Patterns and Predictions of *Barley yellow dwarf virus* Vector Migrations in Central Europe

Jana Jarošová, Wiktor Rafał Żelazny, and Jiban Kumar Kundu[†]

Crop Research Institute, Drnovska 507, Prague 6, 16106 Czech Republic

Abstract

The bird cherry-oat aphid (*Rhopalosiphum padi* (L.)) is one of the most detrimental pests of cereals, causing harm mainly by transmitting *Barley yellow dwarf virus* (BYDV). *R. padi* migration has been monitored since 1992 using suction traps at five sites in the Czech Republic. The count data were subjected to the following different analyses: (i) the minimum temperature thresholds for the aphids to take off were determined; (ii) a partial redundancy analysis using the minimum, average, and maximum temperatures, as well as the wind speed, the precipitation total, and past aphid migration descriptors was performed to explain the relationship between aphid occurrences and weather patterns; and (iii) three types of models from the field of machine learning were used to predict aphid

occurrences. According to our findings, (i) in Central Europe, 8°C is the temperature threshold for *R. padi* migration unless insufficient day-light delays the take-off; (ii) weather conditions occurring roughly 9 months before *R. padi* migration influence the migration size; (iii) the duration of the summer migration influences the autumn migration size; and (iv) the daytime and nighttime temperatures in the autumn determine the summer peak, whereas winter frosts and precipitation influence the autumn peak.

Keywords: cereals and grains, field crops, insecticide treatment, machine learning, *Rhopalosiphum padi*, viruses and viroids, weather

Cereal crops in Europe are affected primarily by two major aphid species; namely, the bird cherry-oat aphid *Rhopalosiphum padi* (L.) and the grain aphid *Sitobion avenae* (F.). The aphids cause damage to the plants mainly by transmitting the *Barley yellow dwarf virus* (BYDV) species. BYDV occurs frequently around the world (Burrows et al. 2009; Foster et al. 2004; Gill 1969; Henry et al. 1993; Jones et al. 1990; Kapooria and Ndunguru 2004; Svaneella-Dumas et al. 2013; Zhang et al. 2009) and, together with *Wheat dwarf virus* (WDV), it causes the most widely distributed viral disease of cereals in Central European countries (Beoni et al. 2016; Bukvayová et al. 2006; Kundu et al. 2009; Szunics et al. 2000). The incidence of BYDV is often spasmodic (Kendall et al. 1992) and temperatures after aphid fall-out are probably critical in BYDV infection (Wallin and Loonan 1971).

BYDV is primarily harmful in the autumn, when emerging winter cereals are infected. At this time of the year, *R. padi* is the predominant abundant species in the crops (Jarošová et al. 2013), and it can cause up to 90% of all autumn infections (Fabre et al. 2006), given that the autumn migration of *S. avenae* is very low (Carter et al. 1982). In Central Europe, *R. padi* typically has two distinct flight peaks. The first peak occurs in the autumn, when primary BYDV infections take place in winter crops; the second peak occurs in the summer. Under intensive farming, the prophylactic application of insecticides against the virus vectors is a common practice. However, these applications are often unnecessary and both economically and environmentally undesirable. Therefore, the time of migration in the autumn is of immense importance to the decision makers concerning insecticide treatments of the winter cereal crops, whereas

determining the beginning of the summer migration has a similar impact on spring barley and oat (Dedryver et al. 2010).

Traditional modeling methods were employed extensively in the 1980s and 1990s to create predictive models about cereal aphids. These models were based on different variables, including atmospheric conditions, natural enemies, crop qualities, and landscape characterizations (Carter et al. 1982; Freier et al. 1996; Watson and Carter 1983). The input data for those models were either obtained from experiments (Walters and Dewar 1986; Wikteliuss 1982) or assumed with the values calculated based on the life cycles of the individual insect species (Freier et al. 1996). These data were virtually inaccessible for end users such as farmers or agronomists. Furthermore, the models were usually created using less than 5 years of data, and therefore their reliability was limited.

The installation of suction traps throughout the world has enabled the long-term collection of data on different aphid species migrations. As the amount of data from the suction traps has increased, an overall picture of cereal aphid migration has been developing that might result in more accurate predictions of the migration timing and in more reliable estimates of insect numbers. The data also play an important role in determining the effect of climate changes on the insect populations (Bell et al. 2015; Harrington et al. 2007).

Suction traps are operated in five different locations in the Czech Republic (Table 1). The locations were chosen for their different weather and landscape conditions which, in turn, influence the crop preferences in individual areas. The traps were installed in 1992 by the Central Institute for Supervising and Testing in Agriculture of the Czech Republic.

The relationship between the numbers of aerial aphids caught in the suction traps, the actual aphid numbers in the nearby fields, and the BYDV infection rate has been investigated with varying results. Dean (1973) found no relationship between the numbers of cereal aphids caught in the suction traps in May and June and the peak densities on nearby crops, whereas Howard and Dixon (1990) claimed that there is a strong relationship between suction trap catches of *Metopolophium dirhodum* and aphid populations in the fields. The results of the aphid infectivity tests have been compared with the catches of these aphids by the Rothamsted suction traps, and they show that high numbers of alate aphids are not necessarily followed by increased BYDV incidence (Plumb 1976). By contrast, in New Zealand, there was a significant correlation between the numbers of aphid vectors that were trapped and the proportion of crops

[†]Corresponding author: J. K. Kundu; jiban@vurv.cz

J. Jarošová and W. R. Żelazny contributed equally to this work.

Funding: This work was supported by project numbers QJ1530373, LTACH-17010, and MZe RO0417.

The author(s) declare no conflict of interest.

