and 136 disease cases, each one corresponding to unique combination of year, site, and sampling day, were generated from the first and last three years of the study, respectively. Classification tree is a supervised learning method, hence bootstrap validation (internal validation) method [33] was used to validate the models (trees), and the 136 independent cases collected in 2015, 2016, and 2018 were used to evaluate their prediction accuracy (external validation).

Table 1. Description of the inoculum, host, and weather predictors.

Description	Unit	Abbreviation [a]
Airborne conidia concentration (ACC)	Log_{10} conidia/m^3+1	$Log_{10}(ACC+1)$
Number of susceptible leaves	Number of leaves per 1 m of row	LVS
Daily mean, minimum and maximum temperature	°C	**T**, Tmin, Tmax
Day mean, minimum, and maximum temperature	°C	**DT**, DTmin, DTmax
Night mean, minimum, and maximum temperature	°C	**NT**, NTmin, NTmax
Average mean, minimum, and maximum temperature during the previous 6 days	°C	**6T**, 6Tmin, 6Tmax
Daily mean, minimum and maximum relative humidity	% relative humidity	RH, RHmin, RHmax
Day mean, minimum, and maximum relative humidity	% relative humidity	DRH, DRHmin, DRHmax
Night mean, minimum, and maximum relative humidity	% relative humidity	**NRH**, NRHmin, NRHmax

Table 1. *Cont.*

Description	Unit	Abbreviation [a]
Average mean, minimum, and maximum relative humidity during the previous 6 days	% relative humidity	**6RH**, 6RHmin, **6RHmax**
Daily, day, night and 6-day average total rainfall	mm	**RAIN**, DRAIN, NRAIN, 6RAIN
Daily, day, night and 6-day average number of rainy hours	hours	RAINH, DRAINH, NRAINH, **6RAINH**
Daily, day, night and 6-day average number of hours with temperature below 5 °C	hours	T < 5, DT < 5, NT < 5, 6T < 5
Daily, day, night and 6-day average number of hours with temperature below 13 °C	hours	T < 13, DT < 13, NT < 13, **6T < 13**
Daily, day, night and 6-day average number of hours with temperature above 30 °C	hours	T > 30, DT > 30, NT > 30, 6T > 30
Daily, day, night and 6-day average number of hours with temperature above 35 °C	hours	T > 35, DT > 35, NT > 35, 6T > 35
Daily, day, night and 6-day average number of hours with temperature between 15 and 25 °C	hours	T15–25, DT15–25, **NT15–25**, **6T15–25**
Daily, day, night and 6-day average number of hours with temperature between 18 and 25 °C	hours	T18–25, DT18–25, NT18–25, **6T18–25**
Daily, day, night and 6-day average number of hours with temperature between 18 and 30 °C	hours	T18–30, DT18–30, NT18–30, 6T18–30
Daily, day, night and 6-day average number of hours with relative humidity above 95%	hours	RH > 95, DRH > 95, NRH > 95, 6RH > 95
Daily, day, night and 6-day average number of hours with relative humidity between 70% and 85%	hours	RH70–85, DRH70–85, NRH70–85, 6RH70–85
Daily, day, night and 6-day average number of hours with relative humidity between 70% and 95%	hours	RH70–95, DRH70–95, NRH70–95, 6RH70–95
Daily, day, night and 6-day average mean saturation vapor pressure	Mm Hg	**VP**, **DVP**, NVP, 6VP
Daily, day, night and 6-day average number of hours with saturation vapor pressure below 5 Mm Hg	hours	VP < 5, DVP < 5, NVP < 5, **6VP < 5**
Daily, day, night and 6-day average number of hours with vapor pressure above 10 Mm Hg	hours	VP > 10, DVP > 10, NVP > 10, 6VP > 10
Daily, day, night and 6-day average number of hours with vapor pressure between 10 and 25 Mm Hg	hours	VP10–25, DVP10–25, NVP10–25, **6VP10–25**
Daily, day, night and 6-day average number of hours with vapor pressure 15 and 25 Mm Hg	hours	VP15–25, DVP15–25, NVP15–25, **6VP15–25**

[a] variable in bold are those selected following clustering, intra-cluster correlation, and discriminant analysis.

2.4. Development of Decision Trees

We used the decision tree classification technique, a supervised learning method described in detail in several text books [34,35]. Briefly, the decision tree is the result of an algorithm which produces a model consisting of a set of classification rules that are represented as a tree [36]. The tree is built following a top-down approach and consists of root nodes, which are split into more branches. In a tree, each node represents a value of a predictor, and each branch descending from the node corresponds to one of the possible values that the predictor can take. The decision tree was developed following several steps: reducing redundancy; defining classification rules, constructing the decision trees, implementing the decision trees, evaluating the classification results for both training and

validation data sets, and determining the reliability of the trees with independent data (2015, 2016, and 2018 data).

First, to facilitate segmentation while developing the decision trees, redundancy among the 94 weather-based predictors (independent variables) was reduced using clustering [37]. Clustering was used to identify the groups (clusters) of highly correlated predictors, with the smallest correlation between groups. Clustering was conducted using the VARCLUS procedure in SAS with an eigenvalue threshold of 0.7 (SAS PROC VAR-CLUS) [38]. Clustering results were used to select groups of weather-based predictors within a cluster using two approaches. First, within each cluster, Spearman's rank-based correlation coefficients were computed between weather-based predictors and used to identify highly correlated predictors ($r > 0.95$). Second, a discriminant analysis was performed to identify the predictors that influenced the categories of strawberry powdery mildew severity [38]. All predictors were given an equal weight.

In this study, the CART algorithm was used to build the decision trees [39]. CART uses binary recursive partitioning to systematically identify the best predictor among all predictors that splits the dataset into the best low- and high-risk groups with respect to the classes of severity of SPM [34,39]. During the process, for each variable all possible separations are evaluated to determine which splits are the best at predicting the classes of severity of SPM. Using this procedure, a decision tree was built using the tree function in DTREG (version 10.9.1). The tree starts with parent nodes, which were further split into child nodes based on the next best variable and split criteria. Splitting of the child nodes was continued until within-node deviance was ≤0.01 of the root node [34]. The minimal number of cases in a terminal node was set to 6 [40], and the GINI criterion was used to determine the best split at each node. A 10-fold cross-validation procedure was used to determine the optimal tree by randomly splitting the learning dataset into 10 subsets of data and then repeating the tree-building process ten times [41]. The trees were then pruned to fewer nodes by removing the least-important splits based on the misclassification cost. Because of the differential availability of host, inoculum, and weather data, four decision trees were built using all predictors, weather and inoculum predictors (excluding LVS), weather and host susceptibility predictors (excluding $\log_{10}(ACC + 1)$), and only weather predictors (excluding both $\log_{10}(ACC + 1)$ and LVS).

