



Validation of a spatial liver fluke model under field conditions in Ireland

Amalia Naranjo Lucena,¹ María Pía Munita Corbalán,² Ana María Martínez-Ibeas,²
Guy McGrath,³ Riona Sayers,² Grace Mulcahy,¹ Annetta Zintl¹

¹*School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4;* ²*AGRIC, Teagasc, Moorepark, Fermoy, Co. Cork;* ³*UCD Centre for Veterinary Epidemiology and Risk Analysis, Belfield, Dublin 4, Ireland*

Abstract

Fasciola hepatica is the causative agent of fasciolosis, a global disease of a wide range of mammals, particularly sheep and cattle. Liver fluke infection causes annual losses estimated at around €2.5 billion to livestock and food industries worldwide. Various models have been developed to define risk factors and predict

exposure to this liver fluke in ruminants in European countries, most of them based exclusively on data from dairy herds. The aim of this study was to validate a published theoretical baseline risk map of liver fluke exposure and cluster maps in Ireland, by including further explanatory variables and additional herd types that are spatially more widespread. Three approaches were employed: i) comparison of predicted and actual exposure; ii) comparison of cluster distribution of hotspots and coldspots; and iii) development of a new model to compare predicted spatial distribution and risk factors. Based on new survey data, the published baseline predictive map was found to have a sensitivity of 94.7%, a specificity of 5%, a positive predictive value of 60% and a negative predictive value of 38.2%. In agreement with the original model, our validation highlighted temperature and rainfall among the main risk factors. In addition, we identified vegetation indices as important risk factors. Both the previously published and our new model predict that exposure to *Fasciola* is higher in the western parts of Ireland. However, foci of high probability do not match completely, nor do the location of clusters of hotspots and coldspots.

Correspondence: Amalia Naranjo Lucena, Pathobiology Section, School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland.
Tel.: +35317166138.
E-mail: amalia.naranjo-lucena@ucdconnect.ie

Key words: Validation; Fasciolosis; Hotspot analysis; Spatial modelling; Ireland.

Acknowledgements: the authors would like to thank Teagasc for providing the farm data to complete this study, which were collected for the Flukeless research project. We also thank Met Éireann and Aidan Murphy for supplying the weather data and Avia-GIS for supplying vegetation data, as well as for the training received and valuable assistance with the modelling process. Moreover, we thank the Centre for Veterinary Epidemiology and Risk Analysis (CVERA) for advice on the various analyses. The habitat map was supplied by Teagasc. The CORINE 2012 Land Cover, soils and subsoils datasets were sourced from EPA. Electoral divisions maps were sourced from the Central Statistics Office (CSO) of Ireland.

Contributions: the authors contributed equally.

Conflict of interest: the authors declare no potential conflict of interest.

Funding: this project is being funded under the European Union Horizon 2020 programme (PARAGONE: vaccines for animal parasites. H2020-EU.3.2. SOCIETAL CHALLENGES).

Received for publication: 24 October 2017.
Revision received: 21 November 2017.
Accepted for publication: 28 November 2017.

©Copyright A. Naranjo Lucena, et al., 2018
Licensee PAGEPress, Italy
Geospatial Health 2018; 13:641
doi:10.4081/gh.2018.641

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (CC BY-NC 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Introduction

Fasciolosis is a continuing threat to the sustainability of European agriculture. The parasite is highly prevalent and increasing numbers of anthelmintic resistant isolates are reported worldwide (Kelley *et al.*, 2016). Furthermore, the effects of climate change are predicted to increase the incidence of fasciolosis in north-western and central Europe (Caminade *et al.*, 2015). For example in the UK, sub-acute fasciolosis is increasingly diagnosed in areas where it was not present before, as the parasite has expanded its transmission range (van Dijk *et al.*, 2010; Fox *et al.*, 2011). In Ireland, several studies have shown very high prevalence rates in dairy herds (76.1%-82%) and sheep flocks (61.6%) (Mooney *et al.*, 2009; Bloemhoff *et al.*, 2015; Rinaldi *et al.*, 2015; Selemetas *et al.*, 2015a).

Due to the long grazing season, the specific landform and climate of Ireland render it highly suitable for livestock farming (O'Mara, 2008). Ireland is one of the main exporters of beef and dairy products in the European Union (EU), with 90% of beef cattle, 75% of dairy products and 60% of sheep meat produced for export mainly within the EU, but also to international markets (O'Mara, 2008; Browne, 2010). Given the focus on grass-based production and the mild and humid Irish climate, infection with helminth parasites is a significant threat in the Irish production system. With a prevalence of infection of 82% in dairy herds (Selemetas *et al.*, 2015a), fasciolosis caused by the liver fluke *Fasciola hepatica* is one of the main parasites affecting Irish livestock associated with significant economic losses. The life cycle

of the parasite involves mud snails as intermediate hosts. In fact, distribution of *F. hepatica* is largely dependent on the presence of an appropriate habitat for the snail intermediate host, which in turn is reliant on climatic and environmental conditions. In Europe, *Galba truncatula* represents the main intermediate snail host. This amphibious freshwater snail lives at the edges of small water bodies with loamy, muddy or clayey ground at pH values between 5 and 9 and 10–25°C temperature (Deplazes *et al.*, 2016).

The definitive host of *F. hepatica* is usually a herbivore, although humans can also act as definitive hosts. Animals become infected by ingesting encysted metacercariae from pasture. Juvenile flukes excyst in the duodenum and then penetrate the gut wall and peritoneal cavity and move towards the liver. Once there, they migrate through the parenchyma to mature and become established in the bile ducts, producing eggs that are shed with the faeces into the environment. Miracidia hatch from the eggs and penetrate the foot of the intermediate snail host where they develop through the redia and sporocyst stages to cercariae, which complete the life cycle by encysting on herbage as metacercariae (Skuce and Zadoks, 2013; Deplazes *et al.*, 2016; Taylor *et al.*, 2016). Management factors such as herd size, flukicide treatment protocols and length of grazing season have been shown to be of importance in the epidemiology of fasciolosis (Bennema *et al.*, 2011; Howell *et al.*, 2015; Olsen *et al.*, 2015; Munita *et al.*, 2016).

