Modelling the distribution of bluetongue vectors

M. Baylis, L. O'Connell & B.V. Purse

Institute for Animal Health, Ash Road, Pirbright, Surrey GU24 0NF, United Kingdom

Summary

Recent epizootics of *Culicoides*-borne disease in the Mediterranean Basin have stimulated the development of climate-driven models for vectors. Predictor variables come from two main sources, weather data and satellites. Generally, models for *Culicoides imicola* combine temperature and moisture variables. The best weather models explain 75-85% of the variance in observed data for *C. imicola*, but satellite models perform better (85-95% of variance). Predictions of models for other regions appear mixed, with successes and failures. The failures indicate the need to:

- explore and incorporate other factors that may affect *Culicoides* populations, such as soil characteristics, host type and wind speed

- develop more complex models, recognising that different climate variables affect different stages of the life-cycle e.g. biological models. The very rapid spread in the distribution of *C. imicola* in recent years suggests that global warming may be a less important driver of change than other, currently unknown, factors.

Keywords

Recent epizootics of *Culicoides*-borne disease in the Mediterranean Basin, such as the 1987-1991 outbreak of African horse sickness (AHS) in Spain, Portugal and Morocco (2), and the ongoing 1998-2004 bluetongue (BT) epizootic (1), have stimulated the development of models of the spatial distribution of the *Culicoides* vectors, principally *C. imicola*. The models are climate-driven, as the life-history traits (e.g. survival and fecundity) of *Culicoides* are highly influenced by climatic factors, and this, in turn, affects their distribution and abundance, and the distribution and intensity of the diseases they transmit. The purposes of the modelling are as follows:

- to improve our understanding of the biotic and abiotic determinants of the distribution of the vectors

- to identify the limits to the distribution permitted by suitable climate, and thereby to define new areas at risk and areas that might remain disease-free

- to investigate how the distribution might change under scenarios of future global warming.

This review is limited to consideration of models of *C. imicola*.

Many weather variables influence *Culicoides* populations via their effects on survival and fecundity (11). High temperatures yield smaller, less fecund adults and both high and low temperatures lead to larval and adult mortality (17). The pupae of *C. imicola* may drown if breeding sites become too wet following rain. Equally, lack of rain and the concomitant low soil moisture may desiccate larvae. Low relative humidity at high temperature causes low saturation deficit, which will desiccate adults. There is evidence that high windspeeds affect local *Culicoides* populations, perhaps via dispersal (blowing midges away, both literally and metaphorically), starvation (adults unable to find feeding opportunities during consistently windy conditions) or desiccation (3, 4).

The development of climate-driven models for the distribution of *Culicoides* vectors requires uncovering the statistical relationship between climate variables and the presence/absence, or abundance, of the vectors. The vector picture to be modelled is blurred, however, by the understandable tendency to estimate

the vector population using light traps. The number of *Culicoides* spp. caught per night in a light trap is proportional to the size of the local population (which is what needs to be measured), multiplied by the activity rate and multiplied again by the efficiency of the trap. The latter two quantities are themselves affected by the weather (11, 12), and this confounds attempts to determine the effect of climate on the population. Activity rate is defined as the proportion of the population that is active in a given night and, for *Culicoides*, this rate is highly variable (2). Thus, *Culicoides* tend to be less active when nights are very hot or cold or when relative humidity is very low; they tend not to fly when it is raining; and they are not active when windspeeds exceed certain levels. Activity is also dependent upon light levels.

Trap efficiency is a measure of the ability of a trap to catch the vectors in its immediate area: if trap efficiency is 100%, the trap catches all vectors that approach it on a given night. For *Culicoides*, it is likely that trap efficiency is reduced by light sources other than that of the trap, such as moonlight, as this may distract the midges from approaching the trap. More significantly, the use of a suction device in light traps means that their efficiency reduces with an increase in windspeed. This effect has been known for other insects for many years (16) and the small size of *Culicoides* means that the effect is especially significant. Controlled experiments conducted at Pirbright (12), in which *Culicoides* were released into a room and caught in a trap exposed to different windspeeds, indicated a logarithmic relationship between windspeed and the reduction in trap efficiency (Fig. 1). A repercussion of this effect in the field is that trap catches outside a stable reduce at higher windspeeds, while those inside the stable do not (10).