The performance of the decision trees was assessed based on the reliability of the trees in assessing classes of SPM severity. Within each SPM class, all sampling dates (disease cases) were divided in two groups, with cases and controls defined as the sampling dates associated with SPM severity within the class or not, respectively. Each tree was evaluated for its ability to classify the severity of SPM by recording the number of cases and controls that were correctly classified [27]. For each SPM class and for each tree, true positive (TP) and true negative (TN) were calculated as the number of cases and controls correctly classified by the decision tree, while the false positive (FP) and false negative (FN) were calculated as the number of cases and controls incorrectly classified by the decision tree. The true positive proportion (TPP, sensitivity) was calculated by dividing the number of TP classifications by the total number of cases. The true negative proportion (TNP, specificity) was calculated by dividing the number of TN by the total number of controls. The false positive proportion (FPP) was calculated as 1-TNP, and the false negative proportion (FNP) was calculated as 1-TPP. The overall accuracy was calculated as the proportion of correct classifications (TP+TN) [27]. This procedure was conducted using the training, internal validation, external validation data sets.

3. Results

For the data collected in 2006, 2007, and 2008 used as training and internal validation data, 170, 289, and 222 cases were analyzed respectively, for a total of 681 cases. Among all cases, 163, 144, 177, and 197 cases fell within the SPM severity classes 1, 2, 3, and 4, respectively. For the data collected in 2015, 2016, and 2018 used as external validation data, 44, 46, and 46 cases were analyzed, for a total of 136 cases from which 29, 26, 31,

