Preventive Veterinary Medicine 204 (2022) 105667

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

Development of a syndromic surveillance system for Irish dairy cattle using milk recording data

Alexander P. Douglass [a], Luke O'Grady [a,b], Guy McGrath [c], Jamie Tratalos [c], John F. Mee [d], Damien Barrett [e], Cosme Sánchez-Miguel [f], Simon J. More [c], Aurélien Madouasse [g], Martin Green [b], Jamie M. Madden [c], Conor G. McAloon [a,*]

[a] School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland
[b] School of Veterinary Medicine and Science, University of Nottingham, Nottingham, UK
[c] Centre for Veterinary Epidemiology and Risk Analysis, School of Veterinary Medicine, University College Dublin, Belfield, Dublin, Ireland
[d] Animal and Bioscience Research Department, Teagasc, Moorepark Research Centre, Fermoy, Co. Cork, P61 C997, Ireland
[e] One Health One Welfare Scientific Support Unit, Department of Agriculture, Food and the Marine, Agriculture House, Dublin, Ireland
[f] Cork Regional Veterinary Laboratory, DAFM, Bishopstown, Cork, Ireland
[g] INRAE, Oniris, BIOEPAR, 44300, Nantes, France

ARTICLE INFO

Keywords:
Syndromic surveillance
Early warning
Schmallenberg virus

ABSTRACT

In the last decade and a half, emerging vector-borne diseases have become a substantial threat to cattle across Europe. To mitigate the impact of the emergence of new diseases, outbreaks must be detected early. However, the clinical signs associated with many diseases may be nonspecific. Furthermore, there is often a delay in the development of new diagnostic tests for novel pathogens which limits the ability to detect emerging disease in the initial stages. Syndromic Surveillance has been proposed as an additional surveillance method that could augment traditional methods by detecting aberrations in non-specific disease indicators. The aim of this study was to develop a syndromic surveillance system for Irish dairy herds based on routinely collected milk recording and meteorological data. We sought to determine whether the system would have detected the 2012 Schmallenberg virus (SBV) incursion into Ireland earlier than conventional surveillance methods. Using 7,743,138 milk recordings from 730,724 cows in 7037 herds between 2007 and 2012, linear mixed-effects models were developed to predict milk yield and alarms generated with temporally clustered deviations from predicted values. Additionally, hotspot spatial analyses were conducted at corresponding time points. Using a range of thresholds, our model generated alarms throughout September 2012, between 4 and 6 weeks prior to the first laboratory confirmation of SBV in Ireland. This system for monitoring milk yield represents both a potentially useful tool for early detection of disease, and a valuable foundation for developing similar tools using other metrics.

1. Introduction

In the last decade and a half, emerging vector-borne diseases have become a substantial threat to cattle across Europe. In 2006, Bluetongue virus (**BTV**) spread across northern Europe affecting many member states for several years until 2010, before remerging again in 2015 in Central France (Meiswinkel et al., 2008; Courtejoie et al., 2018). In 2011, Schmallenberg virus (**SBV**) emerged in Germany and rapidly spread across Europe (Afonso et al., 2014).

Analysis of meteorological data has suggested that SBV most likely entered Ireland on the 11–12th of August 2012 via long-range wind transportation of *Culicoides* from southern England to southeastern Ireland (McGrath et al., 2018). However, the first laboratory confirmation of SBV in cattle was not reported until late October 2012 (Bradshaw et al., 2012).

In contrast, there has been no evidence of BTV incursion into Ireland, either during or following the cross-European Bluetongue disease outbreak of 2006–2007. However, transplacental and contact

Abbreviations: **SBV:**, Schmallenberg virus; **BTV:**, Bluetongue virus; **ICBF:**, Irish Cattle Breeding Federation; **MAE:**, Mean Absolute Error.

* Corresponding author.
E-mail address: conor.mcaloon@ucd.ie (C.G. McAloon).

https://doi.org/10.1016/j.prevetmed.2022.105667
Received 21 December 2021; Received in revised form 1 April 2022; Accepted 6 May 2022
Available online 10 May 2022

transmission did occur in Northern Ireland (Menzies et al., 2008). It is expected that the risks from vector-borne diseases in Europe may increase as a consequence of climate change, due to increases in the range of insect vectors (Purse et al., 2005; Caminade et al., 2019).