Sudden deaths due to acute fasciolosis occur in sheep, when the infective burden is high (Deplazes *et al.*, 2016; Taylor *et al.*, 2016). In cattle, where *F. hepatica* most frequently follows a chronic course, infections are associated with increased calving interval, delayed puberty, reduced milk yield, reduced milk fat content, lower body condition scores and carcass cold weight (López-Díaz *et al.*, 1998; Schweizer *et al.*, 2005; Charlier *et al.*, 2007; Sanchez-Vazquez and Lewis, 2013; Skuce and Zadoks, 2013). This, together with liver condemnations in abattoirs further increases the economic costs due to fasciolosis, which can reach up to €90 million every year in Ireland (Animal Health Ireland, 2013).

The emergence of *F. hepatica* strains resistant to triclabendazole (the only chemical that kills early immature, immature and adult flukes) and the difficulty inherent in developing a vaccine against the parasite are making future prospects for the control of this infection challenging (Mooney *et al.*, 2009; Toet *et al.*, 2014; Hanna *et al.*, 2015; Molina-Hernández *et al.*, 2015). Hence, there is interest in developing detailed, local predictive high-resolution models in order to improve control methods and minimise losses. Geographical Information Systems (GIS) and Remote Sensing (RS) have become valuable tools for investigating spatial patterns of disease and supporting decision-making with regard to selecting the most appropriate approaches. Baseline risk and cluster maps for Ireland were published within the scope of the EU project GLOWORM (funded by the European Commission's Seventh Framework Programme FP7–KBBE-2011-5 under grant agreement no. 288975), which investigated global changes in the emergence of drug resistance, environmental and climatological conditions, and their effects on ruminant livestock helminth infections through the use of GIS and RS for mapping and modelling (Selemetas and de Waal, 2015; Selemetas *et al.*, 2015a). Irish GLOWORM studies were based on dairy herds concentrated in the South-west of the country. The main objective of this new study was to develop a more inclusive and widespread spatial analysis of the epidemiology of fasciolosis in Ireland. We aimed to re-evaluate the maps previously published by Selemetas *et al.*, (2015a); and Selemetas and de Waal, (2015) using survey data from dairy, beef

and sheep farms with a more widespread geographical distribution compared to that employed in the original study.

Materials and Methods

Study area and data provision

Ireland has an area of 69,825 km² with a climate characterized by cool summers and mild winters and defined as temperate oceanic climate, which means mild and moist with environmental and climatologic conditions that can vary greatly over short distances.

The Agriculture and Food Development Authority in Ireland, officially referred to as Teagasc, is responsible for research and development, training and advisory services in the agri-food sector. The authority has a number of county advisory centres, colleges and research laboratories all over the country.

Data were collected during three surveys conducted between October 2014 and February 2015 (Flukeless research project, Research Stimulus Fund under project no 13/S/405) and made available by the Animal and Grassland Research and Innovation Centre (AGRIC), Teagasc Moorepark in Fermoy, Cork County, Ireland.

Sampling strategy

For dairy herds, a target sample size of 318 herds was calculated based on a predicted prevalence of 70% and a dairy population of 18,500 herds [95% confidence limit (CL), 5% precision (CI)]. The 319 herds included in the project by Bloemhoff *et al.*, (2015) were initially targeted, as these geographically represent the Irish national dairy population (O'Doherty *et al.*, 2013; Bloemhoff *et al.*, 2015). Herds were selected from the HerdPlus® database containing 3,500 dairy herds which is a breeding information decision support tool coordinated by the Irish Cattle Breeding Federation (ICBF). Herds were selected using a stratified sampling procedure based on herd size and geographical location. A total of 322 flocks (predicted prevalence of 70%, population 33,500 flocks, 95% CL, 5% CI) and 325 beef herds (predicted prevalence 70%, population 100,000 herds, 95% CL, 5%CI) were targeted. Applications for sheep and beef farms were distributed through Teagasc Newsletters; which is a monthly publication produced by Teagasc Advisory Services. A stratified sampling strategy, based on herd size and geographical location, was also used for selection of these farms. To increase the study power, an additional 44 dairy herds, applied through Teagasc Newsletters, were randomly selected to take part in the study. The participation of herds in the current study was on a volunteer and non-incentive basis. All sampled farms had just one piece of land where their livestock was kept, *i.e.* all animals experienced the same environmental and climatologic conditions.

Farms from which incorrect or incomplete geographical coordinates were noted were not included. Consequently, data from 312 dairy herds, 194 beef herds and 290 sheep flocks were finally used (Figure 1A). Exposure to *F. hepatica* in the various operations was investigated as follows. For dairy herds, bulk tank milk (BTM) samples were collected and tested for specific anti-*F. hepatica* antibodies by the Ildana Biotech *F. hepatica* recombinant mutant Cathepsin L1 (CL1) ELISA (<http://www.ucd.ie/conway/innovation/ildanabiotech/>). With this kit antibodies start to be detectable between 2 and 4 weeks after infection, and decline 3 to

4 months after flukicide treatment (Bloemhoff *et al.*, 2015). The tests were carried out as per manufacturer's instructions and previously described by Bloemhoff *et al.* (2015). The ELISA sample-to-positive (S/P) ratio was calculated for each sample by first subtracting the optical density (OD) reading of the CL1 uncoated well from the reading of the coated well. Subsequently, the ratio of the sample OD to the positive control was calculated (S/P ratio). Herds with an S/P ratio higher than 15 were considered as positive, whereas herds with an SP ratio of 15 or less were considered negative. For beef herds, serum samples from at least 6 animals per herd were tested using the same ELISA kit. Average SP ratios per herd was calculated. Those herds with one serum sample value higher than 20 were considered a positive herd, and those with all sera samples with values of 20 or lower were considered negative herds (Bloemhoff *et al.*, 2015). Finally, data produced by Martinez-Ibeas *et al.*, (2016) were used for exposure in sheep flocks (Martinez-Ibeas *et al.*, 2016). Briefly, a standardised sampling kit was posted to each farmer with a request to submit 20 fresh faecal catch samples from 20 different mature ewes. Faecal egg counts (FEC) of liver fluke eggs were determined using the sedimentation technique (Taylor *et al.*, 2016). Flocks where at least one liver fluke egg was observed were considered positive. Herd status data based on these surveys are shown in Figure 1B.

Comparison of predicted probability of exposure and actual exposure

Mapping and georeferencing

Using ArcGIS version 10.1 (ESRI, Redlands, CA, USA) geo-

graphical coordinates and exposure/infection status of all surveyed farms were mapped, and a shapefile created. Since the prediction by Selemetas *et al.*, (2015a) was based on district electoral division levels (DEDs), the shapefile for the new farms was also converted from point locations to DED polygons. DEDs are the smallest legally defined administrative areas in the state. A DED that contained at least one positive farm was considered positive. Overall exposure status was recorded for 658 DEDs.