and 50 cases fell in SPM severity classes 1, 2, 3, and 4, respectively. The distribution of cases for each year, in each class is presented in Figure 1. From the initial set of 92 weather-based predictors (Table 1), 21 were selected following clustering, intra-cluster correlation, and discriminant analysis (Table 1) for building the trees. The tree with the highest accuracy was developed with the following five predictors: airborne inoculum concentration ($\log_{10}$ ACC + 1), number of susceptible leaves (LVS), mean relative humidity during the previous 6 days, mean daily number of hours at temperature between 18 and 30 °C during the previous 6 days, and mean daily number of hours at saturation vapor pressure between 10 and 25 mmHg during the previous six days (h). The distributions of these predictors for each SPM class are presented in Figures 2 and 3. An increase in predictor values with increasing classes of SPM severity was observed for inoculum and number of susceptible leaves (Figure 2A,B). Average $\log_{10}$(ACC + 1) was 0.09, 0.71, 1.48, and 2.38 log number of airborne conidia per m3 +1 for classes 1, 2, 3, and 4, respectively (Figure 2A), while the number of susceptible leaves per meter of row was 0.80, 3.40, 12.39, and 17.94 for classes 1, 2, 3, and 4, respectively (Figure 2B). For the selected weather-based predictors, an increase in the number of hours at temperature between 18 and 30 °C during the preceding 6 days, with increasing class of SPM severity, was observed with 9.25, 11.27, 12.16, and 15.40 h for classes 1, 2, 3, and 4, respectively (Figure 3A). However, such a pattern was not observed for mean relative humidity during the previous 6 days, with 75.06, 77.14, 78.88, and 78.44(%), and for number of hours at saturation vapor pressure between 10 and 25 mmHg during the previous 6 days, with 20.24, 20.69, 21.48, and 21.51 h, for SPM classes 1, 2, 3, and 4, respectively (Figure 3B,C). For the training, internal validation, and external validation data, the sensitivity ranged from 0.68 to 0.90, the specificity ranged from 0.87 to 0.99, and the accuracy ranged from 0.82 to 0.97 (Table 2). The tree built with these predictors is represented in Figure 4.
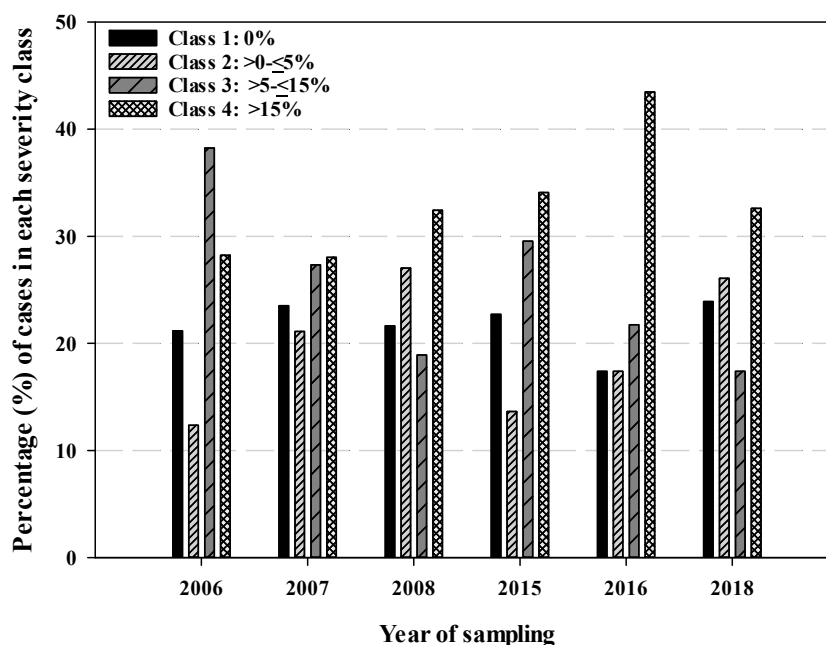
Figure 1. Percentage of cases in each strawberry powdery mildew severity class (1 to 4) for the years of sampling used to develop the trees (2006, 2007, and 2008) and years used for external validation (2015, 2016, and 2018). Disease severity was assessed as the proportion of leaf area diseased on the three youngest fully expanded leaves, mean over 25 plants per site. Strawberry powdery mildew was assessed 170, 289, 222, 44, 46, and 46 times in 2006, 2007, 2008, 2015, 2016, and 2018, respectively.
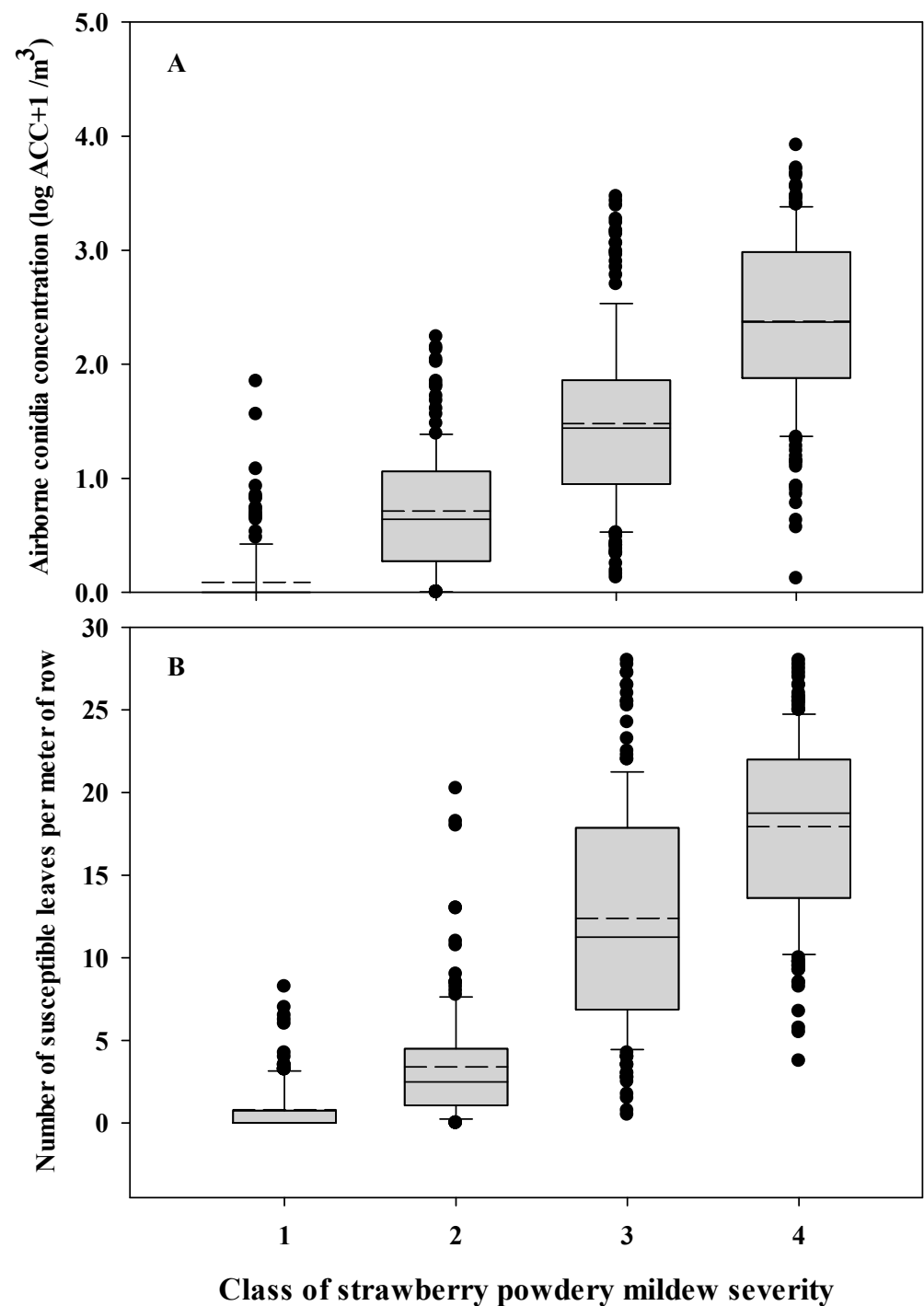
Figure 2. Distribution of airborne inoculum concentrations (**A**) and number of susceptible leaves (**B**) corresponding to the four classes of strawberry powdery mildew. Classes are 1: PLAD = 0; 2: PLAD > 0 and ≤ 5%; 3: >5% and ≤ 15%; and 4: PLAD > 15%, where PLAD is the proportion of leaf area diseased. For each boxplot, from bottom to top, except for the thicker line and whisker caps, each horizontal line marks 25%, 50%, and 75% of data points, respectively. The dashed horizontal line is the mean value and black circles are outliers.
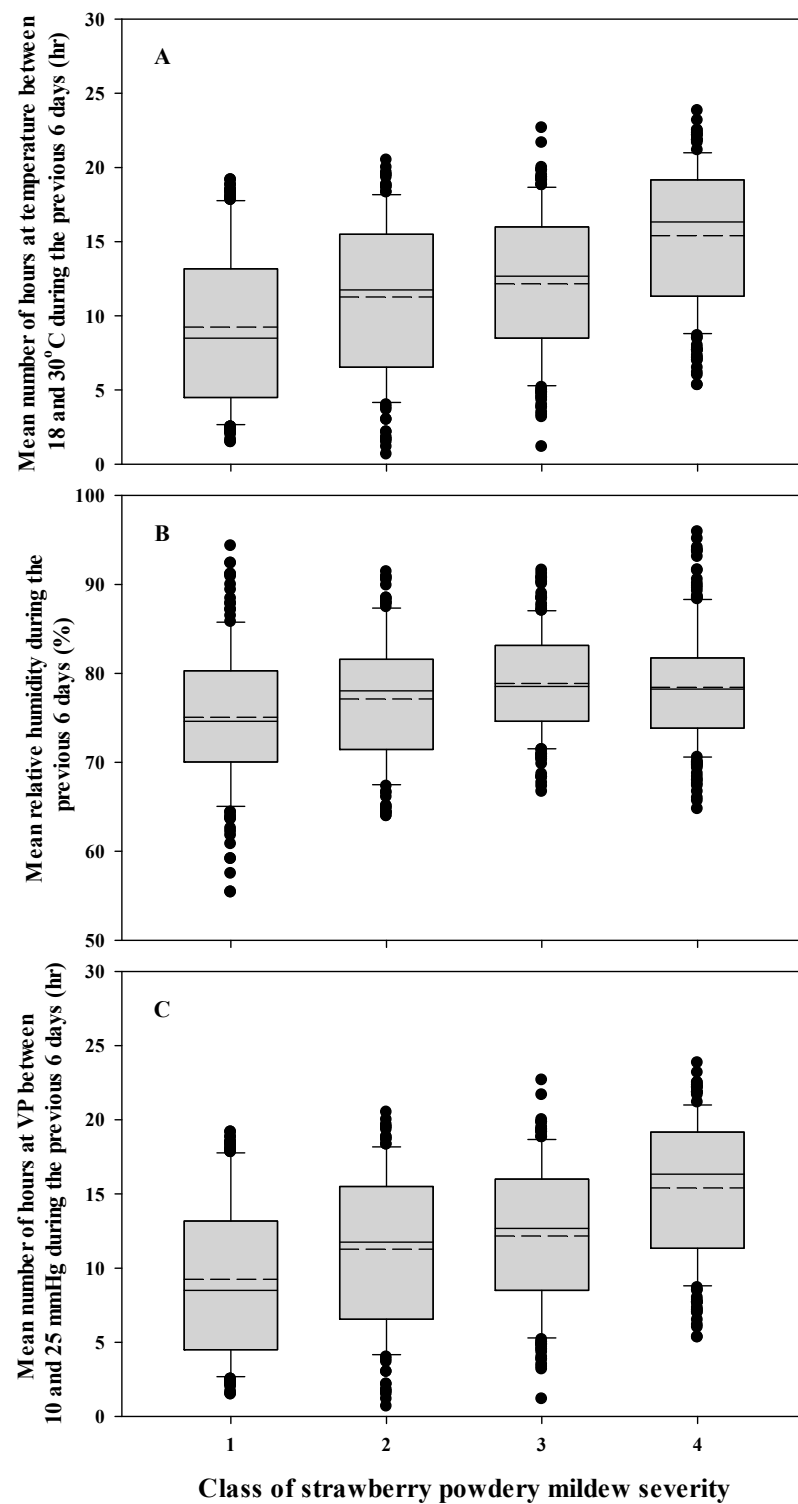
Figure 3. Distribution of the number of hours at temperature between 18 and 30 °C during the previous six days (**A**), mean relative humidity (%) during the previous six days (**B**) and mean number of hours at VP between 10 and 25 mmHg during the previous six days (**C**) for the four classes of strawberry powdery mildew. Classes are 1: PLAD = 0; 2: PLAD > 0 and ≤5%; 3: >5% and ≤15%; and 4: PLAD > 15%, where PLAD is the proportion of leaf area diseased. For each boxplot, from bottom to top, except for the thicker line and whisker caps, each horizontal line marks 25%, 50%, and 75% of data points, respectively. The dashed horizontal line is the mean value and black circles are outliers.