To mitigate the impact of the emergence of new diseases, outbreaks must be detected early. However, the clinical signs associated with many diseases may be nonspecific. Furthermore, there is often a delay in the development of new diagnostic tests for novel pathogens which limits the ability to detect emerging disease in the initial stages. Syndromic surveillance has been proposed as an additional surveillance method that could augment traditional methods by detecting aberrations in non-specific disease indicators.

Milk recording data represent one source of routinely collected data that could be used as an early disease indicator. Milk yield is ideal as a potential indicator because a large quantity of data are collected regularly, and milk yield can be affected rapidly by infectious diseases (Elgersma et al., 2018). There are no known examples of milk yield-based syndromic surveillance of dairy cattle in existence prior to 2013. Since then, systems have been tested in France (Madouasse et al., 2013); Belgium and the Netherlands (Poskin et al., 2016; Veldhuis et al., 2016).

There is no syndromic surveillance system for dairy cattle in Ireland thus far, whether based on milk yield or otherwise. The development of a syndromic surveillance model based on milk recording presents a particular challenge in Irish dairy herds given that a pasture-based, seasonally calving dairying system predominates in Ireland. Such systems are potentially prone to greater impact of meteorological changes that might impact on grass growth for example, as well as reducing the number of milk recording observations when a large proportion of the dairy herd is dry: December and January.

However, the development of a surveillance system for vector-borne diseases in Ireland has previously been recommended (Collins et al., 2019). A study of the efficacy of surveillance of sentinel flocks of Irish sheep for disease detection showed that such a system could help to supplement existing passive surveillance systems in Ireland, and the same could be true for a syndromic surveillance system of dairy cattle (Murray et al., 2019).

The aim of this study was to develop a syndromic surveillance system for Irish dairy herds based on routinely collected milk recording and meteorological data. We sought to determine whether the system would have detected the 2012 SBV incursion into Ireland earlier than conventional surveillance methods.

2. Materials and methods

2.1. Data selection and sampling

2.1.1. Milk recording data

All milk recording data from Irish dairy herds are held within the Irish Cattle Breeding Federation (ICBF) database. Approximately one third of Irish dairy farmers conduct herd milk recording with an average of 4–8 recordings per herd per year (Carty et al., 2017). The milk recording service in Ireland is provided to farmers by Milk Recording Organisations who transfer the milk recording data to the ICBF database as soon as sample processing is complete. We extracted all the milk recording data that were collected from 1st of January 2007 to the 19th of November 2019 and managed within the ICBF database. The data consisted of milk records of 2,256,853 cows from 11,260 herds, and included recording date, measurements of milk yield, fat, protein, lactose content, somatic cell count (SCC), parity, calving date and calving ease, in addition to anonymous identifiers for herds and individual animals.

2.1.2. Location data

Spatial summary data were provided for each herd. To maintain herd anonymity, the location of the centroid of the largest fragment of land for each herd was randomly positioned within national grids 2.5 km, 5 km and 10 km radii. Grid hexagons with less than five herds were removed from the dataset. This location ‘jittering’ process was conducted prior to the data being made available to the researchers, to comply with data protection rules.

2.1.3. Weather data

Weather data were extracted from each of 461 weather stations reporting usable data from the Met Éireann database. These stations were based around the Ireland and did not include the offshore buoy stations. Stations were selected based on proximity to each herd in the dataset. In each case, the average daily rainfall and the average daily temperature were extracted during the twelve-year sample period for each herd on each day that a recording was taken. Every station contained rainfall data, whereas not every station contained temperature data and some stations were missing data for specific dates. In these latter cases where data were missing, data from the nearest station with complete data were used. In addition to daily data for the 28-days prior to the recording date, the mean rainfall (in mm) and temperature (in °C) at each weather station were calculated for the 7, 14, 21 and 28-day periods prior to the milk recording date.