Accepted for publication 20 March 2019.

with a >5% BYDV incidence (Farrell and Stufkens 1992). In a recent study, van Toor et al. (2016) stated that, when the aphid numbers were below a relatively low threshold during the risk period for each crop, the BYDV incidence was low regardless of the insecticide treatments in the fields. In this way, a high number of unnecessary insecticide applications could be avoided.

Modeling how a phenomenon is shaped by weather conditions involves working with numerous variables, which are characterized by multicollinearity, the presence of intricate interactions, and high information redundancy, to name just a few challenges. Informative yet simple approaches to the data analysis commonly found in experimental research such as linear models are not suitable for coping with these data. Acquiring insight into observable patterns requires methods for dimensionality reduction such as principal component analysis (PCA) (King and Jackson 1999), which assumes linear relationships between variables. In cases where prediction rather than explanation is the goal, solutions from the field of machine learning can be employed instead. Many of these “black-box” models make it possible to account for nonlinear relationships between variables, or they include algorithms of variable (feature) selection, at a price of substantially reduced interpretability (Kuhn and Johnson 2013).

Here, our aim was to explain aphid occurrences using Czech weather data and to develop predictive models that could enable farmers in Central Europe to predict the aphid migration time, duration, and total aphid numbers in the weeks ahead of their arrival.

Materials and Methods

Data collection. This study was based on the daily aphid counts provided by automatized traps that have been operated at five sites (Čáslav, Chrlice, Lípa, Věrovany, and Žatec) (Table 1) since 1998 in the Czech Republic. The dataset was provided by the Central Institute for Supervising and Testing in Agriculture of the Czech Republic, and it is available at <https://www.gitlab.com/wzelazny/jarsova2019patterns>. Some site and date combinations occurred more than once in the dataset. If the counts differed by 10 or less, the first value was treated as valid. If the difference exceeded this value or there were more than two records with the same date, all the values were discarded. No observations were made during the first and last days of a year. If the first or last available count was equal to 10 or less, the missing counts were replaced with a value of zero. The remaining missing data were imputed using linear interpolation (Fig. 1). For each site, the daily weather data were obtained from the Agri4cast Resources Portal (Biavetti et al. 2014). The weather descriptors included the average, minimum, and maximum temperatures, the total precipitation, and the average wind speed at 10 m. The relevant data subset is available at <https://www.gitlab.com/wzelazny/jarsova2019patterns-s1>.

Migration temperature threshold hypothesis. The daily data were used to verify the migration temperature hypotheses (Dry and Taylor 1970; Wikteliuss 1981). For each site and year combination for each summer and autumn flight period, the starting dates were derived. For each station, the moving average of order five was calculated with respect to the mean temperature to reduce the effects of fluctuations. The temperatures occurring on the starting days of summer migrations were then extracted, and their distribution was assessed visually and compared with those of published reports.

Patterns of aphid occurrence. For each combination of the site, year, and migration season, the aggregated sums of the aphid counts were calculated. The counts were logarithmically transformed to account for their skewness and the nonnegativity constraint. The daily weather records were aggregated on a per-decade basis. It was assumed that the characteristics of aphid flights in a given year were determined by the weather conditions occurring between September of the preceding year and October of the year in question. Similarly, as in Worner et al. (2002), a set of cascades comprising paired adjacent months was generated by averaging or summing up the per-decade measurements.

A partial redundancy analysis (pRDA) was performed, with the response matrix consisting of the starting dates, durations, and sizes of the summer and autumn migrations. The matrix of explanatory variables was based on the characteristics of the previous year's autumn migration and weather descriptors centered and scaled to the unit variance. A single conditioning variable was made by site. The obtained pRDA loadings were analyzed to search for patterns explaining the variations in the aphid migrations.

Predictive modeling. For each response variable, three types of models from the machine learning field were fitted to the training dataset; namely, an artificial neural network (Worner et al. (2002), a random forest, and multivariate adaptive regression splines. Unlike the pRDA, the models were intended to have a predictive rather than an explanatory character. For this reason, the weather variables for predicting the summer migration were restricted to the months between the preceding year's September and the given year's April whereas, for the autumn migration, January-to-August weather descriptors and summer migration characteristics were included in the model.

The dataset was randomly divided into training and validation subsets, with a size ratio of 3:1. The correlation coefficients between all the combinations of independent variables were calculated for the training subset, and variables were removed from both subsets until no coefficient values exceeded 0.4.

Fig. 1. Raw *Rhopalosiphum padi* counts from suction traps by site. Interpolated values are marked in red. Values exceeding 2,000 aphids are truncated to save space.

Table 1. Characteristics of the suction trap locations^a

Site	Region	Altitude (masl) ^b	Temp (°C) ^c	Precip (mm) ^d	GPS location
Čáslav	Central	260	8.9	555	49°54'10.015"N, 15°24'53.193"E
Chrlice	South Moravia	190	9.0	451	49°7'25.856"N, 16°38'2.599"E
Lípa	Vysočina	505	7.5	594	49°33'22.133"N, 15°32'13.146"E
Věrovany	Olomoucký	207	8.7	502	49°28'24.380"N, 17°16'27.069"E
Žatec	Ústecký	285	9.0	439	50°18'12.020"N, 13°31'16.407"E

^a Data were taken from the Central Institute for Supervising and Testing in Agriculture of the Czech Republic.

^b Meters above sea level (masl).

^c Long-term average temperature.

^d Long-term average precipitation sum (1972 to 2002).

The model parameters were tuned using 10-fold cross-validation. For each type of dependent variable (migration starting date, size, and duration), a model with the lowest residual mean square error was picked and refitted with the additional site factor included. The performance of the resulting six models (three model types, each with and without the site factor) was assessed using the validation dataset. Finally, the best-performing models were examined to identify the most influential predictors. Because these were artificial neural network models, a modified Garson-Goh approach (<https://beckmw.wordpress.com/2013/08/12/>) was used for this purpose.