Table 2. Characteristics of the classification trees built with all predictors, weather and inoculum predictors, weather and host predictors, and only weather predictors.

Trees and Predictors Selected [a]	Data Set [b]	Reliability [c]	SPM Class			
			1	2	3	4
All: $\log_{10}(ACC + 1)$, LVS, 6RH, 6T1830, 6VP1025	Training	Sensitivity	0.88	0.74	0.70	0.89
		Specificity	0.97	0.96	0.93	0.90
		Accuracy	0.95	0.91	0.85	0.90
	Internal validation	Sensitivity	0.88	0.72	0.71	0.87
		Specificity	0.98	0.94	0.93	0.92
		Accuracy	0.96	0.90	0.84	0.90
	External validation	Sensitivity	0.90	0.73	0.68	0.80
		Specificity	0.99	0.89	0.87	0.97
		Accuracy	0.97	0.86	0.82	0.90
Weather and Inoculum: $\log_{10}(ACC + 1)$, 6T1825, 6RHMAX, 6VP1025, 6RAINH	Training	Sensitivity	0.88	0.63	0.66	0.79
		Specificity	0.97	0.92	0.93	0.90
		Accuracy	0.95	0.86	0.81	0.87
	Internal validation	Sensitivity	0.88	0.57	0.56	0.78
		Specificity	0.97	0.90	0.93	0.89
		Accuracy	0.95	0.83	0.77	0.86
	External validation	Sensitivity	0.83	0.65	0.71	0.76
		Specificity	0.97	0.91	0.84	0.94
		Accuracy	0.94	0.86	0.81	0.88
Weather and Host: LVS, 6T13, 6RH, 6RAINH, 6T1525, 6RHMAX	Training	Sensitivity	0.80	0.76	0.68	0.89
		Specificity	0.97	0.93	0.93	0.89
		Accuracy	0.93	0.90	0.86	0.89
	Internal validation	Sensitivity	0.74	0.62	0.59	0.80
		Specificity	0.95	0.89	0.93	0.88
		Accuracy	0.90	0.84	0.79	0.86
	External validation	Sensitivity	0.76	0.65	0.61	0.76
		Specificity	0.98	0.88	0.82	0.93
		Accuracy	0.93	0.90	0.86	0.89
Weather: 6T13, 6RH, 6T1525, 6RAINH, 6T1825, 6RHMAX, 6T, 6VP5, 6T1830, NTMIN, 6VP1025, NRH, RAINH, T, NT1525	Training	Sensitivity	0.76	0.72	0.67	0.76
		Specificity	0.93	0.93	0.93	0.90
		Accuracy	0.89	0.88	0.82	0.86
	Internal validation	Sensitivity	0.54	0.57	0.45	0.60
		Specificity	0.90	0.85	0.93	0.85
		Accuracy	0.81	0.79	0.70	0.78
	External validation	Sensitivity	0.69	0.42	0.58	0.68
		Specificity	0.93	0.85	0.81	0.90
		Accuracy	0.88	0.76	0.76	0.82

[a] The trees were built using all, weather and host, weather and inoculum, and weather only predictors. The predictors listed are those selected to build the best tree; [b] data collected in 2006, 2007, and 2008 at one experimental farm and at six commercial farms for a total of 681 cases were used to build the trees, bootstrap validation (internal validation) method [33] was used to validate the models (trees), and the 136 independent cases collected in 2015, 2016, and 2018 were used to evaluate their prediction reliability (external validation); [c] the sensitivity was calculated as the true positive proportion (TPP = number of true positive classification/total number of cases). The specificity was calculated as the true negative proportion (TNP = number of true negatives/total number of controls). The overall accuracy was calculated as the proportion of correct classifications (TP + TN) [27].

When the classification tree was built with inoculum- and weather-based predictors (excluding LVS), five predictors were retained (log10(ACC + 1), 6T1825, 6RHMAX, 6VP1025, 6RAINH). For the training, internal validation, and external validation data, the sensitivity ranged from 0.56 to 0.88, the specificity ranged from 0.84 to 0.97 and the accuracy ranged from 0.81 to 0.95 (Table 2). The tree built with these predictors is represented in Figure 5. When the classification tree was built with host- and weather-based predictors (excluding log10(ACC + 1)), six predictors were retained (LVS, 6T13, 6RH, 6RAINH, 6T1525, and 6RHMAX). The tree built with these predictors is represented in Figure 6. For the training, internal validation, and external validation data, the sensitivity ranged from 0.59 to 0.89, the specificity ranged from 0.82 to 0.98 and the accuracy ranged from 0.79 to 0.93