2.2. Data analyses

2.2.1. Data management and sampling

The analysis was performed using cow-level measurements of milk yield for each day. Entries with no location or milk yield data were deleted from the dataset. For each measurement, the days in milk (DIM) was calculated as the numeric value of the date of calving subtracted from the numeric value of the recording date. To prevent excessive computational load and analysis time, a stratified random sample of 10% of animal recordings per herd recording were used to train the model. In other words, for each date of milk recording for each herd, 10% of the recordings were randomly sampled for model training. This sample was taken from a ‘reference period’ of all data prior to 31st December 2011, prior to the incursion of SBV into Ireland. Only complete records were used for analysis, and milk recordings with missing data were omitted.

2.2.2. Temporal analyses

2.2.2.1. Linear mixed model. Test day milk yield was modelled at the individual cow level with a linear mixed model using the ‘lme4’ package (Bates et al., 2015). Test day milk yield had a hierarchical structure with milk yield at the individual cow level nested within herd, herd was therefore modelled as a random effect. Year was also modelled as a random effect. This was to facilitate prediction for the ‘next’ year (i.e. year k + 1), since if a fixed effect was used the coefficient for that year would have had to be estimated before any yields could be predicted.

Recording date milk yield (MY) for the i-th cow in the j-th herd, in the k-th year, was therefore modelled as:

$$MY_{i,j,k} \sim \beta_0 + b_{0j,k} + f(Day_k) + f(DIM_{i,j,k}) + \beta_1 EBI_{i,j,k} + \beta_2 \ln SCC_{i,j,k} + \beta_3 Parity_{i,j,k} + f(Weather_{j,k}) + \varepsilon_{j,k}$$

Where Day was the day of the year as a continuous variable; DIM was the days between calving date and recording day for that cow; Year was the year of recording; EBI was the milk production subindex of the economic breeding index for the i-th cow; lnSCC was the log-transformed somatic cell count for the i-th cow on that recording date and Weather was an aggregate effect based on temperature and rainfall data (see below). Parity was modelled as a categorical variable with values greater than five condensed into a single category (5 +). DAY and DIM were modelled using B-spline functions with the ‘splines2’ package (Wang and Yan, 2018). To select the number of knots for these variables, a range from 5 to 15 was trialled for each, resulting in 10^2 combinations. The pairwise combination of the number of knots for each of the two variables which resulted in the lowest Mean Absolute Error (**MAE**) using 10-fold cross-validation was used. This step was conducted in the full model, i.e. whilst adjusting for all the other covariates. Errors at each level in the model were assumed to be mean zero, normally distributed and an unstructured variance-covariance matrix.

To select the most useful weather variables to include in the model, a lasso regression was initially performed using all weather variables only as predictors of milk yield. The lasso regression was performed using the 'glmnet' package (Friedman et al., 2010). Variables with non-zero coefficients were carried forward into the full model. Following this process, the "Weather" effect represented in the equation was comprised of thirteen fixed effects tested in the model. The effects found were the temperature in ℃ at the nearest weather station with that data 11, 12, 13, 16, 21, 22, 24, 26, and 28 days before the recording date, the amount of rain in mm from 1, 2, and 5 days before recording, and the seven-day rolling average temperature for the week two weeks before the recording date.

The model was fitted using 10% of the data in the 'reference period' (defined under alarm generation). Model validation was performed using R-squared and MAE from both the training (10%) and retained (90%) data. Since the focus of our study was to compare models for predictive rather than explanatory purposes, R-squared was calculated directly on the observed and predicted observations and did not consider variance at different levels in the model.

2.2.2.2. Alarm generation. We used a CUSUM function similar to that used by Veldhuis et al. (2016) to detect anomalies in milk production. Using the model described above, milk production was predicted for the observation window of interest. The average cow error per day was calculated as the sum of predicted milk yields minus the sum of observed milk yields divided by the number of cows recorded on that day. Given a low number of recordings on Saturdays and Sundays, these recordings were amalgamated with the recordings on Mondays.