The analyses were programmed in the R language (R Core Team, 2018) using the packages *vegan* (Oksanen et al. 2018) for pRDA; *nnet* (Venables and Ripley 2002), *randomForest* (Liaw and Wiener 2002), and *earth* (Milborrow 2018) to fit the predictive models; and *caret* (Kuhn 2018) for model tuning. The scripts are available at <https://www.gitlab.com/wzelazny/jarosova2019patterns> for reproducibility purposes.

Results

The threshold mean temperatures for the aphid take-off at the beginning of the migration were compared across five sites based on data taken over 19 years. The obtained temperature values associated with the beginnings of the summer and autumn migrations ranged from 8 to 23°C (Fig. 2). The temperatures for the summer peaks were slightly higher than those of the autumn peaks. The minimum temperatures at the beginning of the migration varied for the individual sites by as much as 13°C. The lowest temperature for the take-off was recorded for both summer and autumn migrations at the Lípa site. By contrast, the highest temperatures required for the take-off in the autumn were recorded at Chrlice and in the summer at Žatec.

A pRDA was performed in which the response matrix consisted of the starting date, duration and size of a summer and autumn migration, and matrix of explanatory variables based on characteristics of previous year's autumn migration and weather descriptors, taking the site effects into account. The investigated weather variables and the characteristics of the preceding year's autumn aphid migration explained 78% of the total variance, with the first four constrained axes explaining 25, 22, 13, and 8% of the variance. The site effect accounted for 9% of the variance, leaving 13% as residual variance.

Two major axes of variation can be identified on the pRDA biplot (Fig. 3). The first axis extends from the northwestern to the southeastern quadrant and differentiates among data points according to the starting date and the duration of the summer migration. Early migration onset had a positive influence on the migration duration and, with the increasing summer migration length, the size of the subsequent autumn migration seemed to be positively affected. According to the analysis, hot summers were associated with shortened migrations but the data do not confirm this relationship.

The autumn peak's abundance might have been predetermined by the harshness of the previous winter, as indicated by the minimum January-to-March temperatures. Moreover, abundant autumn migration can be associated with high February-to-March precipitation and a lack of substantial temperature drops from April to May.

Roughly orthogonal to the first axis, there is a gradient representing the variation in observations with respect to the starting date and the duration of the autumn migration. Similar to the summer migrations, autumn migrations that start early last for a long time. According to the pRDA results, these migrations tend to follow hot springs, as indicated by the maximum temperature weather descriptors, but the raw data do not confirm this pattern.

The summer migration abundance seems to be weakly correlated with the remaining peak attributes but it appears to be closely associated with the preceding autumn temperatures. Interestingly, whereas autumns with high maximum temperatures promote migration, in the case of mean temperatures, the opposite relationship is noted. A striking pattern can be observed with regard to the wind speed measurements; in autumn and winter, a high wind intensity has a positive influence on the migration size but, over time, the relationship becomes negative.

Among the examined models predicting migration abundance, artificial neural networks tended to give the best cross-validation results whereas, for starting dates and durations, the best-performing models were almost exclusively random forests (Table 2). However, in contrast to the artificial neural networks, the latter models performed poorly at the validation stage (Fig. 4). In two cases, random forest validations yielded slightly negative correlation coefficients. Artificial neural networks provided reliable predictions, with estimated correlation coefficients of between 0.60 and 0.73. The most precise predictions pertain to the summer migration abundance. Interestingly, the performance of the autumn migration model does not drop after the site information is removed, suggesting its possible continued employment beyond the five examined experimental sites.

The examination of the relative variable importance within the artificial neural network models (Fig. 5) revealed the strong influence of the Lípa and Věrovany site variables on the predicted summer migration abundance, followed by the preceding year's

Fig. 2. Distribution of *Rhopalosiphum padi* migration threshold temperatures according to the site and season.

October-to-November temperatures, the influence of which is ambiguous. With the site effect not accounted for, other weather descriptors gain prominence. In particular, nonwindy weather in autumn and a lack of high temperatures at the turn of the year seem to promote high numbers of aphids at a summer peak. The influence of the minimum November-to-December temperatures (the southeastern quadrant) is reversed. The models predicting autumn migration exhibit more consistency with regard to the variable importance, because the site factor has a limited influence on the predictions. A closer examination highlights the positive influence of the minimum temperatures during the winter months (January to March), the size and duration of the summer peak, high early spring precipitation, and a lack of extremely high temperatures in May to June on the autumn peak abundance.

When trends in both pRDA and the prediction models are compared, the following relationships are shared by both analyses:

1. The summer peak aphid abundance is determined by the preceding autumn temperatures.
2. An early onset of summer aphid migration positively influences the migration length and leads to abundant autumn migration.
3. A lack of severe frosts in the winter and early spring increases the number of aphids during the autumn peak.
4. Abundant rain or snow in the early spring causes an increase in the autumn peak populations.