(Table 2). When only weather-based predictors were used (excluding both log10(ACC + 1) and LVS) to build the classification tree, 15 predictors were used (6T13, 6RH, 6T1525, 6RAINH, 6T1825, 6RHMAX, 6T, 6VP5, 6T1830, NTMIN, 6VP1025, NRH, RAINH, T, and NT1525). For the training, internal validation, and external validation data, the sensitivity ranged from 0.42 to 0.76, the specificity ranged from 0.81 to 0.93 and the accuracy ranged from 0.81 to 0.89 (Table 2).
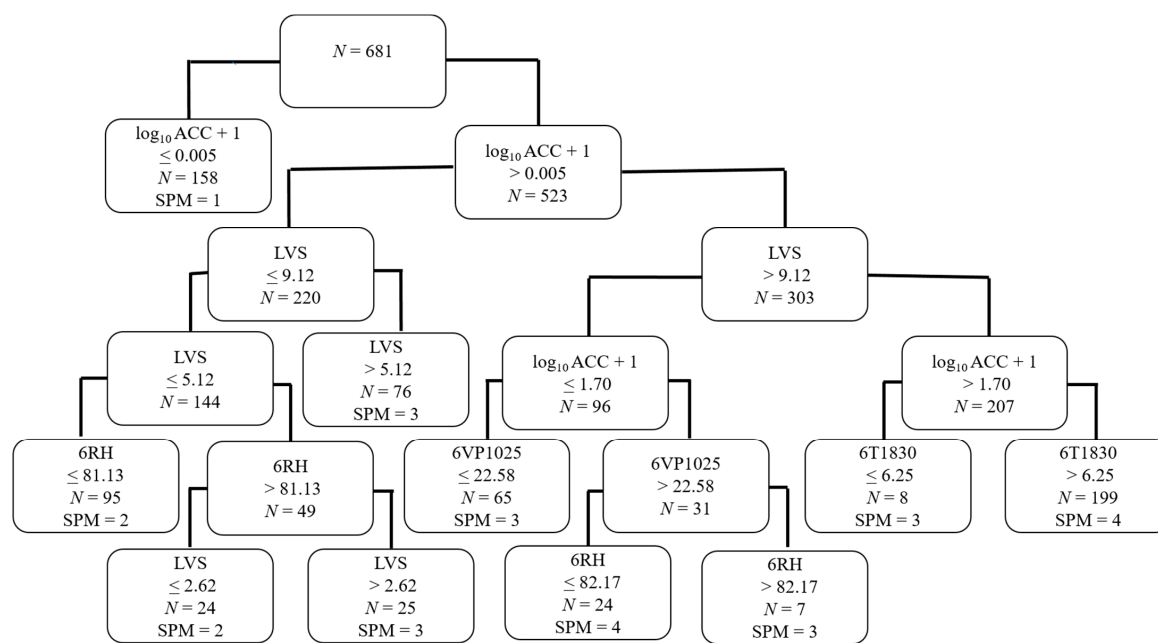
Figure 4. Representation of the classification tree built with airborne inoculum concentration ($\log_{10}$ACC + 1), number of susceptible leaves (LVS), mean % relative humidity during the previous six days (6RH), mean number of hours at temperature between 18 and 30 °C during the previous six days (6T1830), and mean number of hours at saturation vapor pressure between 10 and 25 mmHg during the previous six days (6VP1025).
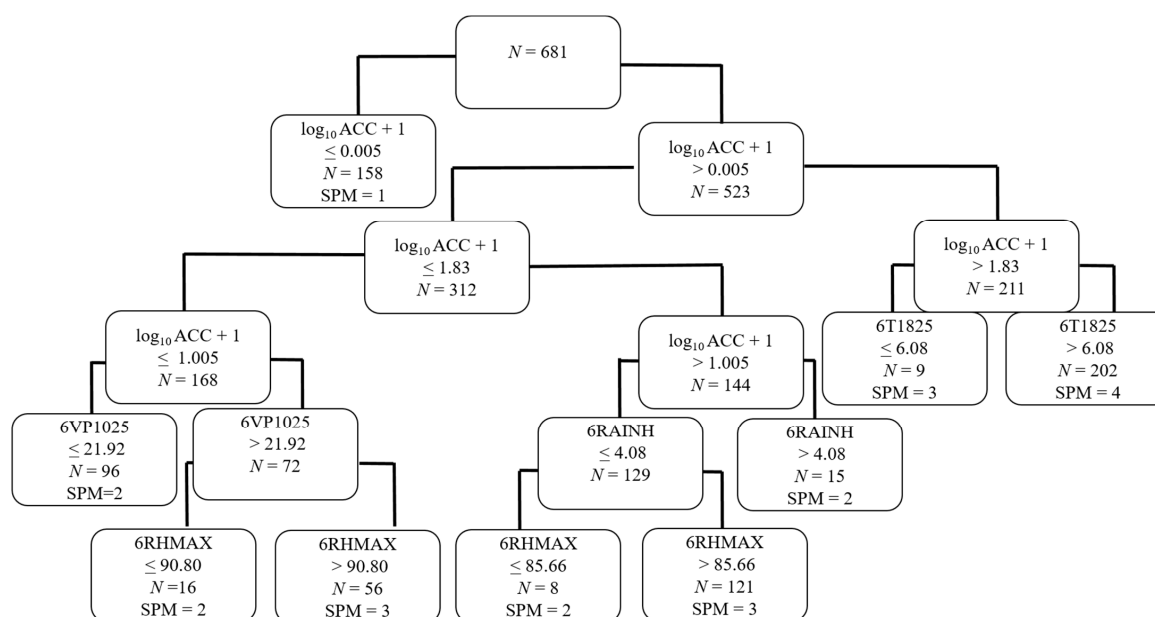
Figure 5. Representation of the classification tree built with airborne inoculum concentration ($\log_{10}$ACC + 1), mean number of hours at temperature between 18 and 25 °C during the previous six days (6T1825), mean maximum % relative humidity during the previous six days (6RHMAX), mean number of hours at saturation vapor pressure between 10 and 25 mmHg during the previous six days (6VP1025), and mean number of rainy hours during the previous six days (6RAINH).
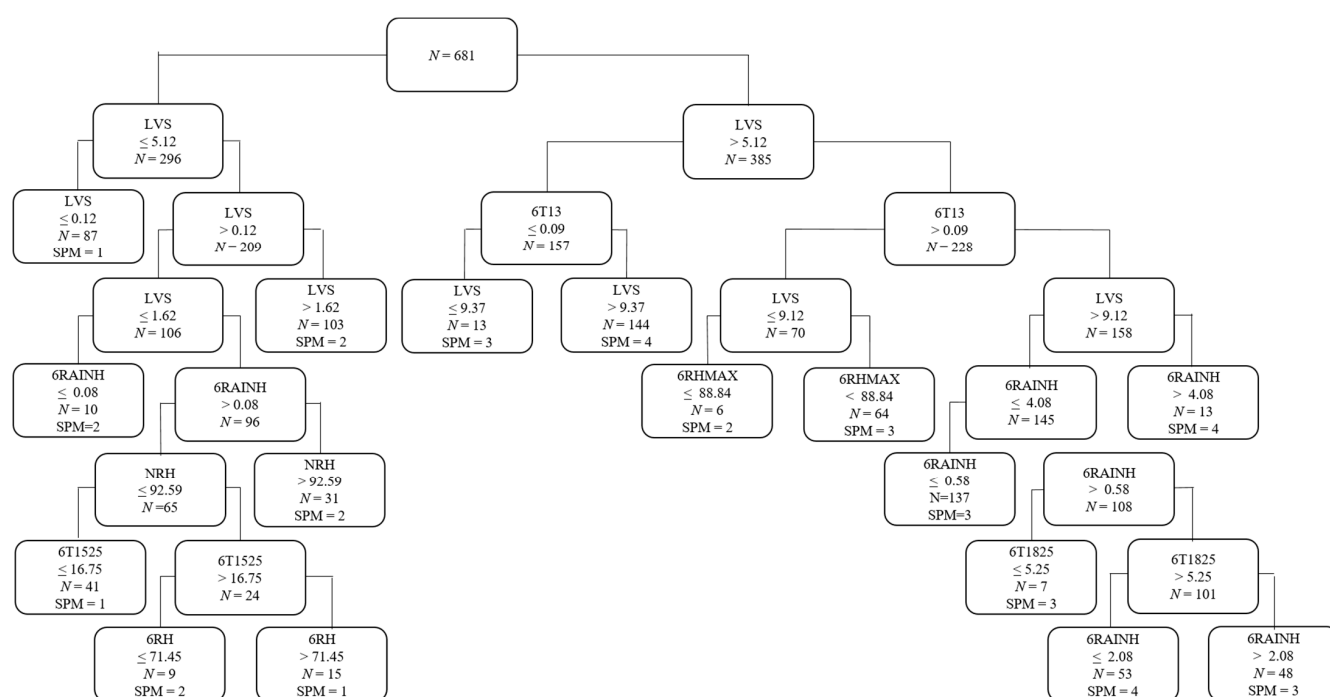
Figure 6. Representation of the classification tree built with number of susceptible leaves (LVS), mean number of hours at temperature above 13 °C during the previous six days (6T13), mean elative humidity during the previous six days (6RH), mean number of rainy hours during the previous six days (6RAINH), mean number of hours at temperature between 15 and 25 °C during the previous six days (6T1525), and mean maximum relative humidity during the previous six days (6RHMAX).