For each day in the time series, a CUSUM value was calculated as the maximum of either 0, or the previous day's CUSUM value plus that day's residual (predicted – observed milk yield) minus a constant (k), aggregated at a national level. The value for k was selected as the 97.5th percentile of the average of the daily residuals in the reference period. Alarms were generated when the CUSUM value exceeded a threshold (h). Following the same methodology as Veldhuis et al. (2016), different values of h were trialled set at $3k$, $5k$ and $7k$.

For the purpose of the alarm generation, the reference period, that is the period over which the model was trained, included all data up until 31st December 2011, with the prediction window being the subsequent 365 days (up until 31st December 2012).

2.2.3. Spatial analysis

We conducted a hotspot analysis based on the Getis-Ord Gi* statistic. Analyses were conducted separately for the months of August and September 2012, corresponding to the time period during which SBV was believed to have been introduced into the country. In each case, analyses were conducted based on aggregating data to the 2.5, 5 and 10 km radius hexagons for the country. For each hexagon, the average percentage prediction error for all of the recordings in that hexagon during each of the two time periods (August and September) were used for the analysis. Spatial analyses were conducted in R (R Core Team, 2019) using the 'spdep' (Bivand et al., 2013), 'sf' (Pebesma, 2018), and 'sp' (Bivand et al., 2013) packages.

3. Results

3.1. Descriptive statistics and linear mixed model evaluation

After initial data cleaning, the data set consisted of 7,743,138 milk recordings from 730,724 cows in 7037 herds. Number of recordings, number of cows, number of herds and average milk yield at recording remained relatively stable over the time period. The reference period (until 31st December 2011) contained 6,552,518 recordings, the sampled dataset for model training included 662,903 recordings, and the testing window consisted of 1,189,805 recordings. For the 10% of the reference period data on which the model was trained, the MAE and R-squared were 3.11 and 0.73, respectively. For the retained 90% of the reference period data, the MAE and R-squared were 3.58 and 0.65, respectively. The MAE and R-squared for the model test period (excluding the reference period data) were 3.65 and 0.65, respectively. (Table 1).

3.2. Alarm generation

After training the data on the reference time period and applying to the test period (2012), six alarms were generated using the $h= 3k$ threshold, five using the $h= 5k$ threshold, and five alarms were generated using the $h= 7k$ threshold. Alarm dates, using each of the 3 thresholds, are shown in Table 2. The prediction error (i.e. predicted-observed) per day (excluding weekends), along with the alarm periods are also shown in Fig. 1. All of the alarm periods using the high thresholds had corresponding time periods using the lower thresholds (Table 2). Of the six alarms generated using the lower threshold, 3 were in the test period (2012), with all of these occurring in late summer to winter of 2012. Three additional alarms were generated in the reference period (prior to 2012) at time periods in late 2007/early 2008, at the beginning of 2010, and in April 2010.

3.3. Spatial analysis

The results of the hotspot analyses are shown in Fig. 2. Hotspots were identified in each of the three different resolutions at each of the 2 different timepoints (August and September 2012). However, there appeared to be little association between hotspots identified at the same time using different resolutions. Furthermore, with the exception of a single hexagon in the southeast identified at a 10 km hexagon resolution in September, there was limited evidence of significant clusters of reduced milk yield clustered around coastal areas.

4. Discussion

This study aimed to develop and evaluate a system for the early detection of outbreaks of disease in dairy cows using regularly collected milk data in Ireland, as previously used elsewhere (Madouasse et al., 2013; Veldhuis et al., 2016). This system represents the first of its kind in Ireland, and incorporates weather data, which has not previously been used in this way.

Using the lower threshold, our model generated an alarm in September 2012 that persisted until December of 2012. During this alarm period, multiple alarms were generated using the middle and higher thresholds, with each of these starting in September 2012. This demonstrates that if our model had been in place in 2012 it would have generated alarms, using all three thresholds, approximately 4–6 weeks prior to the first laboratory confirmation in Ireland (Bradshaw et al., 2012). In addition, a preceding alarm was generated on the lower threshold starting on the 16th July and ending on 6th August 2012. Given that meteorological studies estimate that the SBV incursion occurred subsequent to this, it seems most likely that this was a false positive alarm (McGrath et al., 2018). Whilst the generation of these alarms seems promising, it cannot be considered to be a causal relationship and it is possible that another unidentified event may have resulted in the depression in milk yield detected in this study.