Discussion

Importance of aphid migration modeling. Epidemics of viral diseases transmitted by insect vectors such as BYDV or WDV are sporadic and do not justify routine prophylactic treatments of winter cereals, as are often performed. Considering the negative effects of

Fig. 3. Ordination biplot of the first two partial redundancy analysis (pRDA) axes. The response matrix consists of the *Rhopalosiphum padi* migration starting dates (start_{sum} and start_{aut}, respectively), migration durations (dur_{sum} and dur_{aut}), and migration sizes (count_{sum} and count_{aut}), each for summer and autumn, respectively. Abbreviations: P = precipitation total; W = wind speed; T_{min}, T_{avg}, and T_{max} = minimum, average, and maximum temperatures, respectively. Asterisks signify a previous year.

excessive insecticide use on natural flora and fauna and on human health (Wilson and Tisdell 2001), these applications should be restricted to years in which there is an increased risk of infection. To accomplish this goal, a reliable prediction of aphid behavior in cereal stands prior to the summer and autumn aphid migrations is essential. Abiotic site conditions govern the pattern and timing of infestation by winged viruliferous vectors (primary infection) and the speed of viral dissemination within a field by apterous aphids (secondary spread) (Fabre et al. 2005). We attempted to identify the relationships between the weather patterns and aphid occurrences for the climatic conditions of Central Europe and to develop models that could be used to plan for insecticide treatments. We believe that the predictions should be based on information that is easily available to farmers for a convenient period ahead of time. Because local meteorological data are readily available to the general public, weather variables form the core of these model inputs.

Aphid migrations and temperature. The take-off of the aphids is supposed to be controlled by the temperature thresholds; specifically, the lowest temperatures at which the aphids are willing to fly (Dry and Taylor 1970; Walters and Dixon 1984; Wikteliu 1981). According to Wikteliu (1981), these thresholds vary for *R. padi* between seasons. They are highest during the spring (16 to 17°C), lowest in autumn (9 to 10°C), and intermediate in the summer (13 to 14°C). Dry and Taylor (1970) recorded 14 to 15°C as the threshold in the spring. However, in our study, winged cereal aphids were caught in suction traps during both spring and autumn at temperatures as low as 8°C. This finding is consistent with those of Klueken et al. (2009) in Germany, where *R. padi* individuals were caught in suction traps at the same temperature of 8°C. No aphids were caught in the traps at temperatures below this value during the 20 years of aphid monitoring in the Czech Republic. Therefore, for Central Europe and for the time being, 8°C seems to be the threshold for both summer and autumn *R. padi* migrations. The take-off can be delayed depending on factors other than the temperature. Accordingly, although the observed threshold temperatures did not differ significantly between the seasons, they varied between the sites. The biggest differences, which were as high as 13°C, were recorded for Lpá and Chrlice (and Žatec prior to the summer). The factors contributing to the take-off delay include crowding, plant quality (Walters and Dixon 1982), and the photoperiod (Marcovitch 1924). The first two factors do not seem to differ significantly between the individual sites, leaving the amount of daylight as the potential crucial factor to explain the observed discrepancies. Lpá has the lowest average temperatures among the studied sites, and Chrlice is characterized by having the warmest climate. In addition, Žatec is located in a region with a warm climate. At Chrlice and Žatec, the aphids delayed their take-off until the desired photoperiod conditions were present, which coincided with temperatures much higher than the 8°C threshold.

As expected, the temperature was the primary factor explaining the variation in migration characteristics across the years and sites, and it played a major role in the predictive models. However, the time span

between the period of influence and the period when the outcome was actually evident was quite surprising. The summer peak abundance appeared to be determined by the weather during the preceding autumn, whereas the preceding winter was decisive for the autumn peak. This time is equivalent to a period of approximately 9 months between the cause and the effect. In similar studies in New Zealand, which were performed by Lankin-Vega et al. (2008) and Worner et al. (2002), the abundance of *R. padi* in autumn cereals was determined by the temperatures occurring between 6 and 2 months before the actual peak. However, due to the absence of a primary host and the specific climate, *R. padi* overwinters anholocyclically in that part of the world, hindering the comparison of the patterns with findings from Central Europe.

Summer peak. The summer peak is predetermined by the autumn temperatures, when high maximum temperatures have a positive effect on the peak's size, whereas high mean temperatures decrease the size of the peak. The observed pattern suggests that warm autumn days with high afternoon temperatures and possibly lower night temperatures stimulate the following year's migration. This hypothesis could be explained in terms of a sexual forms ratio. There is still no agreement regarding the proportions of *R. padi* that overwinter anholocyclically in areas such as the Czech Republic, where the winter severity varies over the years. Autumn temperatures might play an important role in the outcome. Sexual morphs are produced during longer night lengths at lower temperatures, resulting in sexual morphs being produced earlier in the year when the autumns are cool (Hardie 2009; Van Emden and Harrington 2017). The days with greater differences in the temperature between day and night can be beneficial for the sexual forms, because their production is stimulated by low night temperatures and their long-distance migration

Table 2. Cross-validation of residual mean square errors obtained for the artificial neural network, random forest, and multivariate adaptive regression spline models without and with the site factor included, according to the predicted variable^a

Predicted variable ^b	Without site factor	With site factor
log count _{sum}	0.34 , 0.35, 0.38	0.28 , 0.33, 0.31
log count _{aut}	0.33 , 0.35, 0.35	0.34 , 0.35, 0.36
dur _{sum} (days)	14.14, 13.65 , 14.14	15.68, 13.62 , 14.22
dur _{aut} (days)	9.09, 8.59 , 8.75	9.63, 9.06, 8.80
start _{sum} (days)	13.78, 9.97 , 10.82	10.64, 9.84 , 11.10
start _{aut} (days)	9.40, 6.58 , 7.45	10.50, 6.68 , 7.58

^a The value associated with the best-performing model of each triplet is type-set in bold.

^b Abbreviations: count_{sum} and count_{aut} = summer and autumn migration sizes, respectively; dur_{sum} and dur_{aut} = summer and autumn migration durations, respectively; and start_{sum} and start_{aut} = summer and autumn migration starting dates, respectively.

Fig. 4. Validation of the artificial neural network models according to the predicted variable, where start_{sum} and start_{aut} = starting date for *Rhopalosiphum padi* migration, dur_{sum} and dur_{aut} = migration durations, and count_{sum} and count_{aut} = migration sizes, each for summer and autumn, respectively.

