ECDC technical report

The spatial distribution of Crimean-Congo haemorrhagic fever in Europe and neighbouring areas

March 2023

This report was commissioned by the European Centre for Disease Prevention and Control (ECDC) to assess the spatial distributions of Crimean-Congo haemorrhagic fever (CCHF) in Europe and neighbouring areas under specific contract No 4 ECD.13254 ID.13285 implementing inter-agency framework contract for services No ECDC/2019/020. The European Union Horizon 2020 MOOD Project N° 874850 (https:// https://mood-h2020.eu/) provided driver covariates and vector distributions and has catalogued this document as MOOD078. The report was managed by Olivier Briet (ECDC)

Authors

Jane Messina (School of Geography and the Environment) and William Wint, Environmental Research Group Oxford).

Acknowledgements

Thanks are due to Marieta Braks and Hein Sprong who managed the contract with ECDC, and Wim Van Bortel as VectorNet Scientific Support Work Package leader. These colleagues, together with Jolyon Medlock, Natalia Fernandez-Ruiz, Augustin Estrada-Pena, Kayleigh Hansford, Frank Sandmann and Olivier Briet also provided invaluable guidance and comments on the technical report.

Contents

Abbreviations

AIC	Akaike information criterion
BRT	Boosted Regression Trees
DEM	Digital Elevation Model
CCHF	Crimean-Congo haemorrhagic fever
CCHFV	Crimean-Congo haemorrhagic fever Virus
ECDC	European Centre for Disease Prevention and Control
ECMWF	European Centre for Medium Term Weather Forecasting
EFSA	European Food Safety Authority
EU	European Union
EVI	Enhanced Vegetation Index
GBIF	Global Biodiversity Information Facility
GIS	Geographic Information System
LST	Land Surface Temperature
MODIS	Moderate Resolution Imaging Spectrometer
MOOD	MOnitoring Outbreaks for Disease surveillance in a data science context
NASA	National Aeronautics and Space Administration
NDVI	Normalised Difference Vegetation Index
NUTS3	Nomenclature of Units for Territorial Statistics, level 3
RF	Random Forest

Executive Summary

Crimean-Congo haemorrhagic fever (CCHF) is a tick-borne viral (Nairovirus, family Bunyaviridae) infection first identified in the Crimean region in 1944. It is one of the most widely distributed arboviral diseases in the world, ranging from southern Russia and the Black Sea region to the southern tip of Africa. The disease is considered as 'emerging' across the globe, with many countries reporting new CCHF cases in humans in recent decades including Georgia, Türkiye, Albania, and most recently, Spain.

The distribution of CCHF was modelled globally in 2015 by Messina *et al.* (1) and in 2020 by Okely *et al.* (2). Messina *et al.* predicted high probability areas for the disease to be in parts of eastern and southern Europe. It also predicted high ecological suitability for CCHF in Spain, which was masked out of the published distribution as the disease had then yet to be found in Iberia. The 2015 model also did not take into account the likely distributions of its main tick vectors in Europe, namely *Hyalomma marginatum* and *H. lusitanicum.* The Okely *et al.* study showed updated risk areas which included Spain, and also statistically compared the distribution of environmental suitability for CCHF human disease to that of the primary tick vectors. Still, these vector distributions were not used to modify or mask the maps of human CCHFV suitability.

ECDC therefore asked that the 2015 CCHF distributions for Europe and its neighbouring areas be re-estimated and updated taking into account both the recent disease occurrence data, and the distribution of its vectors. Two sets of spatial modelling were performed, both using a common covariate predictor data suite, and long-established spatial modelling techniques. These models were a) CCHF itself using Boosted Regression trees (BRT), as used in 2015, and trained on presence data extracted from the published literature and statistically assigned pseudo-absence data b) the two vector species using an ensemble of Random Forest and Boosted Regression Tree models, trained on presence data provided by the ECDC VectorNet project, other online databases and absence data based on calculated habitat suitability.

As the CCHF occurrence in Europe is too limited for its ecological suitability to be reliably spatially modelled based on this data alone, the global distribution of the disease —including Europe, Africa as Asia— was modelled, based on the assumption that a global model better informs a predicted distribution for Europe and its neighbouring areas. The resulting disease prediction suggests substantially more ecological suitability for CCHF in Europe and its neighbouring areas than did the 2015 model, largely because more disease records were found in Spain, Albania, Greece and Western Asia. The environmental suitability for the disease extends into much of northern Europe and the northern Caucasus —significantly beyond the vector distributions, which at worst restrict its northern range to 47°N.

The vector distribution models allow for refined mapping of CCHF ecological suitability compared to 2015 (where entire countries were masked out) by using them to mask out areas in the basic CCHF model where no vectors are present. Using the minimum predicted vector distribution as mask, much of the Mediterranean sea-bord along with Türkiye and the Caucasus are predicted to be widely suitable for CCHF, with a more patchy suitability precited for inland Greece, and the southern Balkans. However, using the maximum predicted vector distribution as mask, the resulting CCHF ecological suitability map matches the observed disease occurrences most closely, and on that map, the predicted suitability for CCHF extends significantly further north into parts of central France and eastward.

This suggests that a number of countries that have yet to record CCHF should be on their guard —most especially those with Mediterranean coastlines (France, Italy, the southern Balkans, countries in western Asia as well as north Africa). In comparison to the (masked) 2015 model, whilst much of Europe's ecological suitability for CCHF remains similar in this updated model, several regions show increased ecological suitability for CCHF, particularly pockets throughout Spain, France, Italy, and the Balkans and the Caucasus.

Background

Crimean-Congo haemorrhagic fever (CCHF) is a tick-borne viral (Nairovirus, family Bunyaviridae) disease first identified in the Crimean region in 1944 (3,4). It was subsequently shown to be the same virus as that causing similar haemorrhagic disease outbreaks in the Congo basin, giving the virus its current name (5,6). CCHF is a widely distributed disease, ranging from Eastern China to southern Russia and the Black Sea region, and to the southern tip of Africa (6).

The disease is considered as 'emerging' across the globe, with many countries reporting infections in humans in recent decades, including Albania (2001) (7), Türkiye (2002) (8), Georgia (2009) (9), and Spain (2013) (10). In some regions, human CCHF infection has also been recently reported after long periods of absence, for example in south-western Russia (11) and Central Africa (12).

The disease is transmitted to humans by ticks, while no apparent disease manifestation occurs in animals (13), both wild and domesticated animals represent an important link in the disease transmission cycle, acting as reservoirs for continued tick re-infection. Many tick species have been associated with CCHF virus (CCHFV), but members of the genus *Hyalomma* are considered the primary vectors and are the most common ticks known to transmit the virus to humans. In Europe the relevant species are *Hyalomma marginatum* and *Hyalomma lusitanicum* (14). These ticks are adapted to hot and dry or semiarid environments, and are found in many parts of Africa, Asia, and Europe (3,15–18) (see Methods section below). Infection of humans is uncommon, although those living or working in close proximity to livestock are at greater at risk and fatality is as high as 40%. No CCHF-specific antiviral drug or vaccine currently exists for animals or humans.

The distribution of CCHF was modelled globally in 2015 by Messina *et al.* (1), a study which predicted the disease to be possible in parts of eastern and southern Europe. Whilst the 2015 model successfully predicted that Spain was suitable for CCHF, the published outputs were adjusted so that ecological suitability for CCHF was only shown in countries that had reported cases. These did not include Spain, which was not therefore shown as a potential area for CCHF distribution. In addition, CCHF predictions from other authors have become available (2). As a result, ECDC identified the need to update the earlier modelling exercise using up-to-date known distribution data to train the model. The 2015 model also did not incorporate vector distributions in the disease predictions, and ECDC therefore asked that the methodology be modified to produce distributions for the two main European vectors (*H. marginatum* and *H. lusitanicum*) which could be used as additional disease model masks.

Methods

Overview

The basic modelling process involves establishing a statistical relationship between known presence (or absence) and the values of a series of selected predictor covariates. These relationships are calculated for a set of sample locations, and the estimated equations then applied to maps of the covariates which provide values at a pixel resolution for the entire area of interest. This results in a modelled spatial distribution showing the probability of presence at the resolution of the covariate maps —which were standardised in this project at 1 kilometre.