How can the precise relationship be determined between *Culicoides* population size and climatic variables when, for example, temperature also reduces activity rates and windspeed reduces both activity and trap efficiency? In most experiments, the best estimate of a population variable is taken to be the average of a number of samples. However, the approach used here was borrowed from the field of remote sensing, called 'maximum compositing'. For satellite images of the earth, the detected level of radiance from the earth may be decreased but not increased (compared to the 'true' level of radiance) by cloud cover or other atmospheric effects. The best estimate of the true level is, therefore, the maximum level across several images recorded at different times. Similarly, *Culicoides* trap catches, as a proportion of the local population size, can only be reduced (and not increased) by the effects of weather on activity rate and trap efficiency (2). Thus, the best

estimate of population size will be the single greatest catch over a time period.

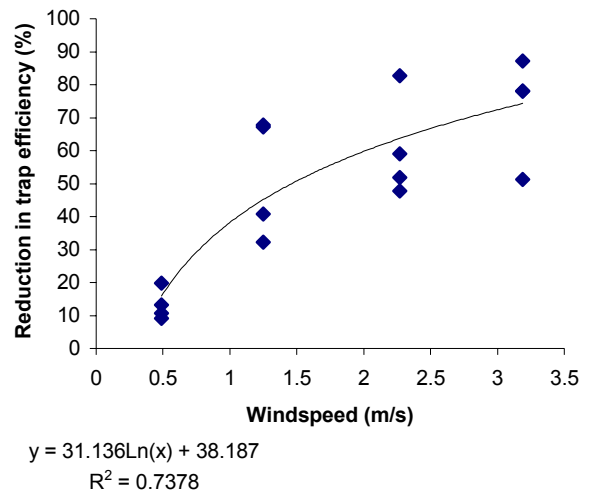$$y = 31.136 \text{Ln}(x) + 38.187$$
$$R^2 = 0.7378$$

Figure 1
The effect of windspeed on the ability of a light/suction trap to catch *Culicoides*
A known number of *Culicoides nubeculosus* were released into a chamber containing a trap, with fans positioned at different distances to create different windspeeds
Reprinted with permission from L. O'Connell (12)

In order to include a range of environmental conditions in the spatial model, population size must be estimated at as many sites as possible. However, maximum compositing of *Culicoides* catches requires repeated trapping at a given site over short time periods. Given that trapping effort is always limited, it is pertinent to ask how many trapping occasions are needed to provide an accurate estimate of population size and/or presence? To address this question, data from twenty-two sites in Morocco were analysed. These were sampled weekly between 1993 and 1995 as part of a vector surveillance campaign set up in response to the 1989-1991 epizootic of AHS. At sites of known *C. imicola* presence, the probability of a positive catch is about 0.22 in January, increasing to 0.77 in October. These monthly probabilities were used to estimate the number of zero catches that are required to give 95% confidence in the absence of *C. imicola* (Fig. 2a). In October in Morocco, two nights with zero catches are required; three nights in August, September and November; four nights in June and five or six nights in July and December; and seven nights in May. This analysis includes data from a small number of sites where only a handful of *C. imicola* were trapped over two years, and in areas that were free from AHS. If these sites are excluded, the probability of a positive catch (at the remaining sites, where disease risk is assumed) is 0.25 in January and 0.86 in October. The number of nights trapping for 95% confidence in absence (Fig. 2b) is two in September and October,

three in June, August and November, four in July and five in April, May and December. In a country such as Morocco, which has experienced significant outbreaks of both AHS and BT, trapping for only one night is not sufficient at any time of year to give confidence that *C. imicola*, or disease risk, is absent.