The relatively low importance of winter temperature characteristics in determining the summer aphid abundance would support the theory according to which the preponderance of the overwintering population is in the Czech Republic in the forms of eggs, because they are particularly well adapted to survive periods of low temperatures (Pritchard and Vickers 2017), whereas the subtle effects of cold stress arising from even a single exposure to subzero temperatures significantly affects aphid life dynamics (Hutchinson and Bale 1994). Furthermore, for the species that overwinter anholocyclically in the regions of interest, the relationship between the winter and early spring temperatures and the size of the summer migration has usually been relatively stable and strong in previous studies (Harrington et al. 2017). Therefore, it could be assumed that the majority of *R. padi* are holocyclic in the Czech Republic.

Furthermore, the build-up of autumn aphid abundance is positively influenced by high February-to-March precipitation. With the precipitation at this time of the year usually taking the form of snow and February being the month with the lowest minimum temperatures, a high build-up can be due to the snow cover protecting the eggs from severe freeze damage or, more likely, protecting them

Wikteliu (1987) described a negative correlation between the sizes of the summer and autumn migrations. However, in our study, it was the length of the first peak duration rather than its intensity that positively influenced the size of the autumn peak. An extended length of summer migration enhances the autumn migration, probably by shortening the gap between the cereal crops. During this period, aphids have to seek refuge on volunteer plants, corn, and wild grasses, and their numbers usually decrease dramatically. When the aphid numbers are kept above a given threshold during the summer, the autumn build-up of the population can reach relatively high levels. This observation is further supported by the shortening of summer migrations and the decreased autumn peaks in years with hot summers, as revealed by both the pRDA and the artificial neural network model. However, high summer rainfall has a positive influence on the duration and size of summer and autumn migrations. Summer rains contribute to high autumn numbers of *R. padi* by reducing the extent of water scarcity, as observed in the dry Mediterranean climate of Australia (Thackray et al. 2009) and, to a lesser degree, in the United Kingdom (Harrington 2002). The moistened

soil supports grass weeds and other crop volunteers, creating a “green bridge” between cereal crops.

Study limitations. Even though the course of BYD epidemics in a field strongly depends on the proportion of plants infested by aphids (Fabre et al. 2003), it is often suggested that BYDV epidemics cannot be predicted from the suction trap catches alone. The relationship between the numbers of cereal aphids caught in suction traps and the peak densities achieved on nearby crops has sometimes been found to be unreliable. Dean (1973) found no relationship between the numbers of cereal aphids trapped in suction traps and the numbers found on crops during May and June. Gillet et al. (1990) found that trap-catch data effectively predicted virus incidence only during some years. Notably, positive correlations have been found between aphids collected in suction traps and aphid activity within 30 km of the trap in the United States (Halbert et al. 1990), New Zealand (Teulon et al. 2004; Thackray et al. 2009), and France (Fabre et al. 2010).

Because only some aphids carry the virus, even an accurate migration prediction does not provide information about the expected extent of infection. Whereas low aphid numbers in suction traps predict no or minimal yield losses due to BYDV infections in cereal crops, high aphid numbers can but do not have to forecast a high BYDV occurrence (Fabre et al. 2003; van Toor et al. 2016). Therefore, when low aphid numbers in suction traps are predicted in the autumn or in the spring, it is very unlikely that high yield losses due to BYDV will occur; thus, prophylactic pesticide application can be avoided. However, the situation is different in the case of high aphid numbers predictions. Fabre et al. (2003) produced an accurate model, according to which high aphid densities in the fields did not always result in an important yield loss due to differences in the aphid infectivity. Plumb (1976) suggested combining data from the suction traps with information on the fraction of aphid migrants carrying the virus to obtain an infectivity index. He confirmed the strong positive correlation between the index value and the extent of viral damage in autumn-sown crops 10 years later (Plumb 1986). Interestingly, in a New Zealand study, there was a relationship between the numbers of aphids trapped in autumn and winter and the proportion of crops with >5% BYDV infection in late spring (Farrell and Stufkens 1992), perhaps because of uniform aphid infectivity.

Aphid infectivity is a variable that might increase the predictive power of the proposed artificial neural network models when accounted for. However, this improvement would contradict the basic premise of the study (i.e., the development of a solution that would be readily available to growers). Making an infectivity estimation requires specialized knowledge and access to laboratory equipment. It is unrealistic to expect this estimation from an average person in our target group. Nevertheless, the proposed approach should suffice to predict years with low BYDV occurrence caused by low aphid abundance. This information alone would enable a significant reduction in preventive insecticide use. Unnecessary treatments would continue only in years with abundant migrations of aphids exhibiting low infectivity.

Comparison with other predictive models. Our approach differs from many earlier models aimed at cereal aphids or BYDV prediction. Those models typically either are not available to end-users due to the very difficult nature of the input data (Carter et al. 1982; Howard and Dixon 1990; Morgan 2000); are based on input data that, by their nature, do not enable long-term prediction (e.g., the need to count the aphids in the fields) (Fabre et al. 2003, 2006); or are designed to be used in an area with very different characteristics (e.g., New Zealand) (Lankin-Vega et al. 2008; Thackray et al. 2009).

The German study by Klueken et al. (2009) resembles ours to the greatest degree but, according to their approach, a migration registered on a single day rather than an entire migration event is the basic unit of interest. We believe that information pertaining to the characteristics of a migration as a whole—that is, its beginning date, duration, and total size—is better suited to the needs of agricultural practice. Although the predictions that can be obtained using our approach are not as precise as those of the cited work, they are based on information that is available weeks before the actual migration. Early prediction is of vital importance for farm management. Perhaps the migration- and single-day-oriented approaches could complement

one another, with the former used for early and the latter for late but more precise predictions.