Two sets of spatial models were implemented for this study —the first for the vectors and the second for the disease, with the specific objective of providing disease suitability assessments for Europe and its neighbouring areas as far east as the Caspian Sea and to include north Africa.

The distribution of the disease within this limited area is somewhat restricted and provides insufficient training data to run an effective model for Europe. It was therefore necessary to run the disease model covering the full extent of the disease distributions so that the global distributions could be used to inform the potential suitability in Europe.

The situation for vectors is different: whilst there are a number of different CCHF vector species within the global range of the disease, only two are widespread in Europe, namely *H. marginatum* which is found throughout southern Europe, the Middle East and North Africa as far east as India; and *H. lusitanicum* which is limited to

south western Europe and neighbouring parts of North Africa (*e.g.* Kolonin (2009) (19), Estrada-Pena *et al.* (2012),(20) VectorNet project (21)). The vectors were each modelled for their entire ranges plus a buffer zoneof at least 200 km.

The methods used to generate vector and disease models were similar (see above) and the driver covariates offered to the models were drawn from the same standard covariate suite that has successfully been used in modelling a range of vectors for ECDC and the European Food Safety Authority (EFSA) (22). These covariates include indicators of level and seasonality derived from MODIS satellite imagery time series using Temporal Fourier Analysis (23) for day and night land surface temperature (LST), vegetation indices, middle infrared and relative humidity; and for rainfall derived from the European Centre for medium Term Weather Forecasting (ECMWF) datasets. Consensus land use, as well as topographic and demographic variables were also part of the covariate suite. These are set out in detail in

Supplementary Table 2 in the Appendix. All datasets were standardised to ensure consistent spatial resolution and extent.

Vectors

The vector spatial distribution modelling was performed using both Random Forest (RF) and Boosted Regression Trees (BRT) implemented through the VECMAP® Software Suite (AVIA-GIS, Belgium), to model presence and absence producing estimates of the probability of presence. Five replicates of each method, with a 25% holdback, were run, and the results combined to produce ensemble mean, median, minimum and maximum predictions of probability of presence. The combination of methods tends to reduce a tendency for BRT to overfit, especially for training data covering relatively restricted areas such as that available for *H. lusitanicum*.

These methods require approximately equal numbers of presence and absence points to be offered to each modelling run. The occurrence data was obtained primarily from VectorNet (21) shown in Supplementary Figure 10, supplemented by very limited data from the Global Biodiversity Information Facility (GBIF, https://www.gbif.org). The VectorNet data consisted of both point and polygon data, differentiated into present, absent and introduced categories. This last category represents records of temporary presence —often from migratory birds and are not indicative of established populations. They were therefore discarded. Five points were defined for each polygon and assigned as present or absent according to the polygon status, and to these were added any point data from either VectorNet or GBIF.

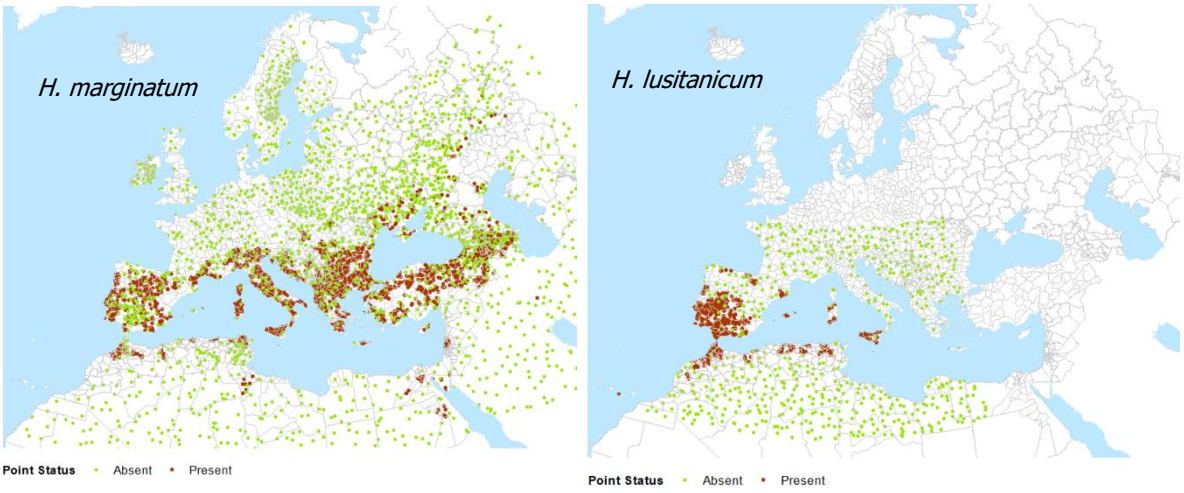
The presence-absence modelling requires the input data to be balanced into approximately equal numbers of presence and absence points. The available vector data did not include the requisite number of absence points, so these needed to be generated. There are a number of geostatistical ways absences can be generated (as used for the disease modelling below), but for these vectors, it has been possible to produce habitat suitability layers based on environmental thresholds. This means that, for vectors, it is possible to assign absences to areas defined as unsuitable.

A location was defined as suitable for each vector as indicated in Supplementary Table 1 (Appendix), and mapped in Supplementary Figure 11, Appendix. These were taken from a number of published sources (20,24–31) and from personal communications with expert colleagues. Whilst both species do well in most woodland, grassland, shrubland and cropland environments, only *H. lusitanicum* occupies dense woodland and only *H. marginatum* is associated with sparse vegetation. Though *H. marginatum* does require minimum temperatures though the summer and needs relatively moist conditions, *H. lusitanicum* is able to occupy areas with hotter and drier summer conditions and requires a relatively warm autumn.

Once the suitable and unsuitable areas had been defined, and following the methodology used in Wint *et al.* (2020)(22), for *H. lusitanicum*, absence points were randomly assigned to unsuitable areas within approximately 300km of known presences, and to all areas, irrespective of suitability, further away. For *H. marginatum*, absence points were randomly assigned to unsuitable areas inside the range depicted by Kolonin (2009) (19) and within approximately 300km of known presences, and to all areas, irrespective of suitability, further away.

Figure 1: Points used for modelling of tick vectors

Left *H. marginatum*, Right *H. lusitanicum.* Brown is presence, green is absence



All points were then were aggregated to a 10km grid, to combine any multiple overlapping records. The number of presence and absence points was then adjusted by reducing the number of the larger class (in this case absences) to that of the less frequent class to produce a final balanced output dataset for spatial modelling. These are shown in Figure 1.

CCHF

An occurrence database comprising point (*e.g.*, town or city) or polygon (*e.g.*, county or province) locations of confirmed CCHF infection presence was compiled from peer-reviewed literature, Genbank records, and ProMed Mail reports (https://promedmail.org/). A full list of peer-reviewed citation references for all sources is provided alongside this report. A literature search was conducted on PubMed using the terms 'CCHF' or 'Crimean Congo Hemorrhagic Fever' or 'Crimean Hemorrhagic Fever' or 'Congo Hemorrhagic Fever', as well as their British English spelling equivalents (spelled "haemorrhagic"). The same terms were used to search Genbank. An occurrence was defined as one or more laboratory confirmed human infection(s) with CCHFV occurring at a unique location (the same administrative area or 1 km×1 km pixel for points) within one calendar year.

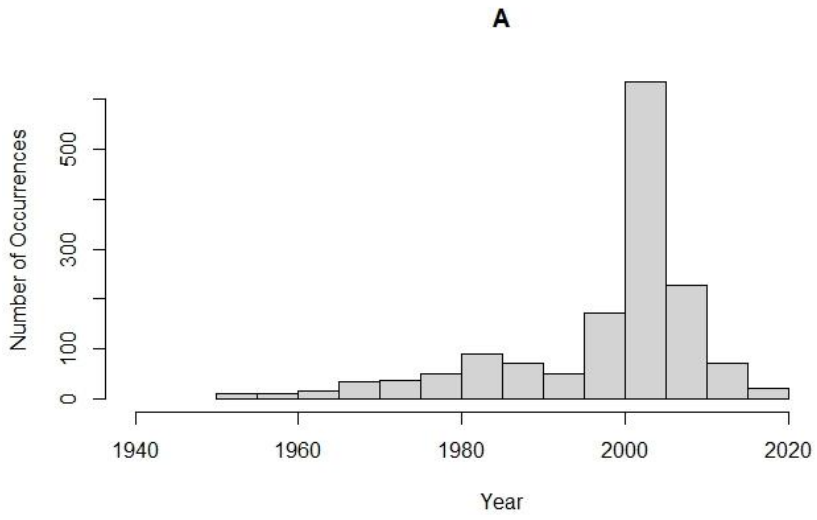
All occurrence data underwent manual review and quality control to ensure information fidelity and precise geo-positioning. Reports of autochthonous (locally transmitted) cases or outbreaks were entered as an occurrence within the country in which transmission occurred. If imported cases were reported with information about the site of infection, they were geo-positioned to the country where transmission occurred. If imported cases were reported with no information about the site of infection, they were not entered into the database. In addition, polygons greater than one square degree in area at the equator (approximately 12,300 km^2 at the equator) were removed from the database, as their inclusion in niche modelling would introduce a large amount of spatial imprecision.