a) All known positive sites in Morocco
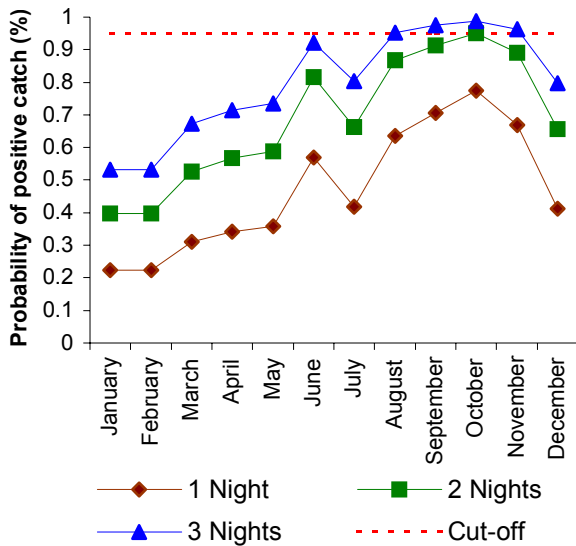
b) Excluding sites where only one or two *Culicoides imicola* were caught over two years, and which did not experience African horse sickness
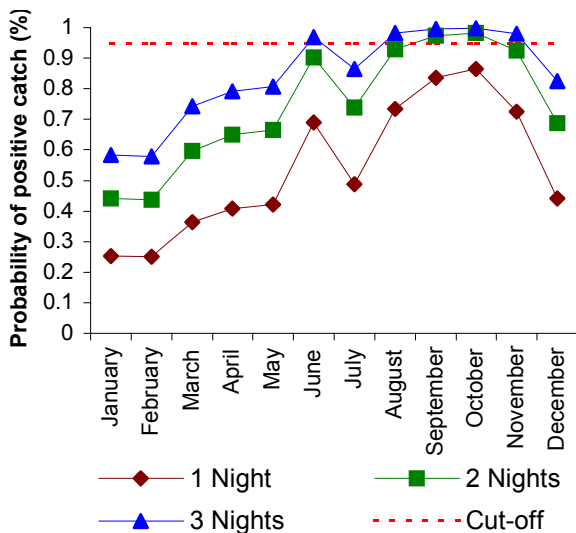

Figure 2
The probability of catching at least one *Culicoides imicola* at known positive sites in Morocco, when setting traps for one to three nights
The dashed line indicates the threshold above which there can be 95% confidence that *Culicoides imicola* is absent from a site, if none have been caught

Vector abundance, for the purposes of disease risk-mapping, must be measured at a spatial and temporal scale that is relevant to disease transmission. For *Culicoides*, population sizes change rapidly over time but generally show annual variation with one or two discrete peaks that for *C. imicola* tend to coincide with the seasonal peak of vector-borne disease in endemic areas. Maximum composited numbers over two-week intervals, averaged across two years, were used as measures of abundance for climate modelling in Morocco (3) but the significant trapping effort required precludes the use of this estimate of relative abundance across extensive sets of sites. Analysis of catch data from Morocco indicates that the maximum of catches over two nights during the late summer peak is significantly correlated with the fuller measure of abundance, and thus, this low effort trapping regime may permit abundance to be estimated at a large number of sites.

Climatic predictor variables used for modelling *Culicoides* presence/absence or abundance have been derived from two main sources, ground-collected weather data and satellite imagery. The former have ready biological significance, but are recorded synoptically at a relatively small number of weather stations that are often distant from trap sites, and between which it is necessary to interpolate, to obtain a continuous layer of climate information. Weather stations are expensive to buy and the data can be laborious to process. Finally, it may be difficult to obtain comparable weather data for other regions to which predictions might be extended. In contrast, satellite images are usually free and give global coverage at scales ranging from many kilometres to a few metres, thereby requiring no interpolation and facilitating extensive prediction. A further benefit is that suitably processed imagery is generally a better predictor of *Culicoides* than are ground-collected climate data. For example, windspeed and the annual minimum normalised difference vegetation index (NDVI, a measure of vegetation biomass obtained from earth-orbiting satellites) were the best predictors of the abundance of *C. imicola* in Morocco (3). In South Africa, a model using NDVI and land surface temperature (LST, a measure of ground temperature, again from satellites) accounted for 67% of variance in the abundance of *C. imicola*, compared to only 45% for a model using temperature and rainfall recorded by weather stations (5).

Nevertheless, models for *Culicoides* continue to be developed using weather data. Logistic regression was used to model the presence/absence of *C. imicola* in Iberia using historic (1930-1960) weather data (as equivalent data were available for all of Europe, thereby facilitating extrapolation) (18). The best model, which correctly predicted presence/absence at 83% of sites in Iberia, comprised the temperature

of the coldest month (the lowest mean daily minimum), the warmest month (the highest mean daily maximum) and the number of months with a mean temperature of ≥12.5°C. This model is unusual in lacking moisture variables, which other studies indicate to be very important for *C. imicola*. Not surprisingly, this temperature-driven model predicts *C. imicola* territory to occur in a broad band across southern Europe, with few predictions of absence at latitudes lower than Madrid. This model has recently been criticised (7). However, it should be noted that the predictions being criticised are substantially different from those presented by Wittmann *et al.* (18), as they were generated from a different source of weather data, without calibration.