Furthermore, Klueken et al. (2009) seem not to account for temporal autocorrelations between observations made on adjacent days. In fact, it could be suspected that a naïve forecast model—that is, one assuming that the size of the migration is going to be the same as it is on the preceding day—would perform even better than the models evaluated in this article. Transforming counts into binary variables according to arbitrary thresholds results in an information loss. In our study, we have used a redundancy analysis, which is related to the linear discriminant analysis but does not assume a qualitative response variable.

Conclusions.

- In Central Europe, 8°C is the temperature threshold for *R. padi* migration. The take-off can be delayed, depending on the amount of daylight.
- Weather characteristics of roughly 9 months preceding the actual migration are of enormous importance to the migration size.
- Autumn temperatures influence the size of the summer migration peak, possibly by controlling the ratio of individual aphid sexual forms.
- In the autumn, large diurnal temperature amplitudes can be beneficial for the sexual forms because their production is stimulated by low night temperatures, whereas high day temperatures enable their long-distance migration.
- During autumns that are warm throughout the day and night, winged viviparous aphids might be produced more readily and the sexual forms tend to be born later in the autumn.
- The majority of *R. padi* is probably holocyclic in the Czech Republic.
- High precipitation in February, March, and summer will increase the duration of the summer migration.
- An extended summer migration leads to an abundant autumn migration.
- Winter precipitation may play a role in aligning the hatching of aphid eggs with the bud burst. Aphids feeding on the unfurling leaves produce offspring that can prolong the summer migration.
- The model can be used to avoid unnecessary pesticide applications in the case of low aphid number predictions. However, in the case of medium and high aphid number predictions, field scouting is recommended to be used simultaneously before the decision to use the pesticide for BYDV management.

Acknowledgments

We thank M. Perutka from the Central Institute for Supervising and Testing in Agriculture of the Czech Republic for providing the daily aphid monitoring data; the high school Gymnázium Jiřího Ortena in Kutná Hora for their kind assistance regarding the meteorological data; and B. Rabochová for her excellent technical assistance with the data preparation.

Literature Cited

- Armstrong, S. J., and Peairs, F. B. 1996. Environmental parameters related to winter mortality of the Russian wheat aphid (Homoptera: Aphididae): Basis for predicting mortality. *J. Econ. Entomol.* 89:1281-1287.
- Bell, J. R., Alderson, L., Izera, D., Kruger, T., Parker, S., Pickup, J., Shortall, C. R., Taylor, M. S., Verrier, P., and Harrington, R. 2015. Long-term phenological trends, species accumulation rates, aphid traits and climate: Five decades of change in migrating aphids. *J. Anim. Ecol.* 84:21-34.
- Beoni, E., Chrpová, J., Jarošová, J., and Kundu, J. K. 2016. Survey of *Barley yellow dwarf virus* incidence in winter cereal crops, and assessment of wheat and barley resistance to the virus. *Crop Pasture Sci.* 67:1054-1063.
- Biavetti, I., Karetos, S., Ceglar, A., Toreti, A., and Panagos, P. 2014. European meteorological data: Contribution to research, development, and policy support. Pages 922-907 in: *Second Int. Conf. Remote Sens. Geoinf. Environ. (RSCy2014)*. International Society for Optics and Photonics.
- Bukvayová, N., Henselová, M., Vajčíková, V., and Kormanová, T. 2006. Occurrence of dwarf virus of winter wheat and barley in several regions of Slovakia during the growing seasons 2001-2004. *Plant Soil Environ.* 52:392-401.
- Burrows, M., Franc, G., Rush, C., Blunt, T., Ito, D., Kinzer, K., Olson, J., O'Mara, J., Price, J., Tande, C., Ziem, A., and Stack, J. 2009. Occurrence of viruses in wheat in the Great Plains region, 2008. *Plant Health Prog.* 10. doi:10.1094/PHP-2009-0706-01-RS