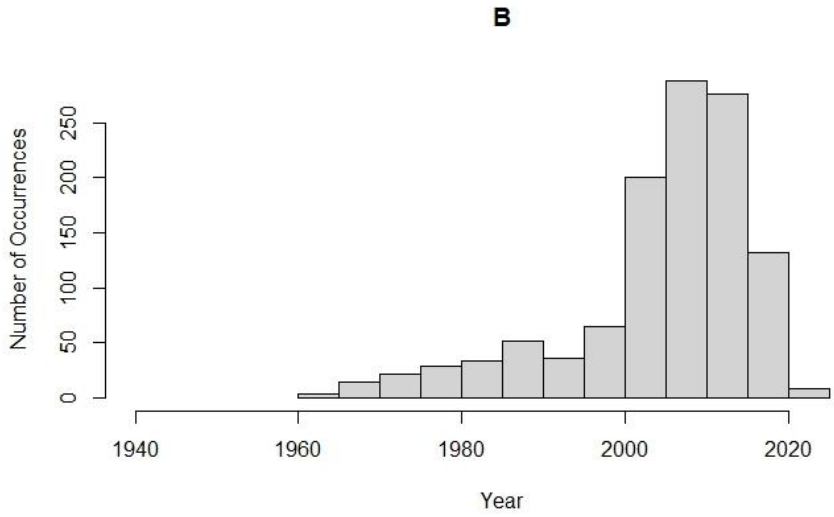
In total, 1,437 point occurrence records and 854 polygon occurrence records were included in our final dataset after performing all quality control procedures (Supplementary Figure 12 for global map). These publications spanned the years 1953 through to July 2022. We assumed that any recorded location of CCHF occurrence, regardless of the date of the record, represented an environment permissible for the occurrence of disease cases.

Figure 2 shows histograms for the occurrence locations by year of publication. Occurrence and generated pseudo-absence locations (described in greater detail below) for Europe are shown in

Figure 3 **Error! Reference source not found.**.

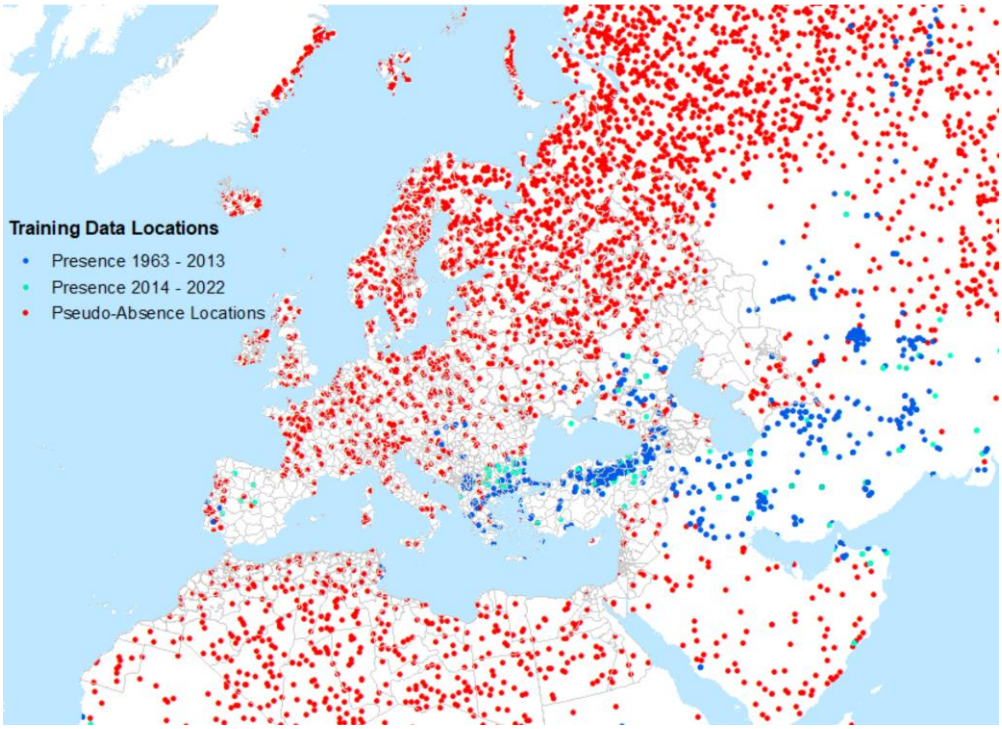
Figure 2: Frequency histograms of CCHFV human occurrence locations by year of publication





Histograms by year of (A) point occurrence locations; (B) polygon occurrence locations

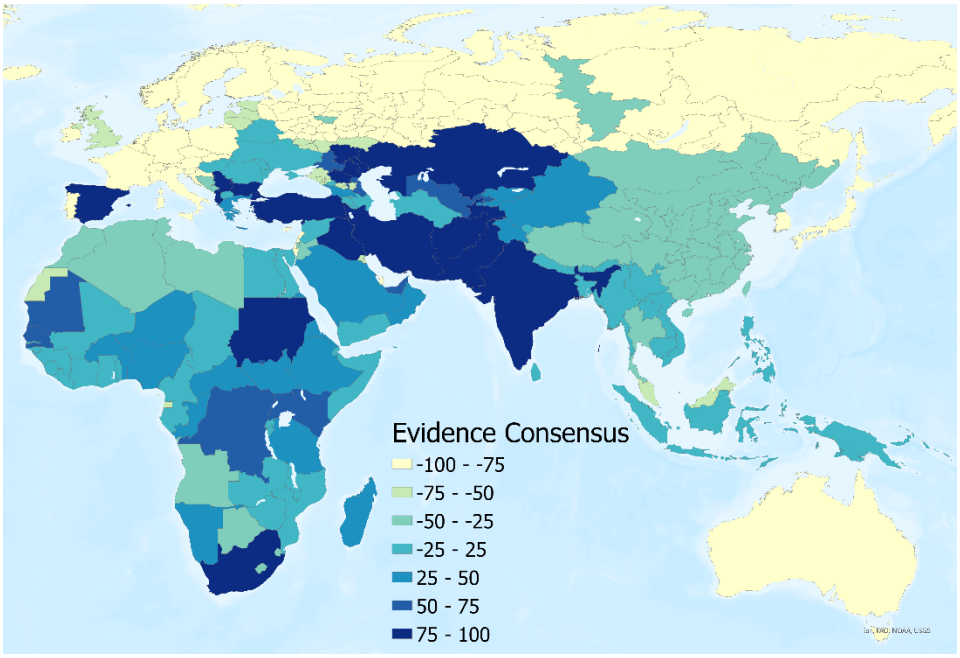
Figure 3: Occurrence and pseudo-absence locations of CCHFV human cases (Europe and neighbouring areas)



Global gridded data (1 km×1 km) were taken from the covariate data suite (Supplementary Table 2) for a set of four explanatory covariates. The covariates were chosen based on factors known or hypothesised to contribute to suitability for CCHF transmission based upon the national-level studies described in the introduction. These included: i) annual mean precipitation interpolated from global meteorological stations, and ii) mean land surface temperature derived from NASA's moderate resolution imaging spectrometer (MODIS) imagery, intended to capture the generally warm and arid climate zones where CCHFV is transmitted; iii) 1 km resolution measure of the mean annual Enhanced Vegetation Index (EVI, also from MODIS); and iv) the proportion of each 1 km×1 km grid cell covered by shrub or grass land cover types derived from the Earthenv consensus land cover datasets. No covariate grids were shown to be adversely affected by multicollinearity.