Another logistic regression model, of the presence/absence of *C. imicola* at >500 sites in Italy (9) was based on 10 km × 10 km grid square data surfaces derived by interpolation between relatively few weather stations that were usually more than 40 km apart. At many sites, trapping was undertaken for a single night only, with the sampling regime biased towards areas in which BT virus (BTV) was present (and thus presumably suitable for vectors), and, indeed, this model performed less well than others (77.5% correct predictions). The best model comprised the annual mean daily minimum temperature, the annual mean daily minimum relative humidity and altitude. The model performed well for many parts of Italy (but not southern Sicily, where *C. imicola* is predicted to be prevalent), but was not externally validated by extrapolation to other countries.

Extrapolation to other countries was a primary objective of an 8-variable model of the abundance of *C. imicola* in Portugal, Spain and Morocco developed by discriminant analysis and 8 × 8 km fourier-processed satellite imagery (6). Three abundance ranges were considered, of which the lowest included 0 (i.e. absence). The best model correctly predicted the abundance range at 93.2% of sites and included, as the most important variables, proxies for both temperature and moisture as well as altitude. The high accuracy of predictions in Iberia/Morocco encouraged more extensive extrapolation (Fig. 3). Suitable conditions for *C. imicola* were predicted for eastern Spain, the Balearics, northern North Africa, Sardinia, Sicily, parts of Lazio and Puglia (Italy), eastern mainland Greece, the Peloponnese, Rhodes and Cyprus. All of these areas are now known to harbour *C. imicola*. There are false-positive predictions for southern Sicily and there are significant false-negative predictions. Most notably, it fails to predict the occurrence of *C. imicola* in Corsica and eastern Calabria (Italy).

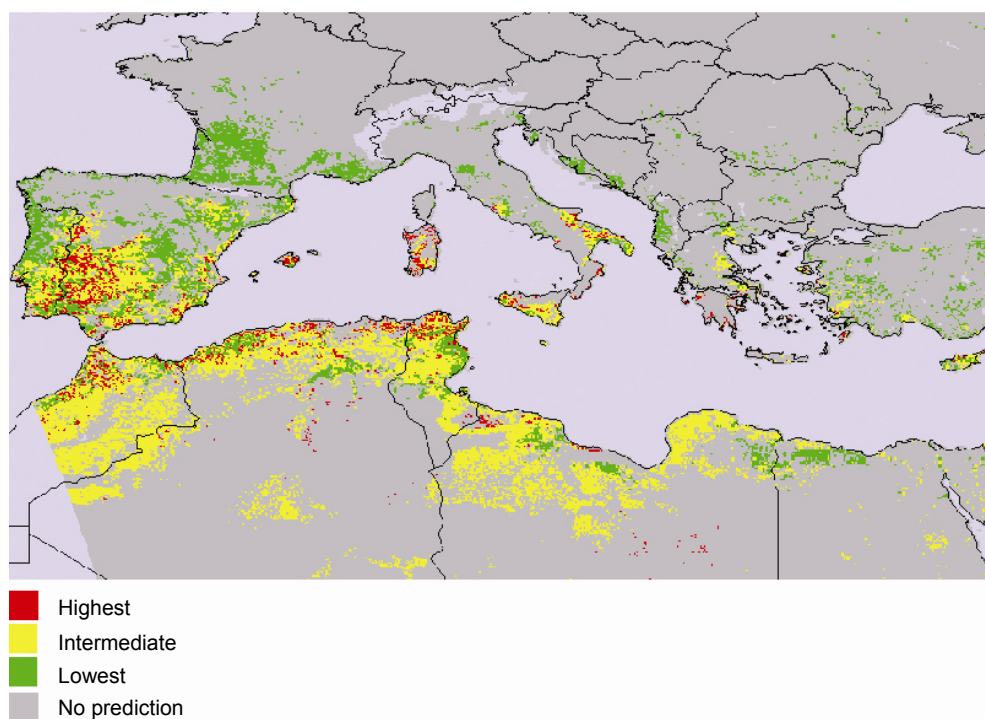Highest
Intermediate
Lowest
No prediction

Figure 3
Abundance of *Culicoides imicola* predicted by a model derived from the observed abundances at 44 sites in Iberia and Morocco
Reprinted with permission from Baylis *et al.* (5)

A similar model, based on recently collected vector data from Portugal (maxima of two summer catches) and fine resolution (1 × 1 km) satellite imagery (15), correctly predicted presence/absence at 95% of sites, and abundance range at 87% of sites. As before, the model comprised correlates of both temperature and moisture. Predictions across Europe are broadly similar to those of Baylis *et al.* (6).