The red, green and blue (RGB) image file of predicted probability (image 6 in Selemetas *et al.*, 2015a) was geo-referenced. The probability for each DED was extracted by using the *feature to point* tool in ArcGIS with the parameter set to INSIDE.

Comparison of risk map with field surveys

Sensitivity, specificity as well as positive and negative predictive values of the risk map by Selemetas *et al.* (2015a) were calculated by comparing the predicted probability of exposure in each DED (for the year 2012) with the data collected in the three surveys carried out in 2014 and 2015. True positive, true negative, false positive and false negative DEDs were obtained by joining the prediction map with the survey map. The result was exported to an excel file to determine number of DEDs with presence or absence as follows: True positives: prediction +/ survey +, true negatives: prediction -/ survey -, false negatives: prediction -/ survey +, false positives: prediction +/ survey -. Sensitivity was calculated as: true positives/ true positives + false negatives, specificity as: true negatives/ true negatives + false positives, positive predictive value as: true positives/ true positives + false positives, and negative predictive value as: true negatives/ true negatives + false negatives.

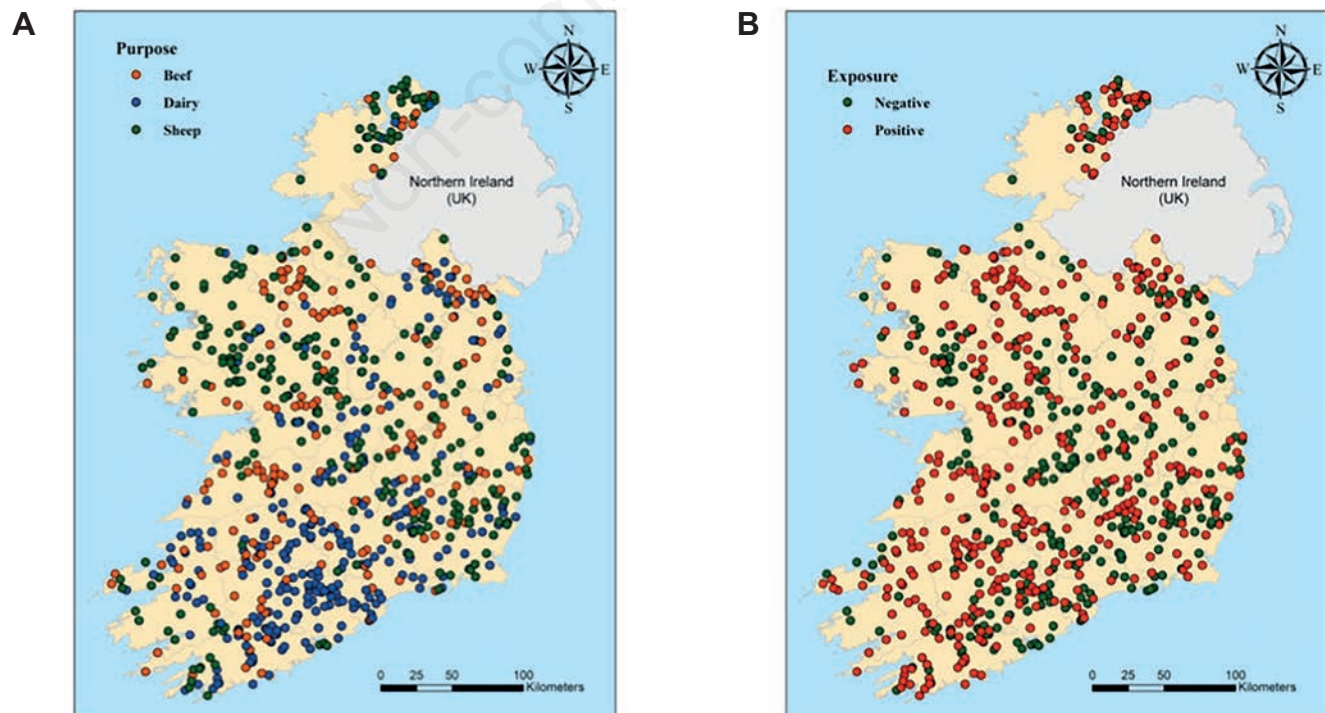


Figure 1. Location of farms from AGRIC Teagasc Moorepark surveys. A) Type of farms; B) Exposure.

Hotspot analysis

Clusters of exposure were identified using the Getis-Ord G_i^* statistic (Getis and Ord, 1992; Ord and Getis, 1995) based on the location of the 452 observed liver fluke-positive herds from the three surveys mentioned above. The Getis-Ord G_i^* statistic can be used to determine whether the spatial clustering of high or low values is more pronounced than expected in a random distribution of those same values. Z-scores values, the outputs of the statistic, represent standard deviations for clustering intensity. Different significance levels (P-values) apply to specific z-scores values, so that P-values of <0.1, <0.05 and <0.01 correspond to z-scores of <-1.65 or >+1.65, <-1.96 or >+1.96 and <-2.58 or >+2.58, respectively.

As DEDs are artificial geographical divisions, it was decided to group herds by proximity instead of DED for this analysis. As a result, two herds located in different DEDs but very close to each other, were grouped together. This gave a higher resolution to the analysis and was more logical biologically. In addition, Irish herds are generally relatively small (81 animals on average). For these reasons, and to minimize edge effects, small distance thresholds were studied to aggregate herds spatially. The distances used were: 1, 3, 5 and 10 km. Global Moran's I statistic was applied to test spatial auto-correlation at increasing distances by using the *incremental spatial autocorrelation* tool in ArcGIS. Statistically significant peaks of clustering were only detected on the dataset that had been aggregated by 3 km (Figure 2), which was the final dataset employed for the hotspot analysis (Getis-Ord G_i^*) performed in ArcGIS. The critical distance was set to 46,148.26 m, the first peak given by the Global Moran's I statistic (Figure 2). Finally, an inverse distance weighting (IDW) interpolation was performed using the Z score of each point for visualization purposes.

Modelling

Data from the AGRIC Teagasc Moorpark surveys were used to develop a new predictive model that was subsequently compared with the older, published model. Main risk factors identified by the two models were also compared.