Having assembled occurrence and covariate datasets, a BRT approach was used to establish a multivariate empirical relationship between the probability of CCHF presence and the environmental conditions (as determined by the set of covariates described above) sampled at each occurrence location for points, or averaged within each polygon-level occurrence, as mentioned above for the vector modelling, BRT models require not only presence data but also pseudo-absence data defining areas of potentially unsuitable environmental conditions at unsampled locations, since data on absence of disease cases are rarely reported.

Figure 4: Evidence consensus map for CCHF presence or absence by country



An evidence-based probabilistic framework for generating pseudo-data was used see (Figure 4, methodology as per Messina *et. al.* (1)). To represent the environmental conditions in locations where the disease has not been reported, 5,000 background points were randomly generated and weighted based on a continuous raster surface derived from the national (and sometimes sub-national) CCHF evidence consensus scores (ranging -100 (consensus on absence) to 100 (consensus on presence)). As such, more background points were located in areas with high consensus on absence.

To increase the robustness of model predictions and quantify model uncertainty, BRT models were fitted to fifty separate bootstraps of the data. We then evaluated the central tendency as the mean across all 50 BRT models. Each of the 50 individual models was fitted using the ***gbm.step*** subroutine in the ***dismo*** package in the R statistical programming environment. All other tuning parameters of the algorithm were retained as in *Messina et al.* (1) .

Results

Vector Distributions

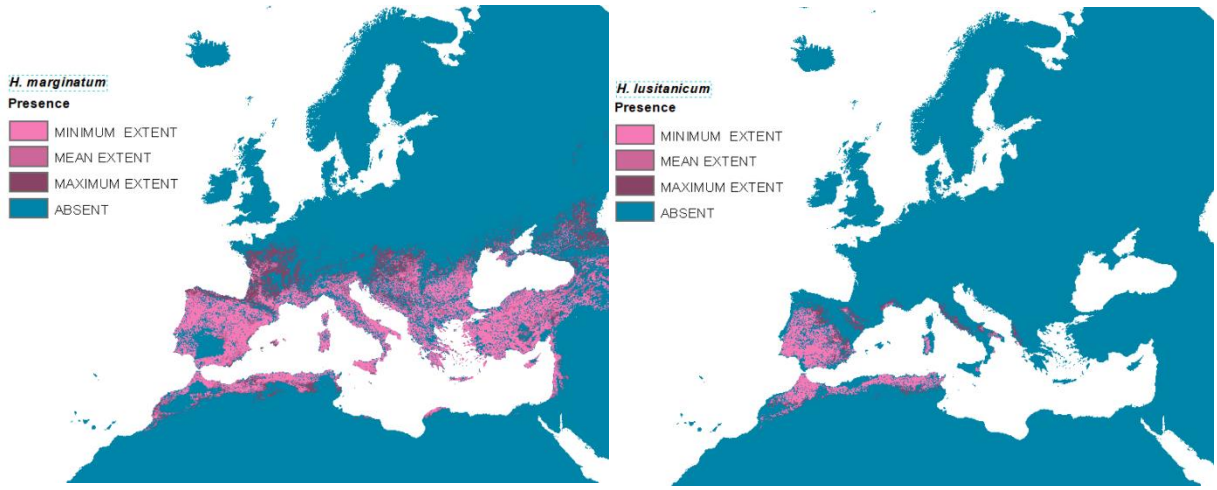
Predicted presence (and absence) for each vector species are shown in Figure 5. Three estimates of presence were derived from the model replicates —the mean, minimum and maximum values, each coded as present if the calculated probability of presence value exceeds 0.5. This resulted in four distribution categories in each map: minimum, mean and maximum predicted vector extent, and absence. All are masked by the suitability limits described in the methods above.

The vector models generally reflect the training data well, especially for *H. lusitanicum* where both sensitivity and specificity or the RF models both average well above 0.9 and the receiver operating characteristic (ROC) values of the BRT models exceed 0.95. This is perhaps to be expected as the species range is relatively restricted and less environmentally heterogeneous than that of *H. marginatum*. The *H. marginatum* models also perform well with associated Kappa values of >0.84 for the RF models and ROC values > 0.89 for the BRT models. Whilst these models have a high sensitivity (> 0.83), their specificity is a little lower at 0.69–0.72, meaning that the models predicted a significant number of false negatives. These are largely to the East of the modelled area and outside the project area of interest —and are further discussed below.

The range of predicted vector distributions is shown by the minimum and maximum predicted extent (Figure 5): the maximum predicted extent for *H. marginatum* extends substantially further north into parts France, Central Europe and the Caucasus than do the mean predictions. It is tempting to suggest that this reflects an increased likelihood of northward shifts in range driven by climate change. It is also worth noting that these maximum

predicted extents overlap some of the training distributions assigned as 'introduced' (Supplementary Figure 10) because they were thought to be too far beyond the current northern distribution limits to represent established populations. It could be that these presence points are in fact the first harbingers of future spread.

Figure 5: Predicted vector distribution maps



The most frequently used predictors for *H. marginatum* are a) the minimum values of the Normalised Difference Vegetation Index (NDVI) b) aspects of the day and night-time temperature, c) timing and levels of rainfall and d) aspects of minimum relative humidity. For *H. lusitanicum*, the best predictors relate to a) monthly rainfall seasonality and the degree to which it varies over the year; and b) the levels and timing of day and night-time temperature. The importance of the top ten predictors for each vector and modelling method is given in Supplementary Table 3.