Our spatial analysis demonstrated that the location of hotspots of depressed milk yield were dependent on which method of spatial aggregation was used. Consistent patterns were not observed between

Table 1
Descriptive statistics of milk recording data from 2007 to 2012.

	2007	2008	2009	2010	2011	2012
Number of recordings	1,258,888	1,379,382	1,235,936	1,320,344	1,357,968	1,190,620
Average yield per recording (Kg/d)	23.4	22.7	22.1	23.0	23.5	22.4
Median DIM at recording	134	136	134	139	133	136
Number of cows	262,836	308,782	287,045	300,389	330,766	306,343
Number of herds	4465	4912	4297	4240	4461	5088

Table 2
Alarm periods generated in both the reference (before 2012) and test (2012) periods using 3 different alarm thresholds (3*k*, 5*k* and 7*k*).

Alarm period	Threshold = 3*k*	Threshold = 5*k*	Threshold = 7*k*
1	26/12/2007–4/1/2008	26/12/2007–27/12/2007; 28/12/2007–31/12/2007	
2	6/1/2010–7/1/2010		
3	1/4/2010–20/4/2010	6/4/2010–16/4/2010	
4	16/7/2012–6/8/2012		
5	6/9/2012–18/12/2012	13/9/2012–12/12/2012	24/9/2012–25/9/2012; 26/9/2012–18/10/2012; 22/10/2012–25/10/2012; 29/10/2012–1/11/2012; 5/11/2012–30/11/2012
6	20/12/2012–31/12/2012	21/12/2012–24/12/2012	

analyses at the same time point with different aggregations nor across timepoints at the same spatial aggregation. Spatial aggregation degrades information and impacts on the ability to monitor spatial disturbances (Jeffery et al., 2014). Further analyses using spatial regression models which require the inclusion of fully specified spatial random effects at either the areal or point level would be required to fully investigate the use of spatial data for an Irish syndromic surveillance system (Rue et al., 2009). In this study, spatial aggregation was conducted so that data could be shared in a way that concealed the location and identity of individual farms, however, it is possible that spatial point analyses could be conducted on raw data, with the subsequent results then adapted or displayed such that individual farms are not identifiable.

Previous methods of spatial detection of anomalies have often used the SaTScanTM software based on the Kulldorf statistic (Kulldorff and Nagarwalla, 1995). One potential drawback of this approach is that it generally identifies clusters as circular patches. However, for islands such as Ireland, vector-borne diseases may be introduced at coastal areas, potentially producing non-circular clusters that may be difficult to detect because of an 'edge effect' (Tango and Takahashi, 2005). Therefore, for our study we used a hotspot analysis based on the Getis-Ord Gi* statistic, on the basis that fixed, non-overlapping spatial scales may be more sensitive to new infection emergence (Getis and Ord, 1992; Ord and Getis, 1995).

Specificity of a syndromic surveillance system is important when considering system utility. Our model detected three additional alarms in 2007/2008 and 2010 in each of the low and medium threshold systems. However, in two of these cases, the alarms occurred between December and January. In the Irish seasonal dairy production system, this period represents the point at which the vast majority of the Irish dairy herd is dry. These alarms were therefore generated on a smaller number of animals that may be considered 'atypical' with respect to the standard pasture-based production system. Therefore, alarms generated at this time of the year would need to be interpreted with particular caution.

Both the specificity and sensitivity of such alarm systems depend on the ability of the underlying model to accurately predict milk yield. In our case a mixed effects linear regression model was used to predict milk yield at an individual milk recording. However, further studies could seek to improve predictive capacity by considering incorporating an ensemble of modelling methods potentially including machine learning techniques.