Presence/absence data

For the development of the new risk map using the 2014/2015 survey data, the exact location of each farm (n=796) rather than the DED was used. The shapefile containing geographical coordinates

and presence/absence information (1/0) was projected to the WGS1984 geographical coordinate system using ArcGIS.

Explanatory variables

Table 1 shows the variables used in the model. Habitat, land cover and soil-related variables are categorical, whereas rainfall, temperature, elevation and vegetation are numerical variables. Vegetation indices, such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI), use near infrared and red wavelengths reflected by the vegetation to compute values that quantify plant biomass and/or vigour for each pixel in a RS image. These values vary between -1.0 and 1.0 and can be used to estimate vegetation amount or damage, *i.e.* infections, defoliation *etc.* (Cringoli *et al.*, 2004; Weier and Herring, 2000). Both are calculated similarly, although EVI corrects for some distortions by particles in the air or the ground cover below, and does not become saturated when analysing highly vegetated areas (Weier and Herring, 2000).

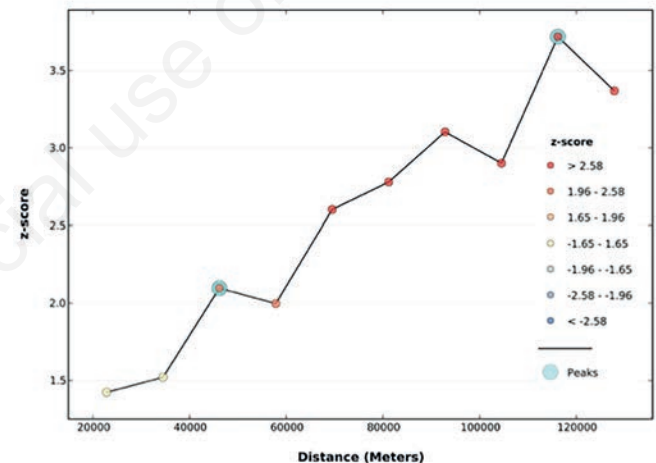


Figure 2. Spatial autocorrelation for the dataset grouped by 3 km distance. Produced by Moran's I statistic showing intensity of spatial clustering at increasing distances. Peaks represent distances where the spatial processes promoting clustering are most pronounced.

Table 1. Data-layers included in the modelling approach.

Variable	Description	Source and resolution
Period 2010-2015 climatic variables	Averages of annual, seasonal and monthly mean temperatures (°C), total rainfall (mm) and annual total number of rain-days (daily rainfall ≥ 0.2 mm)	Met Éireann ^a (1×1 km) interpolated values
Year 2014 climatic variables	Averages of annual, seasonal and monthly mean temperatures (°C), total rainfall (mm) and annual total number of rain-days (daily rainfall ≥ 0.2 mm)	Met Éireann ^a (1×1 km) interpolated values
Soils, subsoils and soil drainage	National Soil Database	EPA ^b (scale 1:250,000)
Habitat	National Habitat Indicator Map	Teagasc (25×25 m)
Land Cover	2012 CORINE ^c land cover dataset	EPA (25 ha minimum mapping unit)
Elevation	National Elevation Map	UCD ^d maps and GIS library. Processed by CVERA ^e (25×25 m).
Vegetation	NDVI ^f and EVI ^g from 2014	Avia-GIS (250×250 m)

^aThe Irish National Meteorological Service; ^bEnvironmental Protection Agency; ^cCoordination of information on the environment (provided by the European Environment Agency (EEA)); ^dUniversity College Dublin; ^eCentre for Veterinary Epidemiology and Risk Analysis; ^fMonthly normalized difference vegetation index; ^genhanced vegetation index.



The various data files were projected to the WGS1984 geographical coordinate system and converted to raster file format using either the *point to raster* or *feature to raster* tool in ArcGIS. All final rasters had the same number of columns and rows (494, 258), the same cell size (0.016) and were saved in the same folder for subsequent analysis with VECMAP (<http://www.vecmap.com/>), version 1.5.16209.2382.

Random Forest methodology

The training dataset used to build the model was first *balanced* using the *Balance/code dataset* tool in VECMAP. To create a reliable model, it is recommended to include an even number of presence and absence points in the input. This is done to avoid creating a bias in the model prediction. Predictor values were extracted using the *extract* tool.

As different predictive methodologies have been shown to vary in their predictions even when using the same datasets (Elith and Graham, 2009), the same algorithm employed by Selemetas *et al.*, (2015a) – Random Forest (RF) – was employed (Breiman, 2001; Selemetas *et al.*, 2015a). RF is a machine-learning algorithm that uses classification and regression trees (CART) where each tree is constructed from a random sample of cases, and a randomly selected subset of available variables are tested to split each node into *child nodes* creating a tree. To assess variable importance, the RF tool in VECMAP employs the Gini index (Breiman *et al.*, 1984). This index uses the impurity criterion as a measure of how each variable contributes to the homogeneity of the *nodes* and *leaves*. In RF, it is used to search for the split that most reduces node and tree impurity. Important variables lead to higher reduction of the Gini index when comparing original and child nodes.

A preliminary model was generated to select potentially important variables to cluster the data. Selected variables included EVI, temperature and rainfall variables from 2014 and soil type. Then, using the 106 predictors, RF was performed in the *classification mode* (presence/absence). Data scaling was set to probability scale (presence/absence). A variable reduction forest was initially performed to identify the most important variables. Subsequently, a second RF was created including only a 50% subset of the most significant variables, which improves model performance. In VECMAP, model validation was carried out for each tree by using the points not sampled from the full training set. These *out-of-the-bag* (OOB) samples are classified. Comparison of the observed and predicted results permits the generation of accuracy statistics.

The following default settings were used: reduction forest (100 points sampled for each tree, 600 trees to grow and 10 variables at each node), prediction forest (100 points samples for each tree, 400 trees to grow and 10 variables used at each node). Each node was split into two child nodes. The output raster file obtained was projected to TM65/Irish Grid using VECMAP.

Results

Validity of the previous liver fluke distribution prediction

Table 2 shows the number of true positive, true negative, false positive and false negative DEDs resulting from the comparison of the predictive map and the 2014/2015 survey data, as well as results for sensitivity, specificity, positive predictive value and negative predictive value.

The predicted distribution was tested in two ways; first by only considering DEDs in the 1st category (predicted probability 0–15.6% according to Selemetas *et al.*, (2015a) as negative, and secondly by considering all DEDs in categories 1 and 2 as negative (*i.e.* including predicted probability of up to 42.8%), and only categories 3 to 10 as positive. The results indicated that overall sensitivity of the predictive map by Selemetas *et al.* (2015a) was very high (>0.9), while its specificity was very low (<0.1). The positive and negative predictive values, ranging between 0.38 and 0.6, were considered moderate in both comparisons.