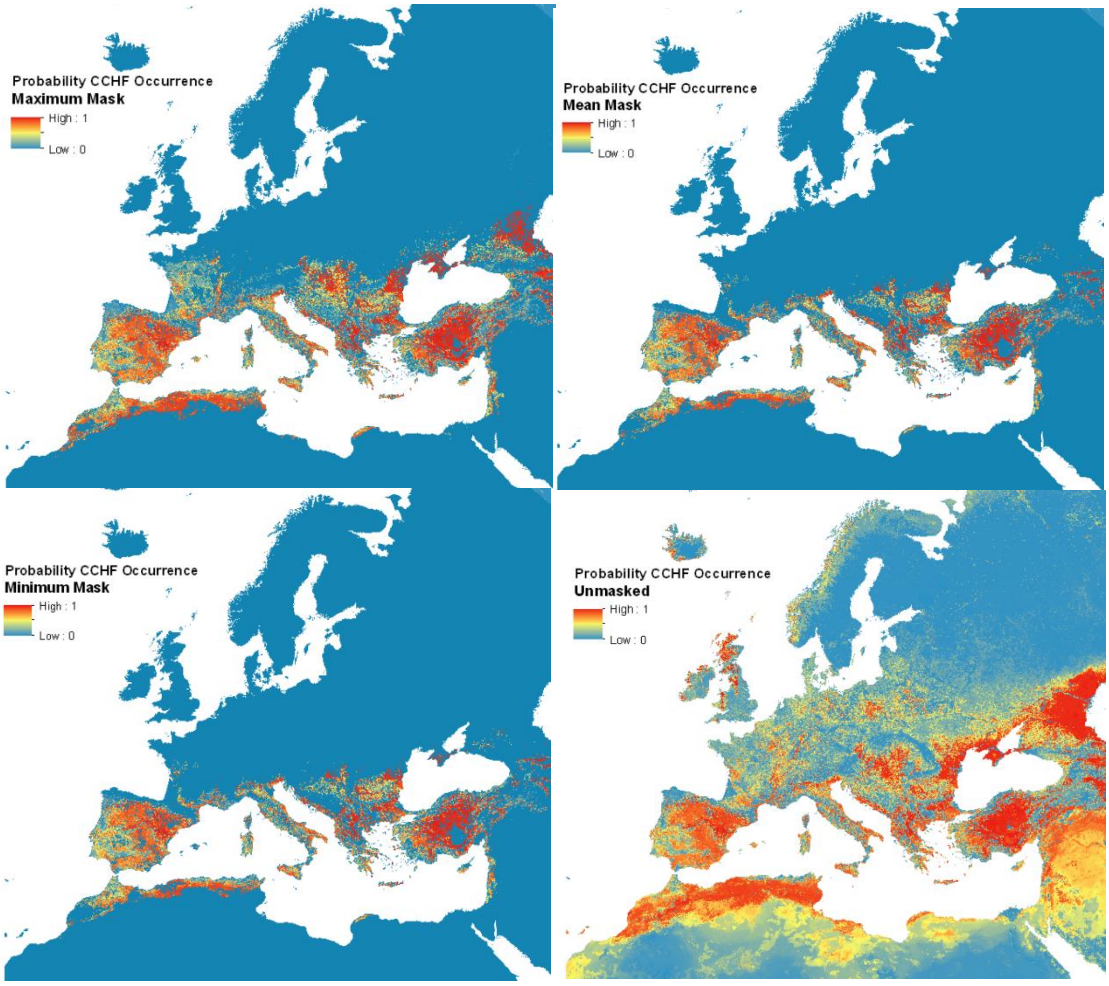
CCHF Distribution in humans

Figure 6 shows the resulting mean predictions for CCHFV probability of occurrence (ranging 0–1), displayed with masks of minimum, maximum, and mean predicted vector presence (where at least one of the probability values for either vector's presence was greater than or equal to 0.5), as well as the unmasked version. Supplementary Figure 13 shows the full global modelled distribution (beyond Europe; unmasked). Uncertainty estimates for each pixel were possible from the ensemble approach (by using the interquartile range of all 50 predictions), and are shown in the Appendix (Supplementary Figure 14).

The unmasked disease predictions suggest that the environmental suitability for the disease extends well into northern Europe and the northern Caucasus —far beyond the distributions of the two vector species considered. Limiting the predictions to the area covered by the vector distributions implies at best (mask with minimum vector distribution area) that Mediterranean France, Spain, Türkiye, and much of the Mediterranean seaboard are all predicted to be widely suitable for the disease, with a more patchy suitability precited for inland Greece, and the southern Balkans.

Looking at the recorded disease locations (Supplementary Figure 12) which extend well into the Caucasus and central Europe, it seems that the vector masking using the maximum predicted extents provides the closest match to the known data. This level of masking suggests the predicted disease range extends significantly further north to 47°N (north central France and eastward). It is also evident that the vector masking "removes" predicted disease occurrence from Israel, Palestine, Lebanon, Syria, and Iraq. These are areas for which there are no presence records of either vector, and are also not predicted to support either vector.

Figure 6: Predicted CCHFV suitability maps



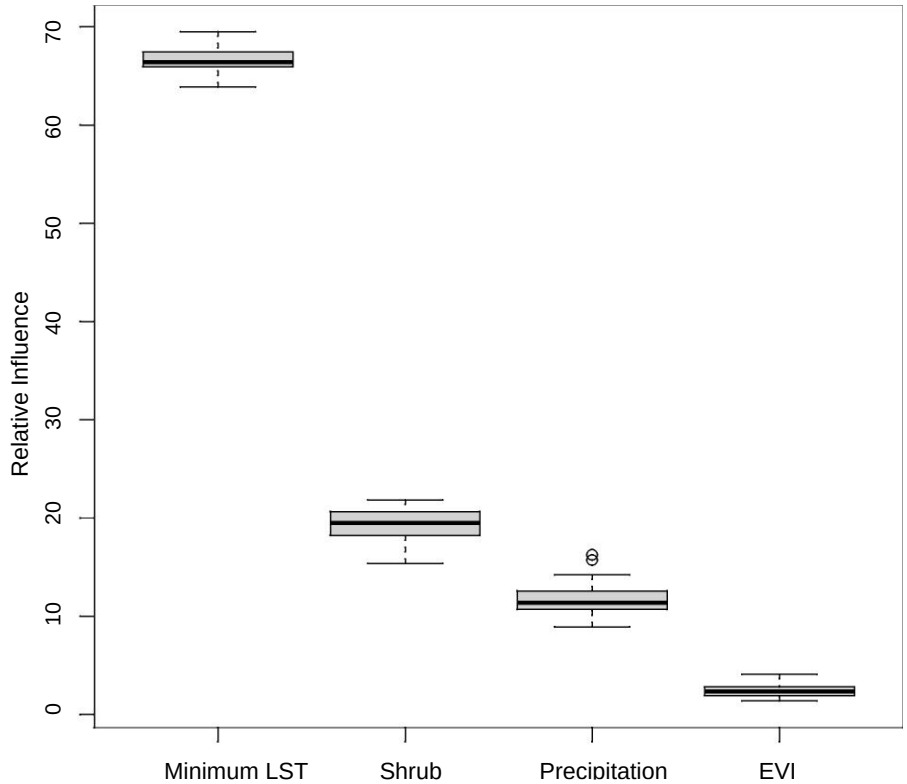
Top left: masked limiting CCHFV occurrence predictions to the maximum vector extent layer; Top right: masked limiting CCHFV occurrence predictions to the mean vector extent layer; Bottom left: masked limiting CCHFV occurrence predictions to the maximum vector extent layer; Bottom right: unmasked CCHFV occurrence predictions

Because these maps are not masked by areas where the consensus layer suggests (high) consensus, they show France and Italy as having a high probability of CCHF, even though there have not been any reports of cases thus far. Still, Figure 6 shows several patches of environmental suitability within the predicted vector ranges in these countries. Currently, Northern Europe is not predicted to be ecological suitability for CCHFV due to a lack of vector presence.

Supplementary Figure 13 in the Appendix shows the unmasked probability of CCHFV occurrence for the full modelled extent. Uncertainty estimates for the updated CCHFV probability of occurrence maps for each pixel were possible from the ensemble approach (by using the interquartile range of all 50 predictions) and are shown in the Appendix (Supplementary Figure 14).

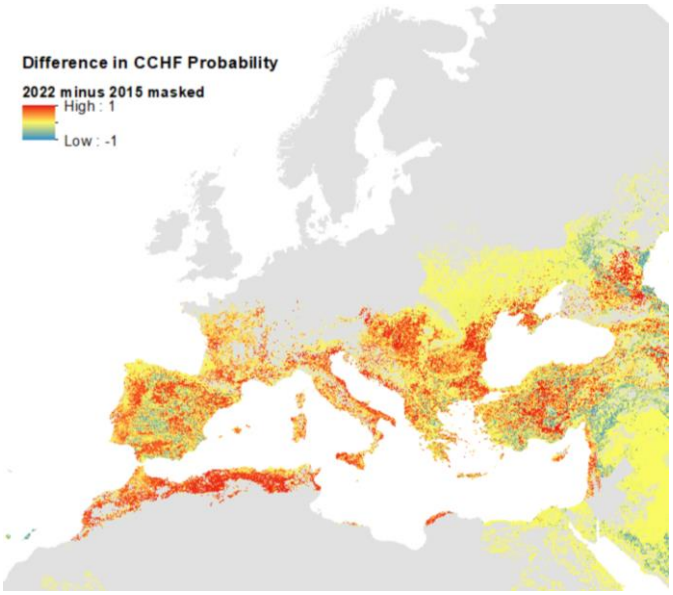
The proportional influence of covariates towards the probability predictions in the ensemble of models is shown in Figure 7, with the total proportion equalling 100 percent. The average Area Under the Curve (AUC) across all models was 0.74 (range 0.67–0.79). AUC ranges from 0 to 1 (with an AUC of 0.5 being as good as a random prediction), and provides an aggregate measure of performance across all possible classification thresholds, such that as the probability that the model ranks a random positive example more highly than a random negative example. Values closer to 1 indicate better model fit. Land surface temperature was by far the greatest contributor to the models, followed by precipitation, shrub land cover, and then EVI.

Figure 7: Proportional influence of covariates on CCHFV suitability predictions



The difference between the unmasked predictions for 2022 and 2015 (Supplementary Figure 13 and Supplementary Figure 15) is presented in Figure 8 and compares the current predictions masked with the maximum predicted vector extent with those of 2015 —the red areas indicating where the current model predicts substantially higher ecological suitability for CCHF. This reflects the additional CCHFV presence records recorded for the current models in the eastern part of the region (Supplementary Figure 12). Whilst much of Europe's predicted ecological suitability remains the same in this updated model, much of the south is predicted to be substantially more suitable: particularly pockets throughout Spain, France, Italy, Greece, the Balkans, Türkiye and North Africa.