- Carter, N., Dixon, A. F. G., and Rabbinge, R. 1982. Cereal Aphid Populations: Biology, Simulation and Prediction. Pudoc, Wageningen, The Netherlands.
- Dean, B. Y. G. J. 1973. Aphid colonization of spring cereals. *Ann. Appl. Biol.* 75: 183-193.
- Dedryver, C. A., Le Ralec, A., and Fabre, F. 2010. The conflicting relationships between aphids and men: A review of aphid damage and control strategies. *C. R. Biol.* 333:539-553.
- Dixon, A. F. G. 1976. Timing of egg hatch and viability of the sycamore aphid, *Drepanosiphum platanoidis* (Schr.), at bud burst of sycamore, *Acer pseudoplatanus* L. *J. Anim. Ecol.* 45:593-603.
- Dry, W. W., and Taylor, L. R. 1970. Light and temperature thresholds for take-off by aphids. *J. Anim. Ecol.* 39:493-504.
- Fabre, F., Dedryver, C. A., Leterrier, J. L., and Plantegenest, M. 2003. Aphid abundance on cereals in autumn predicts yield losses caused by *Barley yellow dwarf virus*. *Phytopathology* 93:1217-1222.
- Fabre, F., Dedryver, C.-A., Plantegenest, M., Hullé, M., and Rivot, E. 2010. Hierarchical Bayesian modelling of plant colonisation by winged aphids: Inferring dispersal processes by linking aerial and field count data. *Ecol. Modell.* 221:1770-1778.
- Fabre, F., Pierre, J. S., Dedryver, C.-A., and Plantegenest, M. 2006. Barley yellow dwarf disease risk assessment based on Bayesian modelling of aphid population dynamics. *Ecol. Modell.* 193:457-466.
- Fabre, F., Plantegenest, M., Mieuze, L., Dedryver, C., Leterrier, J. L., and Jacquot, E. 2005. Effects of climate and land use on the occurrence of viruliferous aphids and the epidemiology of barley yellow dwarf disease. *Agric. Ecosyst. Environ.* 106:49-55.
- Farrell, J. A., and Stufkens, M. W. 1992. Cereal aphid flights and Barley yellow dwarf virus infection of cereals in Canterbury, New Zealand. *N. Z. J. Crop Hortic. Sci.* 20:407-412.
- Foster, G. N., Blake, S., Tones, S. J., Barker, I., and Harrington, R. 2004. Occurrence of *Barley yellow dwarf virus* in autumn-sown cereal crops in the United Kingdom in relation to field characteristics. *Pest Manage. Sci.* 60:113-125.
- Freier, B., Triltsch, H., and Rossberg, D. 1996. GTLAUS—Ein Modell für die Interaktion Weizen-Getreideblattläuse-Prädatoren und dessen Nutzung für komplexe agrarökologische Studien [GTLAUS—A model of wheat-cereal aphid-predator interaction and its use in complex agroecological studies]. *Z. Pflanzenkrankh. Pflanzenschutz [J. Plant Dis. Prot.]* 103:543-554.
- Fu, Y. H., Piao, S., Zhao, H., Jeong, S., Wang, X., Vitasse, Y., Ciais, P., and Janssens, I. A. 2014. Unexpected role of winter precipitation in determining heat requirement for spring vegetation green-up at northern middle and high latitudes. *Glob. Change Biol.* 20:3743-3755.
- Gill, C. C. 1969. Annual variation in strains of Barley yellow dwarf virus in Manitoba, and the occurrence of greenbug-specific isolates. *Can. J. Bot.* 47: 1277-1283.
- Gillet, H., Dedryver, C. A., Robert, Y., Gamon, A., and Pierre, J. S. 1990. Assessing the risk of primary infection of cereals by barley yellow dwarf virus in autumn in the Rennes basin of France. *Ann. Appl. Biol.* 117:237-251.
- Halbert, S., Connelly, J., and Sandvol, L. 1990. Suction trapping of aphids in western North America (emphasis on Idaho). *Acta Phytopathol. Entomol. Hung.* 25:411-422.
- Hardie, J. 2009. Photoperiodism in insects: Aphid polyphenism. Pages 342-364 in: *Photoperiodism The Biological Calendar*. R. J. Nelson, D. L. Denlinger, and E. D. Somers, eds. Oxford University Press, New York, NY, U.S.A.
- Harrington, R. 2002. BYDV: The heat is on. Pages 34-39 in: *Barley yellow dwarf disease: Recent advances and future strategies*. Proc. Int. Symp. M. Henry and A. McNab, eds. Texcoco, Mexico.
- Harrington, R., Clark, S. J., Welham, S. J., Verrier, P. J., Denholm, C. H., Hulle, M., Maurice, D., Rounsevell, M. D., Cocu, N., and Consortium, E. U. E. 2007. Environmental change and the phenology of European aphids. *Glob. Change Biol.* 13:1550-1564.
- Harrington, R., Hullé, M., and Plantegenest, M. 2017. Monitoring and forecasting. Pages 515-536 in: *Aphids as Crop Pests*. H. van Emden and R. Harrington, eds. CAB International, Wallingford, U.K.
- Henry, M., George, S., Arnold, G. M., Dedryver, C. A., Kendall, D. A., Robert, Y., and Smith, B. D. 1993. Occurrence of Barley yellow dwarf virus (BYDV) isolates in different farmland habitats in western France and south-west England. *Ann. Appl. Biol.* 123:315-329.
- Howard, M. T., and Dixon, A. F. G. 1990. Forecasting of peak population density of the rose grain aphid *Metopolophium dirhodum* on wheat. *Ann. Appl. Biol.* 117:9-19.
- Hutchinson, L. A., and Bale, J. S. 1994. Effects of sublethal cold stress on the aphid *Rhopalosiphum padi*. *J. Appl. Ecol.* 31:102-108.
- Jarošová, J., Chrpová, J., Šíp, V., and Kundu, J. K. 2013. A comparative study of the *Barley yellow dwarf virus* species PAV and PAS: Distribution, accumulation and host resistance. *Plant Pathol.* 62:436-443.
- Jones, R. A. C., McKirdy, S. J., and Shivas, R. G. 1990. Occurrence of barley yellow dwarf viruses in over-summering grasses and cereal crops in Western Australia. *Australas. Plant Pathol.* 19:90-96.
- Kapooria, R. G., and Ndunguru, J. 2004. Occurrence of viruses in irrigated wheat in Zambia. *EPPO Bull.* 34:413-419.
- Kendall, D. A., Brain, P., and Chinn, N. E. 1992. A simulation model of the epidemiology of barley yellow dwarf virus in winter sown cereals and its application to forecasting. *J. Appl. Ecol.* 29:414-426.