Hotspot analysis

The analysis revealed that 7 groups of herds represented statistically significant hotspots at the 99% confidence level, 18 at the 95% level and 10 at the 90% level (Figure 3A). Conversely, none of the groups represented statistically significant coldspots at the 99% confidence level, and only 4 and 7 were significant at the 95% and 90% level, respectively. Regarding the geographic distribution of clustering, significant hotspots were mainly observed in the South-west and North-west of the country, whereas significant coldspots were located in the Mid-west and Mid-east.

Spatial prediction

The predicted spatial distribution of fasciolosis is shown in Figure 3B. Overall, probability of exposure is higher in the western part of the country. A cluster of highly predicted probabilities of exposure was found in the North-west. Areas with higher elevation have a lower predicted probability of exposure, although elevation was not included among the main risk factors. The model performance was assessed by Cohen's Kappa (0.52) with values of >0.4 regarded as acceptable, area under the curve (AUC) (0.69), sensitivity (0.67) and specificity (0.61) (Landis and Koch, 1977).

Risk factors

According to the mean decrease of the Gini Index, the 10 most important variables (in order of importance) were: annual mean temperature in 2014 and for the period 2010–2015, mean temperature in September 2014, vegetation in April 2014, mean

Table 2. DED parameters according to the survey data when compared to the prediction map.

	True positives	True negative	False positives	False negatives	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Only 1 st class negative	375	13	21	249	0.95	0.05	0.6	0.38
1 st and 2 nd classes negative	365	23	31	239	0.9	0.09	0.6	0.43

temperature in spring and autumn for the period 2010–2015, vegetation in March and January 2014, and mean rainfall during June and July 2014.

Discussion

This study aimed to refine and validate a published model which chiefly considered dairy herds located in the Munster region (South-West) (Selemetas *et al.*, 2015a) by including data from dairy herds, beef herds and sheep flocks from a total of 796 farms from 658 DEDs (just over 19% out of a total of 3,409 DEDs in the Republic of Ireland) located throughout the country. It was hoped that inclusion of diverse farm types would reduce bias of the modelled prediction due to differences in management procedures between them. For instance, the length of grazing season is usually longer in sheep flocks than cattle herds as some flocks are not housed, or are housed for shorter periods during the winter. Dairy cattle in Ireland are usually housed from November to March, although substantial regional differences in grazing practices have been reported (Bloemhoff *et al.*, 2014). Similarly, housing periods for beef cattle range between 3 and 9 months (Teagasc, 2015). In addition, timing and type of flukicide treatment also vary between dairy, beef and sheep enterprises. Dairy herds are usually dosed (with a restricted number of products licensed) once yearly, *i.e.* during the dry season which is generally coincident with winter housing (Bloemhoff *et al.*, 2014). Flukicide treatments in beef herds and sheep flocks are not as restrictive, although meat with-

drawal periods apply. In this case, rafoxanide is also available, and animals are often dosed twice a year (Parr and Gray, 2000; Teagasc, 2015, 2016).

In our study, ELISA applied to BTM or serum was used to characterize disease exposure in dairy and beef herds, respectively. For sheep flocks, FEC data were used. Although this added an extra level of variability it was felt that including these data provided a greater number of true positive results and broadened the scope of the study.

A comparison of the probability of exposure predicted by Selemetas *et al.* (2015a) to the survey data revealed a high sensitivity (0.95, or 0.92) but low specificity (0.05, or 0.09) of the original predictive model, *i.e.* the model was good at predicting positive DEDs but failed to predict negative ones (Selemetas *et al.*, 2015a). Moreover, hotspot analysis based on the recent survey data indicated a slightly different geographical pattern with respect to that published by Selemetas and de Waal, (2015). While results for clustering obtained from the north-western region (Figure 3A) are similar for both studies, especially for those areas close to the border with Northern Ireland (UK), with z-scores between +1.65 and +2.58, our study identified two additional significant hotspot clusters in the northernmost region of the Republic of Ireland. Furthermore, our results indicated significant clustering further inland in the South-west (Munster) than what had been obtained by Selemetas and de Waal, (2015) who located them closer to the western coast. Interestingly, when Selemetas and colleagues performed a similar analysis exclusively considering the Munster Province (where most of the samples originated), the pattern