There are some drawbacks to this method. As the temporal alarm system uses national milk recording data, it is possible that smaller local reductions in milk yield may go unnoticed, since reductions in the national average are most likely to be caused by either smaller reductions across the whole country or large reductions in one area. In addition, since not all herds milk-record, there could be regions with relatively poor data coverage. While the model was adapted to account for declines in milk yield based on weather, it is possible that this aspect of the model could be improved by explicitly modelling grass growth. Whilst grass growth data are collected across Ireland (Hurtado-Uria et al., 2013), these data were not available for the present study. Additionally, whilst our model appears well suited for diseases similar to SBV, the efficacy of this approach to detect a 'new' disease will be impacted by infection dynamics within the population. Previous work has demonstrated that such approaches are best suited for diseases that transmit relatively quickly between herds, affecting animals (and herds) over a short time frame (Madouasse et al., 2013). In contrast, the approach will have a reduced sensitivity for diseases which are transmitted slowly within and between herds.

Finally, whilst our study was designed with a reference period for model fitting and a 'future' prediction window for alarm detection to reflect how such an approach would be used in practice, our approach was nonetheless conducted on an existing dataset, and therefore retrospective in nature. The potential for this system to be deployed for real-time detection would depend on the ability for model predictions to be compared with observed data in as close to real-time as possible. The implementation of such a model in practice therefore presents challenges that would require multi-stakeholder engagement to address. There are likely to be delays of at least 1–2 days before milk recording data are collected and reported. Predictions would then need to be generated for each animal in the herd according to animal-specific covariates, compared with observed data and subjected to CUSUM algorithms and spatial analyses. However, Ireland is well positioned to deploy such a model given the existence of a centralised (ICBF) database containing the majority of all herd- and animal-level data in the country. The specific covariates required for prediction are therefore already present in the same database in which milk recordings are reported. Secondary analyses based on CUSUM and spatial scanning, could potentially be deployed within the ICBF system, with outputs summarised on regular intervals (e.g. weekly) for those responsible for disease surveillance (e.g. the Department of Agriculture Food and the Marine in Ireland), presenting the opportunity for delays to be minimised as much as possible.

In conclusion, this system for monitoring milk yield represents both a potentially useful ancillary tool for early detection of disease, and a valuable foundation for developing similar tools using other metrics.