A 10-variable model for Sicily correctly predicted the presence/absence of *C. imicola* at 87% of sites (13). The model correctly predicted the presence of *C. imicola* in parts of Iberia (including the Balearics), Sardinia and parts of Greece. However, the predicted distribution was generally much more restricted than the observed, with little prediction of presence in southern Spain, Corsica, mainland Italy and North Africa. This limitation probably results from the low number of presence sites for *C. imicola* in Sicily, such that only a restricted range of potential *C. imicola* habitats were included in the training set. This raises the question of why *C. imicola* is not more widespread in Sicily, despite most models predicting the island to be climatically suitable. It has been suggested that the porosity of the volcanic soils in Sicily may be unfavourable for breeding sites (7).

Finally, a 4-variable model for Corsica correctly predicted the *C. imicola* abundance range at 78% of sites (14). The predictions across Europe were less restrictive than in the study conducted in Sicily. The presence of *C. imicola* was correctly predicted in southern Portugal, south-western Spain, the east coast of Spain, the Balearic islands, the south coast of mainland France, the Corsican lowlands, most of Sardinia, northern and eastern Sicily, the west coast of mainland Italy, much of Calabria and Basilicata and parts of Greece, Turkey and Cyprus. There were some significant false-positive results, however: a low risk is associated with western Greece and Crete.

In summary, the best *C. imicola* models correctly predict 80%-95% of observations when built from satellite data and 75%-85% when based on weather data. In most cases, the models incorporate both temperature and moisture variables. Most models indicate that presence/abundance is most likely at moist, lower-altitude sites that are warm but not too hot. There is disparity between satellite-derived models in different regions as to whether mean, minimum, phase or amplitude NDVI variables are most important or whether minimum land surface temperature is an important predictor. While this may arise due to differences in the resolution of the imagery used or due to differences in the trapping regimes, it may also reflect biological differences in habitat requirements across a species range. This variation in importance of particular variables

between regions and models may be unimportant for the production of preliminary risk maps, but detailed maps will require investigation of the relationship between satellite-derived variables and biological processes that determine population performance. To date, no spatial models have attempted to incorporate the effects of windspeed as a possible variable that may control population size. Satellite-derived data layers for windspeed are becoming available, but their utility has not yet been confirmed. There is a further need to investigate and include certain soil characteristics that may mitigate against *C. imicola* breeding, irrespective of the suitability of the climate.

To date, only one model has been used to examine the possible effects of global warming on the future distribution of *C. imicola* (18). With a putative 2°C increase in mean temperature, the distribution of *C. imicola* is predicted to expand to the north, with significant probabilities of occurrence in southern France and northern Italy. This prediction is not surprising. As described earlier, this model is entirely temperature-driven and any effects of global warming on moisture (rainfall, humidity) or other weather variables cannot be taken into consideration. Indeed, it is the temperature-dependence of this model that almost certainly explains why it is the only model that has been used to investigate possible effects of climate change. Most studies have detected effects of both temperature and moisture on the occurrence of *C. imicola*. All climate change scenarios include effects on rainfall and soil moisture, but these are still difficult to predict with accuracy, and vary considerably geographically. Until the expected effects of climate change are known more clearly, it is too early to derive predictions on the future distribution of the moisture-sensitive *C. imicola*.

More importantly, the scale of change in recent years, with the supposed spread of *C. imicola* to the Balearics, Corsica, Sardinia, Sicily, mainland Italy, France and Greece and a significant northward spread, is very rapid compared to the observed rate of global warming and, to date, there has been no detectable spread in Portugal in the last decade (8). While climate change may be in-part driving the spread of *C. imicola* (1), it is important to consider that there may be other contributing factors, as yet unidentified, and also that the pattern may arise from increased surveillance effort such that there is an increased rate of discovery of *C. imicola* populations.

Finally, recent preliminary models have indicated that climatic determinants of distribution differ between *Culicoides* species, probably due to their differing life histories. Predictive risk maps for *Culicoides*-borne disease must be based on species-

specific spatial models for *C. imicola*, and novel vector species to avoid omitting extensive regions at risk of transmission by the latter.