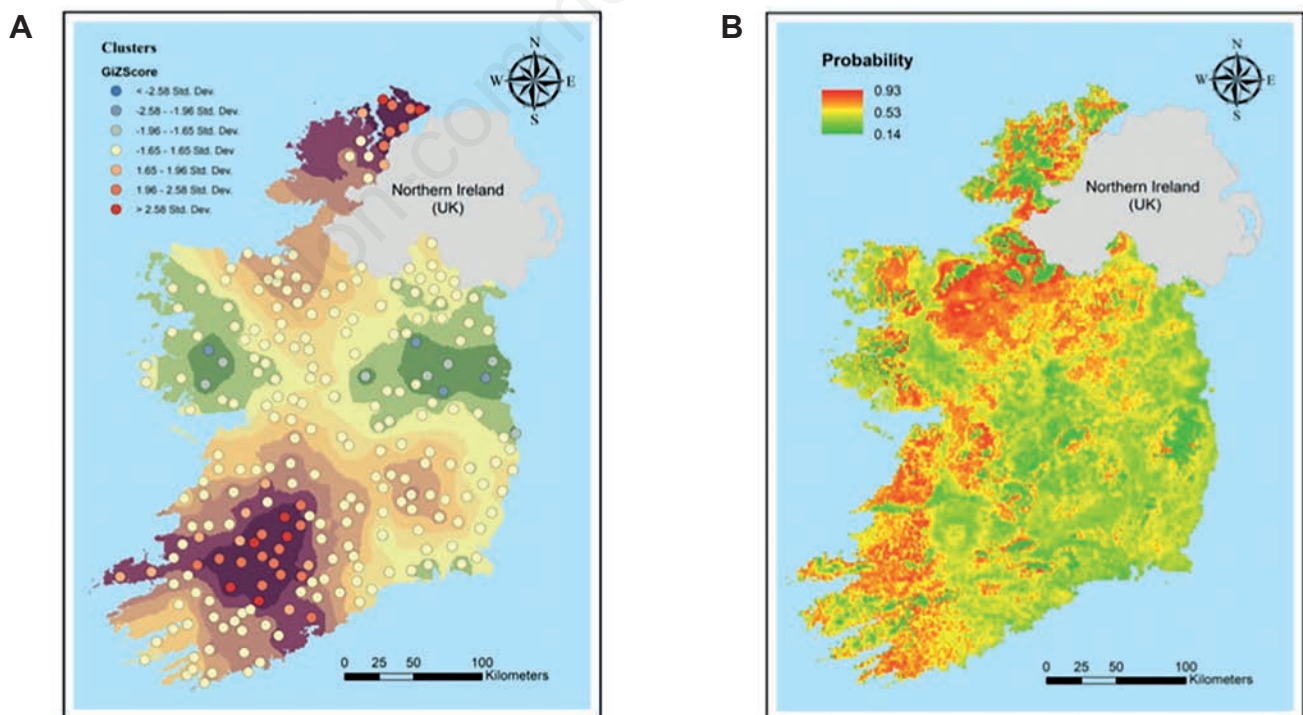


Figure 3. Spatial analysis of dairy, beef and sheep exposure to *Fasciola hepatica* in Ireland. A) Hotspot analysis (Getis-OrdGi* statistic) for each cluster of exposure to *F. hepatica*; B) Spatial distribution of predicted probabilities of exposure to *F. hepatica* (spatial resolution: 1.2×1.2 km).



observed for high clustering was more similar to our outcome (Selemetas *et al.*, 2015b). Selemetas and de Waal (2015) described clusters of coldspots mainly on the southern coast and in the centre of the country, whereas our results located these mostly in the Mid-west and Mid-east.

With respect to predicted distributions, the two models agree in that probability of exposure is predicted to be highest in the West of the country. However, the Midlands-North region predicted to have higher probabilities in the older model, was found to be located further west in our study. Lower probabilities in areas with higher elevation can also be observed from the new risk map, although elevation was not one of the main risk factors for exposure. Irish highlands are generally covered by blanket bogs, a type of peat formed by waterlogged soil composed of compacted, partially decomposed, vegetable matter fed exclusively by rainwater. Due to the higher altitude, average temperatures in these regions are also lower. In the past, overgrazing of these areas by sheep caused severe damage to 7% of Ireland's total blanket bogs. However, changes in EU legislation resulted in significant reduction of sheep numbers (IPCC, 2016). Because of this, the number of sample points from these areas is very low in our study. However, as in theory, the model also predicts risk in areas where no sampling is done, we suggest that the reason for the low predicted probabilities in these regions is due to their environmental unsuitability for the intermediate host and/or parasite.

In the present model, climatic and environmental variables were useful predictors for the spatial probability of *Fasciola* exposure. This agrees with Selemetas *et al.* (2015a) who predicted annual rain days from the 1981-2010 period to be the main risk factor, followed by other rainfall variables and annual temperature from the year of sampling (2012). The main risk factors in our study were temperature, vegetation and precipitation, with the annual mean temperature being the best predictor. Other studies in Ireland, the UK and the rest of Europe also highlighted precipitation and temperature as the most important variables (McCann *et al.*, 2010; Ducheyne *et al.*, 2015; Selemetas and de Waal, 2015; Selemetas *et al.*, 2015a; Munita *et al.*, 2016). However, other environmental predictors such as elevation, soil type or snail habitat have also been found to be important (Bennema *et al.*, 2011; Charlier *et al.*, 2011; Ducheyne *et al.*, 2015). In our study, vegetation was included within the 10 best predictors. More specifically, the most important months with respect to vegetation were April, March and January in that order. The average EVI values of all locations for these months were 0.54, 0.48 and 0.42, respectively. In the spring months, the index is higher due to increased chlorophyll content in plants, which preferentially reflect near-infrared wavelengths (Weier and Herring, 2000; Cringoli *et al.*, 2004). The inclusion of vegetation variables within the most important risk factors in the new spatial model reflects the role of grazing in the transmission of infection, and may indicate better conditions for the development of the intermediate host. In fact in addition to adequate moisture, spring growth of *G. truncatula* depends on the quality and quantity of available food, which influences the mean growth of the shell (Dreyfuss *et al.*, 2015). It is worth mentioning that some Irish studies have also found that management aspects, such as smaller herd sizes and herds with heifers calving over 30 months of age (rather than <24 months), are relevant factors (Bloemhoff *et al.*, 2015; Munita *et al.*, 2016). In fact, McCann *et al.*, (2010) reported that climatic and environmental variables could only explain up to 70-76% of the variation in liver fluke infection pressure (McCann *et al.*, 2010).

The Ollerenshaw index is currently the most commonly used method to predict liver fluke incidence in Ireland (Ollerenshaw and Rowlands, 1959; de Waal *et al.*, 2007). It is a seasonal index derived from measured rainfall, number of rain days and potential evapotranspiration for a given month between May and October. Our results indicate that annual liver fluke forecasts could be refined by including vegetation indices. In addition, it is likely that climate change will impact the seasonality of the parasite as higher temperatures may increase the number of overwintering metacercariae on pastures (van Dijk *et al.*, 2010; Fox *et al.*, 2011; Caminade *et al.*, 2015). According to our results annual mean temperatures are the best predictors of liver fluke exposure in Ireland. Also, the fact that vegetation during January was included within the main risk factors might indicate the significance of the preceding winter season for forecasting fluke risk. A wider timeframe than that used by the Ollerenshaw index (May to October) should therefore be considered when forecasting the risk as suitable temperatures for the parasite and intermediate host may occur at other times during the year.