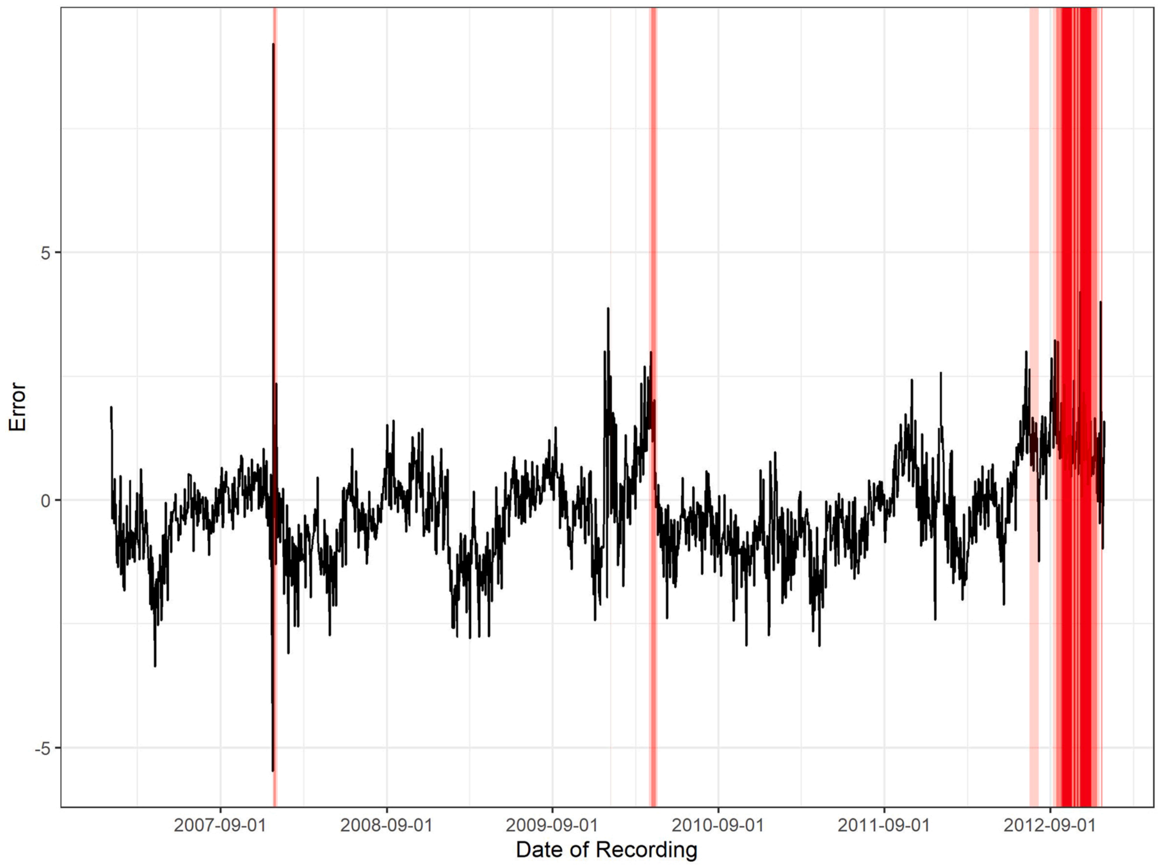

Fig. 1. Average daily prediction error (predicted – observed) and corresponding alarm periods for the reference (prior to 2012) and testing (2012) periods. Alarm periods generated using the 3k, 5k and 7k thresholds are shaded in light, medium and darker shades of red respectively.

Fig. 2. Hotspot analyses for August and September 2012 using the Getis-Ord Gi* statistic. Data are aggregated to hexagonal lattice. Hexagons with less than 5 farms have been removed.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Funding: This study was funded by the Irish Department of Agriculture Food and the Marine (RSF 17/S/230). The authors are very grateful to the Irish Cattle Breeding Federation (ICBF) for providing production data used in this study.

References

Afonso, A., Abrahantes, J.C., Conraths, F., Veldhuis, A., Elbers, A., Roberts, H., Van der Stede, Y., Méroc, E., Gache, K., Richardson, J., 2014. The Schmallenberg virus epidemic in Europe—2011–2013. Prev. Vet. Med. 116 (4), 391–403.

Bates, D., Maechler, M., Bolker, B., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1), 1–48.

Bivand, R.S., Pebesma, E, Gomez-Rubio, V., 2013. Applied spatial data analysis with R, Second ed. Springer, NY. ⟨https://asdar-book.org/⟩.

Bradshaw, B., Mooney, J., Ross, P.J., Furphy, C., O'Donovan, T., Sanchez, C., Gomez-Parada, M., Toolan, D., 2012. Schmallenberg virus cases identified in Ireland. Vet. Rec. 171 (21), 540–541.

Caminade, C., McIntyre, K.M., Jones, A.E., 2019. Impact of recent and future climate change on vector-borne diseases. Ann. N. Y. Acad. Sci. 1436 (1), 157.

Carty, C.I., Fahey, A.G., Sheehy, M.R., Taylor, S., Lean, I.J., McAloon, C.G., O'Grady, L., Mulligan, F.J., 2017. The prevalence, temporal and spatial trends in bulk tank equivalent milk fat depression in Irish milk recorded herds. Irish Vet. J 70 (1), 1–14.

Collins, Á.B., Doherty, M.L., Barrett, D.J., Mee, J.F., 2019. Schmallenberg virus: a systematic international literature review (2011-2019) from an Irish perspective. Irish Vet. J. 72 (1), 1–22.

Courtejoie, N., Durand, B., Bournez, L., Gorlier, A., Bréard, E., Sailleau, C., Vitour, D., Zientara, S., Baurier, F., Gourmelen, C., Benoit, F., Achour, H., Milard, C., Poliak, S., Pagneux, C., Viarouge, C., Zanella, G., 2018. Circulation of bluetongue virus 8 in French cattle, before and after the re-emergence in 2015. Transbound. Emerg. Dis. 65, 281–284.