Figure 8: Difference in probability of CCHFV suitability predictions between 2015 and 2022 models



As emphasised above, this reflects the addition of new training data for the 2022 model and changes in the masking used from country level (2015) to vector presences (2022). It should perhaps be noted that the two models were run at different spatial resolutions (2015 at 5km and 2022 at 1km), and so the differences are not strictly valid at the higher resolution, and the map in Figure 8 should only be used for reference on general areas where ecological suitability predictions are different.

Data availability

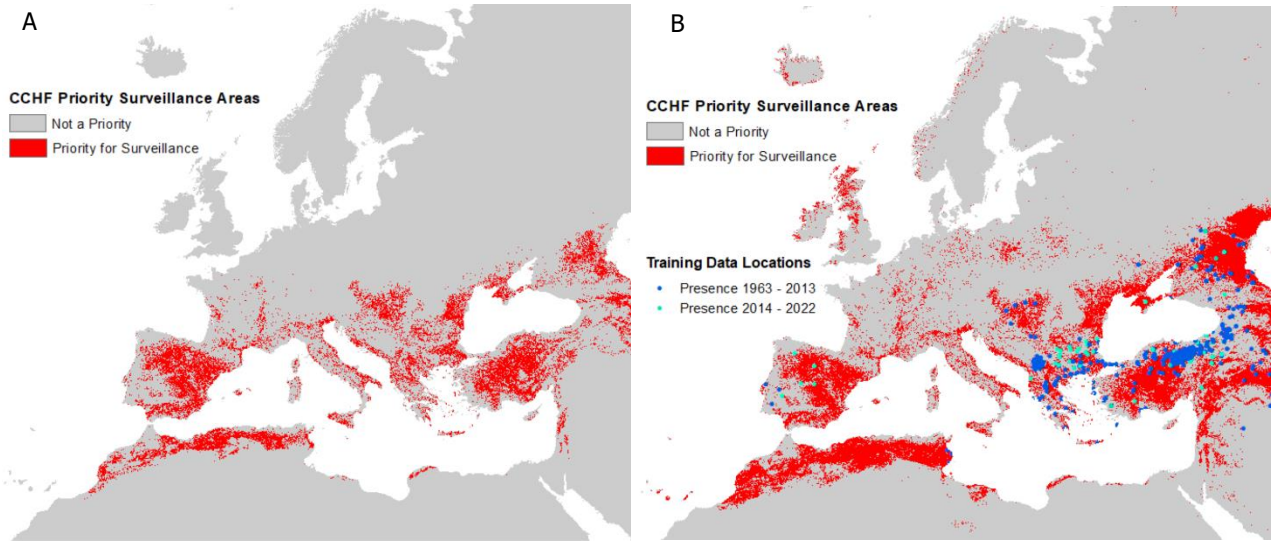
All data inputs and outputs are provided as two ARCGIS 10.8 Packages: 1) "VNEREGOCCHF_Vector_Report_VectorCCHFData_April23.mpk" containing the input and output Vector and CCHF datasets; and 2) VNEREGOCCHF_Vector_Report_CovariateData_April23.mpk. The files can be downloaded from the following link: https://tinyurl.com/VNCCHFDATAFEB23. Individual file descriptions are given in the package's Table of Contents, and the associated filenames can be obtained from the Layer Properties Source Tab. An endnote library (CCHF_2013_2022.xml) provides the references for all literature searched for CCHFV occurrences, and a second file (cchf_report_2023_final.xml) provides all references cited in the report.

Discussion

This report provides updated baseline data for monitoring future changes in the distribution of CCHFV and its associated vectors for Europe and its neighbouring areas. The predicted distributions suggest that a number of countries that have yet to record CCHFV have areas that are environmentally suitable for the disease —most especially those with Mediterranean coastlines (France, Italy, the southern Balkans, Western Asia and North Africa). Both the unmasked predictions and the maximum predicted extents of the two vector species (which better reflects the distribution of reported human CCHFV case locations than the medium extent) imply a potential CCHFV distribution extending significantly northward into mainland western and central Europe.

Further cartographic refinements are required in order to help differentiate endemic from epidemic-prone areas, particularly in areas where there is less certainty about the presence or absence of CCHFV overall. The CCHFV occurrence database used in creating the CCHFV ecological suitability map has been updated for the year 2022, and can continue to be updated with new information as it becomes available, leading to better future predictions. Other potential improvements are the inclusion of other potential CCHFV vector species into the masking process (discussed below) and the addition of more data types, such as those related to CCHFV seroprevalence in animals. For the latter, recent research which has elucidated potentially important CCHFV reservoir species in non-endemic areas (e.g. France, Italy) should also be considered; for example, imported sheep and goats may be more susceptible to CCHFV infection and therefore might promote virus circulation where suitable vectors are present (32,33). These efforts to improve mapping could enable identification of foci of potential future transmission to.

Figure 9: Maps of high predicted CCHFV suitability



A) *Left: areas with predicted probability of disease >= 0.75 masked to maximum predicted vector extent; B) Right: same as left but with recorded occurrence points and no vector mask*)

Already, these maps have been improved from the 2015 Messina *et al.* (1) versions due to not only the updated occurrence database and higher spatial resolution (now 1km x 1km), which includes significant new data from Türkiye, Spain and the Caucasus, but also the inclusion of vector suitability instead of Evidence Consensus to mask disease predictions. The left panel of **Error! Reference source not found.** shows areas where the m asked CCHFV probability of occurrence is high (greater than or equal to 0.75) within the maximum predicted presence of either suitable vector species, and, as such, may be used as a guide for prioritising future surveillance efforts. Although a similar occurrence dataset was used in the current model and that of Okely *et al.* (2020) (2), our predictions do not have the high ecological suitability in Poland demonstrated in their maps; in contrast, greater ecological suitability for CCHFV is seen in southern Russia and Kazakhstan. This is likely to be due to the different set of covariates used in each set of models (for example, Okely *et al.* used a Principal Components Analysis to reduce all available bioclimatic variables to a small number of factors, whilst our models selected those considered most important according to the literature, as well as a measure of shrub land cover). The equivalent high probability maps for each vector species are presented in upplementary Figure 16, Appendix. It can be seen that areas bordering the Adriatic and Mediterranean Seas are likely to be particularly important for vector and disease surveillance, as well as the majority of Portugal, Spain and Türkiye.

It is worth noting that the unmasked map of CCHFV suitability (right panel of **Error! Reference source not f ound.**) suggests many areas (*e.g.* Scotland) to be environmentally suitable for CCHFV occurrence that are far beyond recorded occurrence locations and the predicted range of the two vector species. This underlines the need for vector masking to produce a realistic disease prediction but does also suggest that if the vectors spread to these areas, then the disease might as well.

Whilst the predicted vector and vector-masked disease outputs reflect the known data well for most of the area of interest, recorded (and unmasked predicted) disease occurrence in certain parts of Western Asia is probably incorrectly masked by the vector distribution mask. The literature reports *H. marginatum* to be present (widespread) in Iran (19,34,35), further south than the distribution range boundary predicted by Kolonin (19), so it is likely that the point data available is substantially underrepresented for these regions (Iraq and Iran are not inside the region covered by VectorNet) and also perhaps some of the CCHFV pseudoabsences were incorrectly placed in suitable areas in this region. Otherwise, it may be that the vector (and thus the disease) survives in pockets (perhaps of irrigation) that these spatial suitability models do not 'detect'. If so, this suggests that the unmasked disease models are more reliable in such areas. Also, number of other *Hyalomma* species including *H. impeltatum*, *H. anatolicum*, *H. asiaticum* , *H. excavatum* and *H. dromedarii* have been identified as CCHFV vectors in Iran and its neighbours (19,36). These are not present in Europe (and there are few if any geo-referenced occurrence records) and were not incorporated into the vector masking —with the main focus being on Europe— but could improve the prediction in Western Asia.

A further potential improvement would be masking future maps according to the presence of host species for ticks, such as deer (for *H. marginatum*) and hares (for *H. lusitanicum*) and limiting surveillance priority to within rural areas and/or those which are nearer to livestock populations.

The updated and improved maps presented in this report could serve as a starting point for a wider discussion about the potential future likelihood of occurrence of CCHFV in the European region. By increasing public health awareness in geographic areas which we have newly shown to have a high probability of occurrence, more occurrence records may be generated leading to better models and improved identification of surveillance priority areas. As evidenced by reports of the disease occurring in Spain in 2016 (before which a 2015 publication did not identify this as a suitable CCHF transmission area), it is important for the mapping of high-suitability areas to be updated regularly, taking into consideration advancements in understanding of the disease distribution, as well as new covariate data sources.