There are several potential reasons for differences in the outcome of our study and that of Selemetas *et al.*, (2015a) as indicated by the literature (Selemetas *et al.*, 2015a, 2015b; Selemetas and de Waal, 2015). Discrepancies between hot and cold cluster distributions may largely be due to the fact that our analysis was based on spatially dispersed sample points, while the previous study was chiefly focused on the South-west and had few data points from mid-western and eastern regions. Another important aspect is that between March 2010 and October 2012, *i.e.* when sampling for the Selemetas *et al.* (2015a) study was performed, there was a flukicide license restriction which meant that albendazole and oxclozanide were the only two licensed active ingredients for fluke control in dairy cows (Bloemhoff *et al.*, 2014). Both these drugs are only effective against mature *F. hepatica* stages (Animal Health Ireland, 2013; Williams *et al.*, 2014). After the restriction period other products were made available for pregnant dairy animals (though not cows during lactation), namely triclabendazole, closantel and nitroxynil (O'Brien *et al.*, 2010; IMB, 2013). Hence, it is likely that control measures adopted in dairy herds differed between the two studies. With regard to the explanatory variables used in both models, Selemetas *et al.* (2015a) employed variable averages from the period 1981-2010 as a background climatic dataset, as well as data for the year of sampling (2012) (Selemetas *et al.*, 2015a), whereas we used a background dataset (2010-2015) that spanned the actual dates when the survey herds were sampled, which was likely to result in a better performing model (McCann *et al.*, 2010). Furthermore, climatologic conditions each year also vary. In 2012, the spring and the summer average monthly rainfall according to the Irish National Meteorological Service Met Éireann's online datasets (recorded from 25 weather stations situated in various parts of the country) was 102.53 and 99.06 mm, respectively. However, during 2014, average rainfall during the spring and the summer was recorded to be only 65.08 and 60.60 mm, respectively, and the average monthly temperature values in the same seasons were between 0.8 and 1.7°C higher than in 2012. It is highly likely that these fluctuations resulted in annual differences in liver fluke prevalence (Ollerenshaw and Rowlands, 1959; Ollerenshaw, 1966). Finally the elevation and vegetation variables were not included by Selemetas *et al.* (2015a), but incorporated in the new model.

Conclusions

Our study aimed to investigate further the spatial patterns of liver fluke exposure in Ireland by evaluating a previously published predictive model and hotspot analysis. To our knowledge, this is the first study simultaneously addressing the spatial exposure of Irish dairy, beef and sheep farms to *F. hepatica* across the Republic of Ireland. The use of these data allowed us to detect new hotspots for liver fluke exposure, a different pattern of predicted distribution and further possible predictors. These results will help locate areas where monitoring for development of anthelmintic resistant strains as well as improved control measures may be considered. Furthermore, suggestions for the use of new possible predictors for future forecasting improvement are made.

References

- Animal Health Ireland, 2013. Liver Fluke: the facts. Parasite Control Leaflet Series, 5.
- Bennema SC, Ducheyne E, Vercruyse J, Claerebout E, Hendrickx G, Charlier J, 2011. Relative importance of management, meteorological and environmental factors in the spatial distribution of *Fasciola hepatica* in dairy cattle in a temperate climate zone. *Int J Parasitol* 41:225-33.
- Bloemhoff Y, Danaher M, Forbes A, Morgan E, Mulcahy G, Power C, Sayers R, 2014. Parasite control practices on pasture-based dairy farms in the Republic of Ireland. *Vet Parasitol* 204:352-63.
- Bloemhoff Y, Forbes A, Danaher M, Good B, Morgan E, Mulcahy G, Sekiya M, Sayers R, 2015. Determining the prevalence and seasonality of *fasciola hepatica* in pasture-based dairy herds in Ireland using a bulk tank milk ELISA. *Ir Vet J* 68:16.
- Breiman L, 2001. Random forests. *Mach Learn* 45:5-32.
- Breiman L, Friedman JH, Olshen RA, Stone CJ, 1984. Classification and regression trees. Wadsworth International Group, Belmont.
- Browne P, 2010. Irlanda. Available from: http://www.europeaes-pana.es/almac/_pdf/Repertorio_FP/Irlanda.pdf. Accessed: April 2018.
- Caminade C, van Dijk J, Baylis M, Williams D, 2015. Modelling recent and future climatic suitability for fasciolosis in Europe. *Geospat Health* 9:301-8.
- Charlier J, Bennema SC, Caron Y, Counotte M, Ducheyne E, Hendrickx G, Vercruyse J, 2011. Towards assessing fine-scale indicators for the spatial transmission risk of *Fasciola hepatica* in cattle. *Geospat Health* 5:239-45.
- Charlier J, Duchateau L, Claerebout E, Williams D, Vercruyse J, 2007. Associations between anti-*Fasciola hepatica* antibody levels in bulk-tank milk samples and production parameters in dairy herds. *Prev Vet Med* 78:57-66.
- Cringoli G, Taddei R, Rinaldi L, Veneziano V, Musella V, Cascone C, Sibilio G, Malone JB, 2004. Use of remote sensing and geographical information systems to identify environmental features that influence the distribution of paramphistomosis in sheep from the southern Italian Apennines. *Vet Parasitol* 122:15-26.
- de Waal T, Relf V, Good B, Gray J, Murphy T, Forbes A, Mulcahy G, 2007. Developing models for the prediction of fasciolosis in Ireland. *Mak Sci Work Farm* 60-3.
- Deplazes P, Eckert J, Mathis A, Samson-Himmelstjerna G, von Zahner H, 2016. Parasitology in veterinary medicine. 1st ed. Wageningen, NL: Wageningen Academic Publishers.
- Dreyfuss G, Vignoles P, Rondelaud D, Cabaret J, 2015. The mud snail and its biology, in: *The Mud Snail (Galba truncatula) ecology, parasitism and control*. LAP Lambert Academic Publishing, pp 3-26.
- Ducheyne E, Charlier J, Vercruyse J, Rinaldi L, Biggeri A, Demeler J, Brandt C, de Waal T, Selemetas N, Höglund J, Kaba J, Kowalczyk SJ, Hendrickx G, 2015. Modelling the spatial distribution of *fasciola hepatica* in dairy cattle in Europe. *Geospat Health* 9:261-70.
- Elith J, Graham CH, 2009. Do they? How do they? Why do they differ? on finding reasons for differing performances of species distribution models. *Ecogr Cop* 32:66-77.
- Fox NJ, White PCL, McClean CJ, Marion G, Evans A, Michael R, 2011. Predicting impacts of climate change on *fasciola hepatica* risk. *PLoS One* 6:19-21.
- Getis A, Ord JK, 1992. The analysis of spatial association by use of distance statistics. *Geogr Anal* 24:189-206.
- Hanna REB, McMahon C, Ellison S, Edgar HW, Kajugu PE, Gordon A, Irwin D, Barley JP, Malone FE, Brennan GP, Fairweather I, 2015. *Fasciola hepatica*: A comparative survey of adult fluke resistance to triclabendazole, nitroxylnil and closantel on selected upland and lowland sheep farms in Northern Ireland using faecal egg counting, coproantigen ELISA testing and fluke histology. *Vet Parasitol* 207:34-43.
- Howell A, Baylis M, Smith R, Pinchbeck G, Williams D, 2015. Epidemiology and impact of *Fasciola hepatica* exposure in high-yielding dairy herds. *Prev Vet Med* 121:41-8.
- IMB, 2013. Use of flukicides in dairy cattle.
- IPCC, 2016. Blankets Bogs. Available from: <http://www.ipcc.ie>. Accessed: October 2017.
- Kelley JM, Elliott TP, Beddoe T, Anderson G, Skuce P, Spithill TW, 2016. Current threat of triclabendazole resistance in *Fasciola hepatica*. *Trends Parasitol* 32:458-69.
- Landis JR, Koch GG, 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159-74.
- López-Díaz, CM, Carro CM, Cadórniga C, Díez-Baños P, Mezo M, 1998. Puberty and serum concentrations of ovarian steroids during prepuberal period in Friesian heifers artificially infected with *Fasciola hepatica*. *Theriogenology* 50:587-93.
- Martinez-Ibeas AM, Munita MP, Lawlor K, Sekiya M, Mulcahy G, Sayers R, 2016. Rumen fluke in Irish sheep: prevalence, risk factors and molecular identification of two paramphistome species. *BMC Vet Res* 12:143.
- McCann CM, Baylis M, Williams DJL, 2010. The development of linear regression models using environmental variables to explain the spatial distribution of *Fasciola hepatica* infection in dairy herds in England and Wales. *Int J Parasitol* 40:1021-8.
- Molina-Hernández V, Mulcahy G, Pérez J, Martínez-Moreno Á, Donnelly S, O'Neill SM, Dalton JP, Cwiklinski K, 2015. *Fasciola hepatica* vaccine: we may not be there yet but we're on the right road. *Vet Parasitol* 208:101-11.
- Mooney L, Good B, Hanrahan JP, Mulcahy G, de Waal T, 2009. The comparative efficacy of four anthelmintics against a natural acquired *Fasciola hepatica* infection in hill sheep flock in the west of Ireland. *Vet Parasitol* 164:201-5.
- Munita MP, Rea R, Bloemhoff Y, Byrne N, Martinez-Ibeas AM, Sayers RG, 2016. Six-year longitudinal study of *Fasciola hepatica* bulk milk antibody ELISA in the dairy dense region of