Acknowledgements

References

1. Baylis M. (2002). – The re-emergence of bluetongue. *Vet. J.*, **164**, 5-6.

2. Baylis M., El Hasnaoui H., Bouayoune H., Touti J. & Mellor P.S. (1997). – The spatial and seasonal distribution of African horse sickness and its potential *Culicoides* vectors in Morocco. *Med. Vet. Entomol.*, **11**, 203-212.

3. Baylis M., Bouayoune H., Touti J. & El Hasnaoui H. (1998). – Use of climatic data and satellite imagery to model the abundance of *Culicoides imicola*, the vector of African horse sickness virus, in Morocco. *Med. Vet. Entomol.*, **12**, 255-266.

4. Baylis M. & Rawlings P. (1998). – Modelling the distribution and abundance of *Culicoides imicola* in Morocco and Iberia using climatic data and satellite imagery. *In* African horse sickness (P.S. Mellor, M. Baylis, C. Hamblin, C. Calisher & P.P.C. Mertens, eds). *Arch. Virol.* [Suppl.], **14**, 137-153.

5. Baylis M., Meiswinkel R. & Venter G.J. (1999). – A preliminary attempt to use climate data and satellite imagery to model the abundance and distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in southern Africa. *J. Sth Afr. Vet. Assoc.*, **70**, 80-89.

6. Baylis M., Mellor P.S., Wittmann E.J. & Rogers D.J. (2001). – Prediction of areas around the Mediterranean at risk of bluetongue by modelling the distribution of its vector using satellite imaging. *Vet. Rec.*, **149**, 639-643.

7. Calistri P., Goffredo M., Caporale V. & Meiswinkel R. (2003). – The distribution of *Culicoides imicola* in Italy: application and evaluation of current Mediterranean models based on climate. *J. Vet. Med. B*, **50**, 132-138.

8. Capela R., Purse B.V., Peña I., Wittmann E.J., Margarita Y., Capela M., Romao L., Mellor P.S. & Baylis M. (2003). – Spatial distribution of *Culicoides* species in Portugal in relation to the transmission of African horse sickness and bluetongue viruses. *Med. Vet. Entomol.*, **17**, 165-177.

9. Conte A., Giovannini A., Savini L., Goffredo M., Calistri P. & Meiswinkel R. (2003). – The effect of climate on the presence of *Culicoides imicola* in Italy. *J. Vet. Med. B*, **50**, 139-147.

10. Meiswinkel R., Baylis M. & Labuschagne K. (2000). – Stabling and the protection of horses from *Culicoides bolitinos* (Diptera: Ceratopogonidae), a recently identified vector of African horse sickness. *Bull. Entomol. Res.*, **90**, 509-515.

11. Mellor P.S., Boorman J. & Baylis M. (2000). – *Culicoides* biting midges: their role as arbovirus vectors. *Ann. Rev. Entomol.*, **45**, 307-340.

12. O'Connell L. (2002). – Entomological aspects of the transmission of arboviral diseases by *Culicoides* biting midges. PhD, University of Bristol.

13. Purse B.V., Tatem A.J., Caracappa S., Rogers D.J., Mellor P S., Baylis M. & Torina A. (2004). – Modelling the distributions of *Culicoides* bluetongue virus vectors in Sicily in relation to satellite-derived climate variables. *Med. Vet. Entomol.*, **18** (2), 90-101.

14. Roger F.L. (2002). – Emergence of bluetongue disease in the Mediterranean basin: modelling locations at risk for potential vectors (*Culicoides* spp.) using satellite imagery. MSc, University of London.

15. Tatem A.J., Baylis M., Mellor P.S., Purse B.V., Capela R., Pena I. & Rogers D.J. (2003). – Prediction of bluetongue vector distribution in Europe and North Africa using satellite imagery. *Vet. Microbiol.,* **97** (1-2), 13-29.

16. Taylor L.R. (1962). – The absolute efficiency of insect suction traps. *Ann. Appl. Biol.*, **50**, 405-421.

17. Wittmann E. (2000). – Temperature and the transmission of arboviruses by *Culicoides* biting midges. PhD, University of Bristol.

18. Wittmann E.J., Mellor P.S. & Baylis M. (2001). – Using climate data to map the potential distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in Europe. *Rev. Sci. Tech. Off. Int. Épiz.*, **20** (3), 731-740.