4. Discussion

Regardless of the type of disease management, whether it is based on a pre-established schedule, reasoned, integrated or organic, risk estimation is an essential component. In order to ensure the economic, environmental, and social sustainability of agricultural production, it is essential to rationalize the use of pesticides, including fungicides, whether synthetic or not. Strategically, it is expensive to treat when the risk is low or not to treat when it is high [27]. In other words, a risk-estimation tool must optimize the true positives (sensitivity) and the true negatives (specificity). There are several ways to determine a risk, but in most cases the risk is estimated as being below or above a threshold. Thresholds are established from the relationship between disease intensity and economic damage or yield losses [42,43]. Subsequently, it is possible to determine an action threshold corresponding to the moment when, if a treatment is not applied, the cost in yield losses will be higher than the cost of the treatment. In practice, since it is not generally possible to treat instantly when the action threshold is reached, for example because the treatment conditions are not favorable (rain, wind, availability of workers), a warning threshold is used to choose the best time to act.

In the case of strawberry powdery mildew, the relationship between severity of powdery mildew on leaves and losses of fruit yield has been established [3,10]. More recently, Fall and Carisse [32] reported a linear relationship between severity (PLAD) on leaves and yield losses on fruits (%) and determined that yield losses of 1% and 5% corresponded to severities on leaves of 5% and 15%. In other words, the PLAD of 5% and 15% can be used as warning and action thresholds, respectively. It is from these observations that the severity classes were established in the present study as Class 1: PLAD = 0; Class 2: PLAD > 0 and <5%; Class 3: >5% and <15%; and Class 4: PLAD > 15%.

Over the past decades, many disease risk estimation tools have been developed. These tools vary in complexity from simple decision rules to expert systems including forecasting and dynamic simulation models [28,32,44]. Since many factors related to the

pathogen, host, or environment influence the epidemic dynamics, some risk-estimation tools are based only on weather conditions or a combination of variables related to weather conditions, susceptibility of the plant (varietal and ontogenic resistance), and population of the pathogenic agent (size of the population and virulence). The development of strawberry powdery mildew is no exception. Van der Heyden et al. [21] reported a linear relationship between *P. aphanis* airborne conidial concentration and proportion of leaf area diseased (PLAD). Our observations are in accordance with those reported by Van der Heyden et al. [21], with 0.09, 0.71, 1.48, and 2.39 $\log_{10}$ (conidia + 1)/m^3 for SPM severity class 1, 2, 3 and 4, respectively (Figure 2A). Leaf disease severity is also influenced by leaf age [6,16]. Very young leaves that are still folded (angle of less than 45 between leaflet) are very susceptible, whereas leaves partially unfolded (angle of more than 45 between leaflet) are moderately susceptible, and completely unfolded and pale-green leaves are practically resistant [6]. In our study we expressed leaf susceptibility in terms of number of susceptible leaves per meter of row; a leaf was considered susceptible if leaflets were not completely unfolded [6]. Considering that the window of susceptibility is narrowed, it was expected that leaf susceptibility would be a good indicator of disease risk. In fact we observed that the number of susceptible leaves per meter of row increases with increasing disease severity classed, with an average of 0.80, 3.40, 12.39, and 17.94 for severity classes 1, 2, 3, and 4, respectively (Figure 2B). For the weather conditions, there were not clear relationships between individual weather variables and disease severity (Figure 3). Nevertheless, weather data averaged over the preceding six days generally had a higher correlation with severity classes than daily values (Table 2, Figure 3).

In practice, although these three groups of variables influence the development of strawberry powdery mildew, several combinations of these variables can cause the same disease severity. For example, conditions characterized by low inoculum and few susceptible leaves but highly favorable weather conditions can cause the same severity as conditions characterized by high inoculum, few susceptible leaves and less-favorable weather conditions. The interaction between these variables is therefore very important because, in theory, regardless of whether the weather conditions are favorable or not, if the inoculum is absent or very few susceptible leaves are present, the severity of the disease will be zero, or very low. Classification trees make it possible to identify all combinations of variables that cause the same severity, expressed as classes [29,30,45]. In this study, the best classification tree, based on accuracy, was built using airborne inoculum concentration, anumber of susceptible leaves, mean relative humidity (%), mean number of hours at temperature between 18 and 30 C, and mean number of hours with saturation vapor pressure between 10 and 25 mmHg, all with weather variables being averages over the previous six days (Table 2, Figure 4). The accuracy for this tree ranged from 0.82 to 0.97 (Table 2, Figure 4). The second best tree was built using airborne inoculum concentration, mean number of hours at temperature between 18 and 25 °C, mean maximum relative humidity (%), mean number of hours with saturation vapor pressure between 10 and 25 mmHg, and mean number of rainy hours; all weather variables being averages over the previous six days (Table 2, Figure 5). The accuracy for this tree ranged from 0.81 to 0.95 (Table 2, Figure 5). The tree built with weather and host susceptibility predictors had a similar overall accuracy, but lower sensitivity mainly for SPM classes 2 and 3 (Table 2, Figure 6). Also, for the tree built with only weather predictors, 15 predictors (variables) were needed to classify SPM severities with accuracy ranging from 0.70 to 0.89 (Table 2). Nevertheless, the tree is complex (many branches and splits) which may restrict its implementation and adoption by crop advisors and growers (tree not shown).

Regardless of the complexity of disease management decisions, the key element is the estimation of risk. Informed and rational disease management decision cannot be taken without some knowledge about the risk; which can be defined as the probability that a disease reach a critical level usually expressed as potential yield losses. Prediction of risk allows growers to respond in a timely and efficient manner by adjusting their crop management actions. A prediction of high disease risk may result in reduced yield

losses whereas low disease risk may result in reduced pesticide applications with positive economic and environmental effects. There are several types of disease prediction models with a range of complexity from rule-based to complex simulation models [23–26,32]. Nevertheless, disease risk prediction models are based on the interactions of all or some of the factors (predictors) that govern epidemic development: the host, the pathogen, and the environment, developed from controlled or field experiments. Strawberry powdery mildew is not an exception and there are several different types of disease risk prediction models, each one with their advantages and limitations [28]. In this study we explored a relatively new approach for plant disease risk estimation, classification trees. Because we used only field data, the trees were developed and validated using a large number of observations representing different combinations of SPM classes and inoculum-, host susceptibility-, and weather-based predictors (681 and 136 for a total of 817).