- the Republic Ireland. *Prev Vet Med* 134:16-25.
- O'Brien B, Jordan K, Danaher M, 2010. Update on the use of biologics in vasculitides. *Curr Pharm Biotechnol* 15:558-62.
- O'Doherty E, Sayers R, O'Grady L, 2013. Temporal trends in bulk milk antibodies to *Salmonella*, *Neospora caninum*, and *Leptospira interrogans* serovar hardjo in Irish dairy herds. *Prev Vet Med* 109:343-8.
- O'Mara F, 2008. Country Pasture/Forage Resource Profile. FAO Plant Prod. Prot Div Rep Agric Dept Crop Grassl Serv FAO. Available from: http://www.fao.org/ag/agp/agpc/doc/counprof/PDF%20files/South Africa_English.pdf Accessed: October 2017.
- Ollerenshaw CB, 1966. The approach to forecasting the incidence of fascioliasis over England and Wales 1958-1962. *Agric Meteorol* 3:35-53.
- Ollerenshaw CB, Rowlands WT, 1959. A method for forecasting the incidence of fascioliasis in Anglesey. *Vet Rec* 71:591-8.
- Olsen A, Frankena K, Bødker R, Toft N, Thamsborg SM, Enemark HL, Halasa T, 2015. Prevalence, risk factors and spatial analysis of liver fluke infections in Danish cattle herds. *Parasit Vectors* 15:160.
- Ord JK, Getis A, 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geogr Anal* 27:286-306.
- Parr SL, Gray JS, 2000. A strategic dosing scheme for the control of fasciolosis in cattle and sheep in Ireland. *Vet Parasitol* 88:187-97.
- Rinaldi L, Biggeri A, Musella V, De Waal T, Hertzberg H, Mavrot F, Torgerson PR, Selemetas N, Coll T, Bosco A, Cringoli G, Catelan D, 2015. Sheep and *Fasciola hepatica* in Europe: the GLOWORM experience. *Geospat Health* 9:309-17.
- Sanchez-Vazquez MJ, Lewis FI, 2013. Investigating the impact of fasciolosis on cattle carcass performance. *Vet Parasitol* 193:307-11.
- Schweizer G, Braun U, Deplazes P, Torgerson PR, 2005. Estimating the financial losses due to bovine fasciolosis in Switzerland. *Vet Rec* 157:188-93.
- Selemetas N, De Waal T, 2015. Detection of major climatic and environmental predictors of liver fluke exposure risk in Ireland using spatial cluster analysis. *Vet Parasitol* 209:242-53.
- Selemetas N, Ducheyne E, Phelan P, O'Kiely P, Hendrickx G, De Waal T, 2015a. Spatial analysis and risk mapping of *Fasciola hepatica* infection in dairy herds in Ireland. *Geospat Health* 9:281-91.
- Selemetas N, Phelan P, O'Kiely P, De Waal T, 2015b. Cluster analysis of fasciolosis in dairy cow herds in Munster province of Ireland and detection of major climatic and environmental predictors of the exposure risk. *Geospat Health* 9:271-9.
- Selemetas N, Phelan P, O'Kiely P, Waal TD, 2014. Weather and soil type affect incidence of fasciolosis in dairy cow herds. *Vet Rec* 175:371.
- Skuce PJ, Zadoks RN, 2013. Liver fluke. A growing threat to UK livestock production. *Cattle Pract* 21:138-49.
- Taylor MA, Coop RL, Wall RL, 2016. Parasites of cattle. In: Taylor MA, Coop RL, Wall RL, eds. *Veterinary Parasitology*. Hoboken, US: John Wiley & Sons Ltd; pp 352-435.
- Teagasc, 2016. Hill Sheep Conference 2016.
- Teagasc, 2015. Beef Production System Guidelines.
- Toet H, Piedrafita DM, Spithill TW, 2014. Liver fluke vaccines in ruminants: Strategies, progress and future opportunities. *Int J Parasitol* 44:915-27.
- van Dijk J, Sargison ND, Kenyon F, Skuce PJ, 2010. Climate change and infectious disease: helminthological challenges to farmed ruminants in temperate regions. *Animal* 4, 3779-2.
- Weier J, Herring D, 2000. Earth Observatory.