References

1. Messina JP, Pigott DM, Golding N, Duda KA, Brownstein JS, Weiss DJ, et al. The global distribution of Crimean-Congo hemorrhagic fever. Trans R Soc Trop Med Hyg. 2015 Aug;109(8):503–13.

2. Okely M, Anan R, Gad-Allah S, Samy AM. Mapping the environmental suitability of etiological agent and tick vectors of Crimean-Congo hemorrhagic fever. Acta Trop. 2020 Mar 1;203:105319.

3. Hoogstraal H. The epidemiology of tick-borne Crimean-Congo hemorrhagic fever in Asia, Europe, and Africa. J Med Entomol. 1979 May 22;15(4):307–417.

4. Whitehouse CA. Crimean-Congo hemorrhagic fever. Antivir Res. 2004 Dec;64(3):145–60.

5. Han N, Rayner S. Epidemiology and mutational analysis of global strains of Crimean-Congo haemorrhagic fever virus. Virol Sin. 2011 Aug;26(4):229–44.

6. Ergonul O. Crimean-Congo haemorrhagic fever. Lancet Infect Dis. 2006 Apr;6(4):203–14.

7. Papa A, Bino S, Llagami A, Brahimaj B, Papadimitriou E, Pavlidou V, et al. Crimean-Congo hemorrhagic fever in Albania, 2001. Eur J Clin Microbiol Infect Dis. 2002 Aug;21(8):603–6.

8. Karti SS, Odabasi Z, Korten V, Yilmaz M, Sonmez M, Caylan R, et al. Crimean-Congo hemorrhagic fever in Turkey. Emerg Infect Dis. 2004 Aug;10(8):1379–84.

9. Zakhashvili K, Tsertsvadze N, Chikviladze T, Jghenti E, Bekaia M, Kuchuloria T, et al. Crimean-Congo hemorrhagic fever in man, Republic of Georgia, 2009. Emerg Infect Dis. 2010 Aug;16(8):1326–8.

10. Cases of Crimean–Congo haemorrhagic fever in the EU/EEA, 2013–present [Internet]. [cited 2023 Feb 24]. Available from: https://www.ecdc.europa.eu/en/crimean-congo-haemorrhagic-fever/surveillance/cases-eu-since-2013

11. Maltezou HC, Andonova L, Andraghetti R, Bouloy M, Ergonul O, Jongejan F, et al. Crimean-Congo hemorrhagic fever in Europe: current situation calls for preparedness. Euro Surveill. 2010 Mar 11;15(10):19504.

12. Grard G, Drexler JF, Fair J, Muyembe JJ, Wolfe ND, Drosten C, et al. Re-emergence of Crimean-Congo hemorrhagic fever virus in Central Africa. PLoS Negl Trop Dis. 2011 Oct;5(10):e1350.

13. Appannanavar SB, Mishra B. An Update on Crimean Congo Hemorrhagic Fever. J Glob Infect Dis. 2011;3(3):285–92.

14. Maltezou HC, Papa A. Crimean-Congo hemorrhagic fever: risk for emergence of new endemic foci in Europe? Travel Med Infect Dis. 2010 May;8(3):139–43.

15. Lutomiah J, Musila L, Makio A, Ochieng C, Koka H, Chepkorir E, et al. Ticks and tick-borne viruses from livestock hosts in arid and semiarid regions of the eastern and northeastern parts of Kenya. J Med Entomol. 2014 Jan;51(1):269–77.

16. Okello-Onen J, Tukahirwa EM, Perry BD, Rowlands GJ, Nagda SM, Musisi G, et al. Population dynamics of ticks on indigenous cattle in a pastoral dry to semi-arid rangeland zone of Uganda. Exp Appl Acarol. 1999 Jan;23(1):79–88.

17. Aktas M, Altay K, Dumanli N. A molecular survey of bovine Theileria parasites among apparently healthy cattle and with a note on the distribution of ticks in eastern Turkey. Vet Parasitol. 2006 Jun 15;138(3–4):179–85.

18. Horak IG, Braack LE, Fourie LJ, Walker JB. Parasites of domestic and wild animals in South Africa. XXXVIII. Ixodid ticks collected from 23 wild carnivore species. Onderstepoort J Vet Res. 2000 Dec;67(4):239–50.

19. Kolonin G. Fauna of ixodid ticks of the world (Acari, Ixodidae). 2009.

20. Estrada-Peña A, Venzal J. Climate Niches of Tick Species in the Mediterranean Region: Modeling of Occurrence Data, Distributional Constraints, and Impact of Climate Change. J Med Entomol. 2007 Dec 1;44:1130–8.

21. Wint GRW, et al. VectorNet: Collaborative mapping of standardised distributions and surveillance for arthropod disease vectors in Europe and neighbouring countries. Eurosurveillance. In press(In press).

22. Wint W, Van Bortel W, Schaffner F. RVF vector spatial distribution models: Probability of presence. EFSA Support Publ. 2020;17(2):1800E.

23. Scharlemann JPW, Benz D, Hay SI, Purse BV, Tatem AJ, Wint GRW, et al. Global Data for Ecology and Epidemiology: A Novel Algorithm for Temporal Fourier Processing MODIS Data. PLOS ONE. 2008 Jan 9;3(1):e1408.

24. Fernández-Ruiz N, Estrada-Peña A. Towards New Horizons: Climate Trends in Europe Increase the Environmental Suitability for Permanent Populations of Hyalomma marginatum (Ixodidae). Pathogens. 2021 Feb;10(2):95.

25. Kimball A, Hatfield KM, Arons M, James A, Taylor J, Spicer K, et al. Asymptomatic and Presymptomatic SARS-CoV-2 Infections in Residents of a Long-Term Care Skilled Nursing Facility - King County, Washington, March 2020. MMWR Morb Mortal Wkly Rep. 2020 Apr 3;69(13):377–81.

26. Santos-Silva MM. Hyalomma lusitanicum Koch, 1844 (Figs. 155–157). In: Estrada-Peña A, Mihalca AD, Petney TN, editors. Ticks of Europe and North Africa: A Guide to Species Identification [Internet]. Cham: Springer International Publishing; 2017 [cited 2022 Dec 16]. p. 383–7. Available from: https://doi.org/10.1007/978-3-319-63760-0_72

27. Williams HW, Cross DE, Crump HL, Drost CJ, Thomas CJ. Climate suitability for European ticks: assessing species distribution models against null models and projection under AR5 climate. Parasit Vectors. 2015 Aug 28;8(1):440.

28. Estrada-Pena A, Ayllon N, De La Fuente J. Impact of Climate Trends on Tick-Borne Pathogen Transmission. Front Physiol [Internet]. 2012 [cited 2022 Dec 16];3. Available from: https://www.frontiersin.org/articles/10.3389/fphys.2012.00064

29. Estrada-Peña A, Sánchez N, Estrada-Sánchez A. An Assessment of the Distribution and Spread of the Tick Hyalomma marginatum in the Western Palearctic Under Different Climate Scenarios. Vector-Borne Zoonotic Dis. 2012 Sep;12(9):758–68.

30. Estrada-Peña A, Jameson L, Medlock J, Vatansever Z, Tishkova F. Unraveling the Ecological Complexities of Tick-Associated Crimean-Congo Hemorrhagic Fever Virus Transmission: A Gap Analysis for the Western Palearctic. Vector-Borne Zoonotic Dis. 2012 Sep;12(9):743–52.

31. Gray JS, Dautel H, Estrada-Peña A, Kahl O, Lindgren E. Effects of climate change on ticks and tick-borne diseases in europe. Interdiscip Perspect Infect Dis. 2009;2009:593232.

32. Spengler JR, Bergeron É, Rollin PE. Seroepidemiological Studies of Crimean-Congo Hemorrhagic Fever Virus in Domestic and Wild Animals. PLoS Negl Trop Dis. 2016 Jan 7;10(1):e0004210.

33. Grech-Angelini S, Lancelot R, Ferraris O, Peyrefitte CN, Vachiery N, Pédarrieu A, et al. Crimean-Congo Hemorrhagic Fever Virus Antibodies among Livestock on Corsica, France, 2014–2016. Emerg Infect Dis. 2020 May;26(5):1041–4.