Despite the need to monitor airborne inoculum, the tree built with inoculum, host, and weather predictors was the most reliable and simplest one. Basically, it estimates the risk of SPM, first from airborne inoculum and amount of susceptible leaves, and then from humidity, duration of favorable temperatures and of saturation vapor pressure during the previous 6 days. In other words, the tree is coherent with the epidemiology of strawberry powdery mildew [3,6,10,21–23,26]. Among the four trees, it is the most intuitive and easy to interpret (Figure 4). It uses only three weather-based predictors readily available from weather stations, simple calculation, or forecasts (Table 2, Figure 4). The second best tree that was developed without the information of the number of susceptible leaves, also required data on airborne inoculum and four weather-based predictors readily available (Table 2, Figure 5). The choice of classification tree depends on the objectives of SMP management as someone can look for a very high level of control regardless of the cost (high sensitivity), someone may look for reducing management actions as much as possible (high specificity). In general, we are looking for a balance between acting when needed and not acting when not needed (high accuracy).

In this study, the amounts of airborne inoculum and the number of susceptible leaves were monitored. However, depending on the resources available, monitoring efficiency can be improved or values estimated from models [44]. In fact, monitoring susceptible leaves was easy and rapid. Hence, it can be included into scouting services already available. For monitoring airborne inoculum, it is more difficult and time consuming. However, previous study showed that because of the low level of spatial heterogeneity in *P. aphanis* airborne inoculum, it can be estimated using only one sampler per strawberry field [21]. In addition, with the advances in molecular biology, airborne inoculum can be assessed using qPCR, LAMP or new field DNA analysis technologies, which would allow for simultaneous monitoring of inoculum concentration and fungicide resistance.

5. Conclusions

The approach used in this study to determine risk of strawberry powdery mildew has proven to be reliable. Indeed, the development of classification trees has made it possible to highlight all the conditions that lead to the development of the disease. The classification tree, which essentially brings together the classification rules, makes it possible to estimate risks with a high reliability (Table 2). Although a large number of cases have been used to develop and validate the classification tree, as for all tools, it will need to be validated under commercial conditions, for other varieties of strawberry and under other production conditions. In any case, the rules for estimating the risk of strawberry powdery mildew should be easily integrated into the various phytosanitary warning platforms already available.

Author Contributions: Conceptualization, O.C. and M.L.F.; methodology, O.C. and M.L.F.; software, O.C.; validation, O.C. and M.L.F.; formal analysis, O.C.; investigation, O.C. and M.L.F.; resources, O.C.; data curation, O.C.; writing—original draft preparation, O.C.; writing—review and editing, O.C. and M.L.F.; visualization, O.C.; supervision, O.C.; project administration, O.C.; funding acquisition, O.C. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available by contacting the authors.

Acknowledgments: The authors gratefully acknowledge Annie Lefebvre from Agriculture and Agri-Food Canada's Saint-Jean-sur-Richelieu Research and Development Centre for her advices and help in collecting data. The authors would like to thank all of the summer students who participated in the data collection and the producers who gave us access to their fields. This work was financially supported by Agriculture and Agri-Food Canada.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Statistics Canada. *Table 32-10-0364-01 Area, Production and Farm Gate Value of Marketed Fruits*; Statistics Canada: Ottawa, ON, Canada. [CrossRef]
2. Carisse, O.; Lefebvre, A.; Van der Heyden, H.; Roberge, L.; Brodeur, L. Analysis of incidence–severity relationships for strawberry powdery mildew as influenced by cultivar, cultivar type, and production systems. *Plant Dis.* **2013**, *97*, 354–362. [CrossRef]
3. Carisse, O.; Morissette-Thomas, V.; Van der Heyden, H. Lagged association between powdery mildew leaf severity, airborne inoculum, weather, and crop losses in strawberry. *Phytopathology* **2013**, *103*, 811–821. [CrossRef] [PubMed]
4. Carisse, O.; Maas, J. Les maladies du fraisier. In *Our Strawberries/Les Fraises de Chez Nous*; Khanizadeh, S., DeEll, J., Eds.; Gouvernement du Canada, Les éditions et Services de dépôt: Ottawa, ON, Canada, 2005; p. 551. ISBN 0660623382.
5. Amsalem, L.; Freeman, S.; Rav-David, D.; Nitzani, Y.; Sztejnberg, A.; Pertot, I.; Elad, Y. Effect of climatic factors on powdery mildew caused by *Sphaerotheca macularis* f. sp. *fragariae* on strawberry. *Eur. J. Plant Pathol.* **2006**, *114*, 283–292. [CrossRef]
6. Carisse, O.; Bouchard, J. Age-related susceptibility of strawberry leaves and berries to infection by *Podosphaera aphanis*. *Crop. Prot.* **2010**, *29*, 969–978. [CrossRef]
7. Berrie, A.M.; Burgess, C.M. The effect of post-harvest epidemics of powdery mildew on yield and growth of strawberry cv. Elsanta. *Proc. Third Strawb. Symp. Acta Hortic.* **1997**, *439*, 791–798. [CrossRef]
8. Freeman, J.A.; Pepin, H.S. Effect of postharvest infection of powdery mildew on yield of the strawberry cultivar Northwest. *Can. Plant Dis. Surv.* **1969**, *49*, 139.
9. Spencer, D.M. *The Powdery Mildews*; Academic Press: London, UK; New York, NY, USA, 1978.
10. Blanco, C.; de los Santos, B.; Barrau, C.; Arroyo, F.T.; Porras, M.; Romero, F. Relationship among concentrations of *Sphaerotheca macularis* conidia in the air, environmental conditions, and the incidence of powdery mildew in strawberry. *Plant Dis.* **2004**, *88*, 878–881. [CrossRef]
11. Jhooty, J.S.; McKeen, W.E. Studies on powdery mildew of strawberry caused by *Sphaerotheca macularis*. *Phytopathology* **1965**, *55*, 281–285.
12. Peries, O.S. Studies on strawberry mildew, caused by *Sphaerotheca macularis* (Wallr. ex Fries) Jaczewski: I. Biology of the fungus. *Ann. Appl. Biol.* **1962**, *50*, 211–224. [CrossRef]
13. Peries, O.S. Studies on strawberry mildew, caused by *Sphaerotheca macularis* (Wallr. ex Fries) Jaczewski: II. Host–parasite relationships on foliage of strawberry varieties. *Ann. Appl. Biol.* **1962**, *50*, 225–233. [CrossRef]
14. Sombardier, A.; Savary, S.; Blancard, D.; Jolivet, J.; Willocquet, L. Effects of leaf surface and temperature on monocyclic processes in *Podosphaera aphanis*, causing powdery mildew of strawberry. *Can. J. Plant Pathol.* **2009**, *31*, 439–448. [CrossRef]
15. Bouchard, J. *Épidémiologie et Évaluation de Systèmes Prévisionnels Comme Outil de Lutte Raisonnée Contre le Blanc (Sphaerotheca Macularis) Chez le Fraisier à Jour Neutre et Conventionnel*; Université Laval: Québec, QC, Canada, 2008.
16. Asalf, B.; Gadoury, D.M.; Tronsmo, A.M.; Seem, R.C.; Dobson, A.; Peres, N.A.; Stensvand, A. Ontogenic resistance of leaves and fruit, and how leaf folding influences the distribution of powdery mildew on strawberry plants colonized by *Podosphaera aphanis*. *Phytopathology* **2014**, *104*, 954–963. [CrossRef] [PubMed]
17. Xiao, C.L.; Chandler, C.K.; Price, J.F.; Duval, J.R.; Mertely, J.C.; Legard, D.E. Comparison of epidemics of Botrytis fruit rot and powdery mildew of strawberry in large plastic tunnel and field production systems. *Plant Dis.* **2001**, *85*, 901–909. [CrossRef] [PubMed]
18. Pertot, I.; Fiamingo, F.; Amsalem, L.; Maymon, M.; Freeman, S.; Gobbin, D.; Elad, Y. Sensitivity of two *Podosphaera aphanis* populations to disease control agents. *J. Plant Pathol.* **2007**, *89*, 85–96.
19. Sombardier, A.; Dufour, M.C.; Blancard, D.; Corio-Costet, M.F. Sensitivity of *Podosphaera aphanis* isolates to DMI fungicides: Distribution and reduced cross-sensitivity. *Pest. Manag. Sci.* **2010**, *66*, 35–43. [CrossRef] [PubMed]