- King, J. R., and Jackson, D. A. 1999. Variable selection in large environmental data sets using principal components analysis. *Environmetrics* 10:67-77.
- Klueken, A. M., Hau, B., Ulber, B., and Poehling, H. 2009. Forecasting migration of cereal aphids (Hemiptera: Aphididae) in autumn and spring. *J. Appl. Entomol.* 133:328-344.
- Kuhn, M. 2018. caret: Classification and Regression Training. <https://cran.r-project.org/web/packages/caret/index.html>
- Kuhn, M., and Johnson, K. 2013. *Applied Predictive Modeling*. Springer, New York, NY, U.S.A.
- Kundu, J. K., Jarošová, J., Gadiou, S., and Červená, G. 2009. Discrimination of three BYDV species by one-step RT-PCR-RFLP and sequence based methods in cereal plants from the Czech Republic. *Cereal Res. Commun.* 37:541-550.
- Lankin-Vega, G., Worner, S. P., and Teulon, D. A. J. 2008. An ensemble model for predicting *Rhopalosiphum padi* abundance. *Entomol. Exp. Appl.* 129:308-315.
- Leigh, S., and Van Emden, H. F. 2017. Population dynamics: Cycles and patterns. Pages 262-279 in: *Aphids as Crop Pests*. H. van Emden and R. Harrington, eds. CAB International, Wallingford, U.K.
- Liaw, A., and Wiener, M. 2002. Classification and regression by RandomForest. *R News* 2:18-22.
- Marcovitch, S. 1924. The migration of the Aphididae and the appearance of the sexual forms as affected by the relative length of daily light exposure. *J. Agric. Res.* 27:13-22.
- Milborrow, S. 2018. earth: Multivariate Adaptive Regression Splines. <https://cran.r-project.org/web/packages/earth/index.html>
- Morgan, D. 2000. Population dynamics of the bird cherry-oat aphid, *Rhopalosiphum padi* (L.), during the autumn and winter: A modelling approach. *Agric. For. Entomol.* 2:297-304.
- Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., and Wagner, H. 2018. *vegan: Community Ecology Package*. <https://cran.r-project.org/web/packages/vegan/index.html>
- Plumb, R. T. 1976. Barley yellow dwarf virus in aphids caught in suction traps, 1969-73. *Ann. Appl. Biol.* 83:53-59.
- Plumb, R. T. 1986. A rational approach to the control of barley yellow dwarf virus. *J. R. Agric. Soc. Engl.* 147:162-171.
- Pritchard, J., and Vickers, L. H. 2017. Aphids and Stress. Pages 132-147 in: *Aphids as Crop Pests*. H. van Emden and R. Harrington, eds. CAB International, Wallingford, U.K.
- R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Svanella-Dumas, L., Candresse, T., Hullé, M., and Marais, A. 2013. Distribution of *Barley yellow dwarf virus*-PAV in the sub-Antarctic Kerguelen Islands and characterization of two new *Luteovirus* species. *PLoS One* 8:e67231.
- Szunics, L., Pocsai, E., Szunics, L., and Vida, G. 2000. Viral diseases on cereals in central Hungary. *Acta Agron. Hung.* 48:237-250.
- Teulon, D. A. J., Stufkens, M. A. W., and Fletcher, J. D. 2004. Crop infestation by aphids is related to flight activity detected with 7.5 metre high suction traps. *N. Z. Plant Prot.* 57:227-232.
- Thackray, D. J., Diggle, A. J., and Jones, R. A. C. 2009. BYDV PREDICTOR: A simulation model to predict aphid arrival, epidemics of *Barley yellow dwarf virus* and yield losses in wheat crops in a Mediterranean-type environment. *Plant Pathol.* 58:186-202.
- van Emden, H., and Harrington, R., eds. 2017. *Aphids as Crop Pests*. CAB International, Wallingford, U.K.
- van Toor, R., Butler, R., Stufkens, M., and Teulon, D. 2016. Lower incidence of yellow dwarf disease in autumn-sown wheat crops in New Zealand is linked with sowing dates, insecticide regimes and aerial aphid numbers. *Australas. Plant Pathol.* 45:609-619.
- Venables, W. N., and Ripley, B. D. 2002. *Modern Applied Statistics with S*. Statistics and Computing. Springer, New York, NY, U.S.A.
- Wallin, J. R., and Loonan, D. V. 1971. Low-level jet winds, and vectors, local weather, and barely yellow dwarf virus outbreaks. *Phytopathology* 61:1068-1070.
- Walters, K. F. A., and Dewar, A. M. 1986. Overwintering strategy and the timing of the spring migration of the cereal aphids *Sitobion avenae* and *Sitobion fragariae*. *J. Appl. Ecol.* 23:905-915.
- Walters, K. F. A., and Dixon, A. F. G. 1982. The effect of host quality and crowding on the settling and take-off of cereal aphids. *Ann. Appl. Biol.* 101:211-218.
- Walters, K. F. A., and Dixon, A. F. G. 1984. The effect of temperature and wind on the flight activity of cereal aphids. *Ann. Appl. Biol.* 104:17-26.
- Watson, S. J., and Carter, N. 1983. Weather and modelling cereal aphid populations in Norfolk (UK) I. *EPPO Bull.* 13:223-227.
- Wiktelius, S. 1981. Diurnal flight periodicities and temperature thresholds for flight for different migrant forms of *Rhopalosiphum padi* L. (Hom., Aphididae). *J. Appl. Entomol.* 92:449-457.
- Wiktelius, S. 1982. Flight phenology of cereal aphids and possibilities of using suction trap catches as an aid in forecasting outbreaks. *Swed. J. Agric. Res.* 12:9-16.
- Wiktelius, S. 1987. The role of grasslands in the yearly life-cycle of *Rhopalosiphum padi* (Homoptera: Aphididae) in Sweden. *Ann. Appl. Biol.* 110:9-15.
- Wilson, C., and Tisdell, C. 2001. Why farmers continue to use pesticides despite environmental, health and sustainability costs. *Ecol. Econ.* 39:449-462.
- Worner, S. P., Lankin, G. O., Samarasinghe, S., and Teulon, D. A. J. 2002. Improving prediction of aphid flights by temporal analysis of input data for an artificial neural network. *New Zeal. Plant Prot.* 55:312-316.
- Zhang, W., An, D., and Ren, X. 2009. Progress on occurrence and integrated control of wheat yellow dwarf disease in China. *J. Triticeae Crop.* 2:37.