Elgersma, G.G., de Jong, G., van der Linde, R., Mulder, H.A., 2018. Fluctuations in milk yield are heritable and can be used as a resilience indicator to breed healthy cows. J. Dairy Sci. 101 (2), 1240–1250.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1.

Getis, A., Ord, J.K., 1992. 1992: The analysis of spatial association by use of distance statistics. Geogr. Anal. 24, 189–206.

Hurtado-Uria, C., Hennessy, D., Shalloo, L., O'Connor, D., Delaby, L., 2013. Relationships between meteorological data and grass growth over time in the south of Ireland. Irish Geogr. 46 (3), 175–201.

Jeffery, C., Ozonoff, A., Pagano, M., 2014. The effect of spatial aggregation on performance when mapping a risk of disease. Int. J. Health Geogr. 13 (1), 1–9.

Kulldorff, M., Nagarwalla, N., 1995. Spatial disease clusters: detection and inference. Stat Med. 14 (8), 799–810.

Madouasse, A., Marceau, A., Lehébel, A., Brouwer-Middelesch, H., van Schaik, G., Van der Stede, Y., Fourichon, C., 2013. Evaluation of a continuous indicator for syndromic surveillance through simulation. Application to vector borne disease emergence detection in cattle using milk yield. PLoS One 8 (9), e73726.

McGrath, G., More, S.J., O'Neill, R., 2018. Hypothetical route of the introduction of Schmallenberg virus into Ireland using two complementary analyses. Vet. Rec. 182 (8).

Meiswinkel, R., Baldet, T., de Deken, R., Takken, W., Delécolle, J.C., Mellor, P.S., 2008. The 2006 outbreak of bluetongue in northern Europe—The entomological perspective. Prev. Vet. Med. 87, 55–63.

Menzies, F.D., McCullough, S.J., McKeown, I.M., Forster, J.L., Jess, S., Batten, C., Murchie, A.K., Gloster, J., Fallows, J.G., Pelgrim, W., Mellor, P.S., 2008. Evidence for transplacental and contact transmission of bluetongue virus in cattle. Vet. Rec. 163 (7), 203–209.

Murray, G.M., Fagan, S., Murphy, D., Fagan, J., Ó Muireagáin, C., Froehlich-Kelly, R., Barrett, D.J., Sheehan, M., Wilson, M., Brady, C.P., Hynes, F., 2019. Descriptive analysis of ovine mortality in sentinel sheep flocks in Ireland. Vet. Rec. 184 (21).

Ord, J.K., Getis, A., 1995. Local spatial autocorrelation statistics: distributional issues and an application. Geogr. Anal 27 (4), 286–306.

Pebesma, E., 2018. Simple features for r: standardized support for spatial vector data. R J. 10 (1), 439–446. https://doi.org/10.32614/RJ-2018-009.

Poskin, A., Théron, L., Hanon, J.B., Saegerman, C., Vervaeke, M., Van der Stede, Y., Cay, B., De Regge, N., 2016. Reconstruction of the Schmallenberg virus epidemic in Belgium: complementary use of disease surveillance approaches. Vet. Microbiol. 183, 50–61.

Purse, B.V., Mellor, P.S., Rogers, D.J., Samuel, A.R., Mertens, P.P., Baylis, M., 2005. Climate change and the recent emergence of bluetongue in Europe. Nat. Rev. Microbiol. 3 (2), 171–181.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). J. R. Statist. Soc. B 71, 319–392.

Tango, T., Takahashi, K., 2005. A flexibly shaped spatial scan statistic for detecting clusters. Int. J. Health Geogr. 4 (1), 1–15.

Veldhuis, A., Brouwer-Middelesch, H., Marceau, A., Madouasse, A., Van der Stede, Y., Fourichon, C., Welby, S., Wever, P., van Schaik, G., 2016. Application of syndromic surveillance on routinely collected cattle reproduction and milk production data for the early detection of outbreaks of Bluetongue and Schmallenberg viruses. Prev. Vet. Med. 124, 15–24.