20. Carisse, O.L.; Van der Heyden, H. A new risk indicator for botrytis leaf blight of onion caused by *Botrytis squamosa* based on infection efficiency of airborne inoculum. *Plant Pathol.* **2012**, *61*, 1154–1164. [CrossRef]

21. Van der Heyden, H.; Lefebvre, M.; Roberge, L.; Brodeur, L.; Carisse, O. Spatial pattern of strawberry powdery mildew (*Podosphaera aphanis*) and airborne inoculum. *Plant Dis.* **2014**, *98*, 43–54. [CrossRef] [PubMed]

22. Miller, T.C.; Gubler, W.D.; Geng, S.; Rizzo, D.M. Effects of temperature and water vapor pressure on conidial germination and lesion expansion of *Sphaerotheca macularis* f. sp. fragariae. *Plant Dis.* **2003**, *87*, 484–492. [CrossRef]

23. Dodgson, J.; Hall, A.; Parker, S. Rule based prediction system to optimize fungicide applications for control of *Podosphaera aphanis*. *Acta Hort.* **2009**, *842*, 355–358. [CrossRef]

24. Eccel, E.; Fratton, S.; Ghielmi, L.; Tizianel, A.; Shtienberg, D.; Pertot, I. Application of a non-linear temperature forecast post-processing technique for the optimization of powdery mildew protection on strawberry. *Ital. J. Agrometeorol.* **2010**, *2*, 5–14.

25. Bardet, A.; Vibert, J. L'oïdium du fraisier: Un outil de prévision du risque. *Infos Ctifl* **2011**, *276*, 38–44.

26. Hoffman, L.E.; Gubler, W.D. Validation of the UC Davis Strawberry Powdery Mildew Risk Index. *Calif. Strawb. Comm. Rep.* **2002**, *2*, 19.

27. Carisse, O.; Caffi, T.; Rossi, V. How to develop and validate plant disease forecasting systems. In *Exercises in Plant Disease Epidemiology*, 2nd ed.; Stevenson, K.L., Jeger, M.J., Eds.; APS-Press: St-Paul, MN, USA, 2014; ISBN 9780890544402.

28. Rossi, V.; Sperandio, G.; Caffi, T.; Simonetto, A.; Gilioli, G. Critical Success Factors for the Adoption of Decision Tools in IPM. *Agronomy* **2019**, *9*, 710. [CrossRef]

29. Ilic, M.; Ilic, S.; Jovic, S.; Panic, S. Early cherry fruit pathogen disease detection based on data mining prediction. *Comput. Electron. Agric.* **2018**, *150*, 418–425. [CrossRef]

30. Mehra, L.K.; Cowger, C.; Gross, K.; Ojiambo, P.S. Predicting Pre-planting Risk of Stagonospora nodorum blotch in Winter Wheat Using Machine Learning Models. *Front. Plant Sci.* **2016**, *7*, 390. [CrossRef]

31. Mukerji, K.G. *1968 Sphaerotheca macularis. C.M.I. Descriptions of Pathogenic Fungi and Bacteria*; Commonwealth Mycological Institute: Kew, UK, 2009.

32. Fall, M.L.; Carisse, O. Dynamic simulation modelling for plant disease severity prediction: A novelty toward improving accuracy of strawberry powdery mildew prediction. *Agric. For. Meteorol.* **2021**.

33. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*; Chapman and Hall: New York, NY, USA, 1993.

34. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Wadsworth Inc.: Monterey, CA, USA, 1984.

35. Mucherino, A.; Papajorgji, P.; Pardalos, P. *Data Mining in Agriculture*; Springer: Cham, Switzerland, 2009.

36. Witten, I.; Frank, E. *Data Mining Practical Machine Learning Tools and Techniques*, 2nd ed.; Morgan Kaufmann: Burlington, MA, USA, 2005.

37. Nelson, B.D. Variable Reduction for Modeling using PROC VARCLUS. In *SAS Institute Inc. (1999), SAS/STAT_ User's Guide*; SAS Institute Inc.: Cary, NC, USA, 1999; pp. 261–263.

38. Jackson, J.E. *A Users Guide to Principal Components*; John Wiley & Sons: New York, NY, USA, 1991.

39. Lewis, R.J. An Introduction to Classification and Regression Tree (CART) Analysis. Annual Meeting of the Society for Academic Emergency Medicine in San Francisco, California, 2000, Proceedings. Available online: https://www.researchgate.net/publication/240719582 (accessed on 14 November 2020).

40. Fall, M.L.; Van der Heyden, H.; Beaulieu, C.; Carisse, O. *Bremia lactucae* infection efficiency in lettuce is modulated by temperature and leaf wetness duration under Quebec field conditions. *Plant Dis.* **2015**, *99*, 1010–1019. [CrossRef]

41. Steinberg, D.; Colla, P.; Martin, K. *CART—Classification and Regression Trees: Supplementary Manual for Windows*; Salford Systems: San Diego, CA, USA, 1998.

42. Nutter, F.W.; Teng, P.S.; Royer, M.H. Terms and concepts for yield, crop loss, and disease thresholds. *Plant Dis.* **1993**, *77*, 211–215.

43. Savary, S.; Teng, P.S.; Willocquet, L.; Nutter, F.W., Jr. Quantification and modeling of crop losses: A review of purposes. *Annu. Rev. Phytopathol.* **2006**, *44*, 89–112. [CrossRef]

44. Fall, M.L.; Van der Heyden, H.; Carisse, O. A Quantitative Dynamic Simulation of *Bremia lactucae* Airborne Conidia Concentration above a Lettuce Canopy. *PLoS ONE* **2016**, *11*, e0144573. [CrossRef]

45. Paul, P.A.; Munkvold, G.P. A model-based approach to preplanting risk assessment for gray leaf spot of maize. *Phytopathology* **2004**, *94*, 1350–1357. [CrossRef] [PubMed]