34. Kadir MA, Zangana IK, Mustafa BHS. A study on epidemiology of hard tick (Ixodidae) in sheep in Sulaimani governorate - Iraq. Iraqi J Vet Sci. 2012 May 1;26(Suppl. III):95–103.

35. Ismael S, Omer LT. Molecular identification of new circulating Hyalomma asiaticum asiaticum from sheep and goats in Duhok governorate, Iraq. Iraqi J Vet Sci. 2021 Jan 1;35(1):79–83.

36. Telmadarraiy Z, Chinikar S, Vatandoost H, Faghihi F, Hosseini-Chegeni A. Vectors of Crimean Congo Hemorrhagic Fever Virus in Iran. J Arthropod-Borne Dis. 2015 Mar 11;9(2):137–47.

37. Kraemer MUG, Reiner RC, Brady OJ, Messina JP, Gilbert M, Pigott DM, et al. Past and future spread of the arbovirus vectors Aedes aegypti and Aedes albopictus. Nat Microbiol. 2019 May;4(5):854–63.

Appendix

Supplementary Table 1: Vector habitat suitability conditions

Variable	*H. marginatum*	*H. lusitanicum*	**Dataset**
Forest/woodland > 40%	No	Yes	Corine/ESA CCI
Woodland or Forest 15 - 40%	Yes	Yes	Corine/ESA CCI
Shrubland/Grassland	Yes	Yes	Corine/ESA CCI
Cropland	Yes	Yes	Corine/ESA CCI
Sparse Vegetation	Yes	No	Corine/ESA CCI
Minimum Elevation	< 2000m	< 2000m	GMTED (see URL below)
Max temp	< 35 °C	< 40 °C	Worldclim (see URL below)
Summer minimum Relative Humidity		>10%, < 65%	Derived from MODIS LST (see URL below)
Cumulative temperature above 15°C, April to August	>800, plus 50km buffer		Worldclim (see URL below)
Cumulative minimum temperature above 0 °C, October and November		>400	Worldclim (see URL below)
Mean Annual Vapour Pressure Deficit	< 15%		ECWMF (see URL below)
Corine Land Cover 2018 https://land.copernicus.eu/pan-european/corine-land-cover, ESA CCI land Cover http://maps.elie.ucl.ac.be/CCI/viewer/download.php GMTED https://topotools.cr.usgs.gov/gmted_viewer/gmted2010_global_grids.php Modis Land Surface Temperature https://modis.gsfc.nasa.gov/data/dataprod/mod11.php Worldclim: https://www.worldclim.org/data/index.html ECMWF: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5			

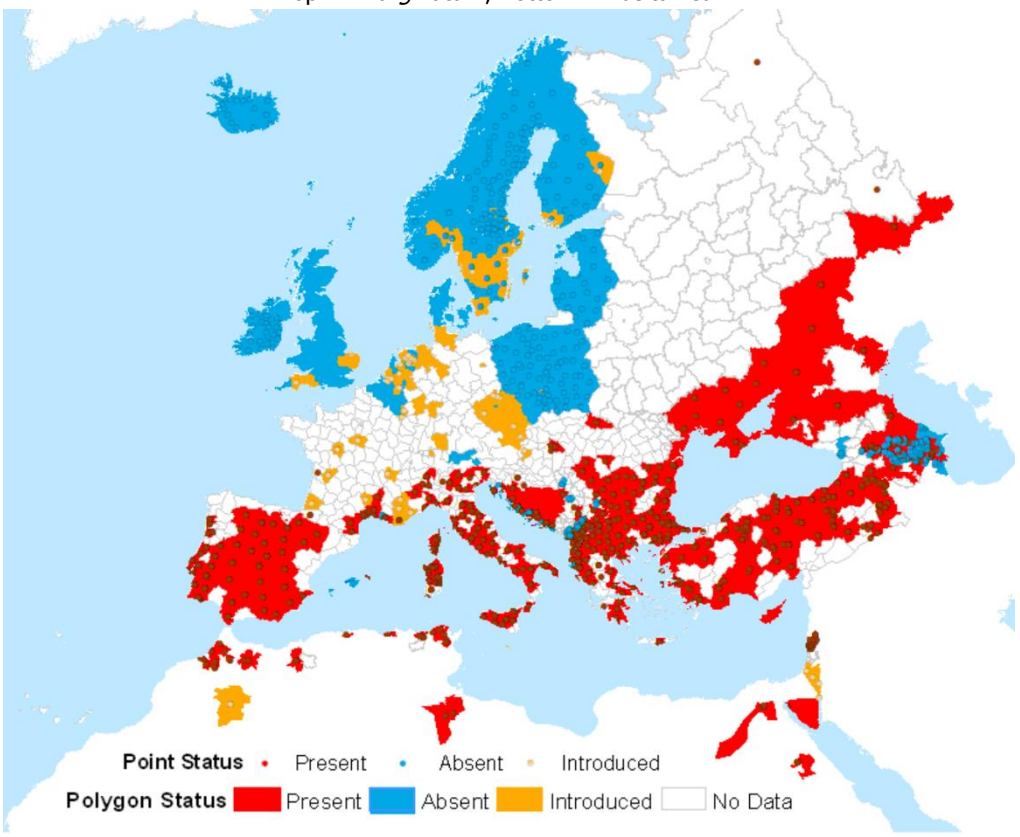
Supplementary Table 2: Covariates offered to modelling procedures

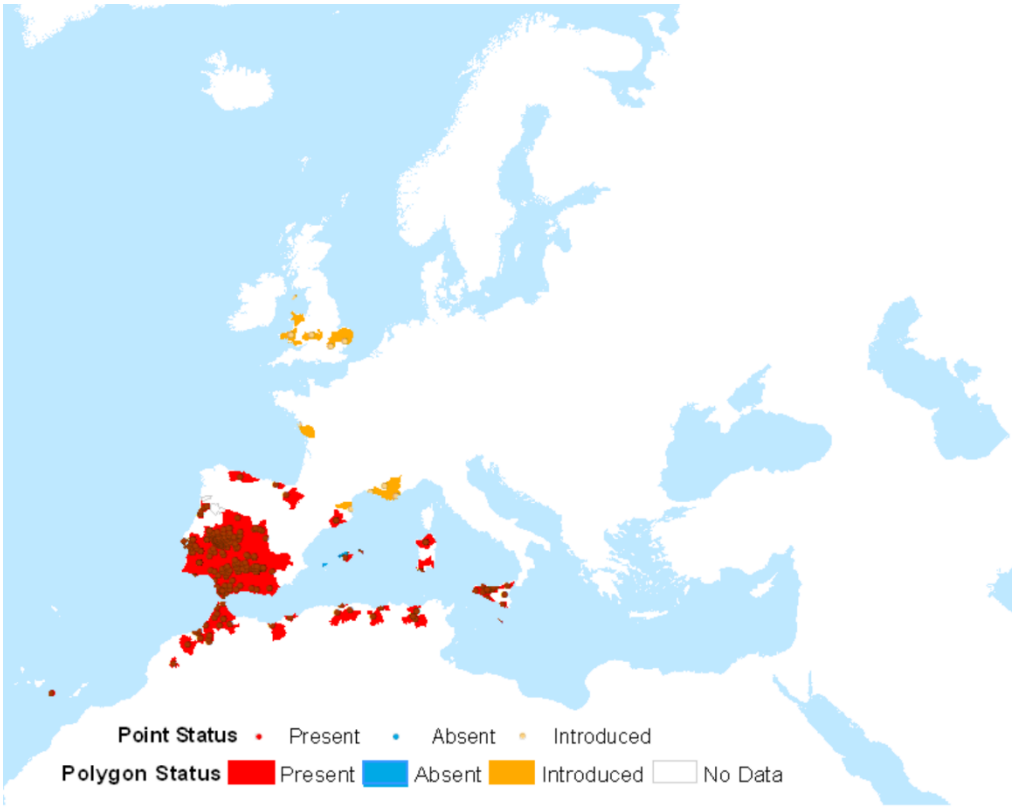
1 VCC1103A0: Middle infra-red mean	45 VCC1115MN: EVI minimum
2 VCC1103A1: Middle infra-red amplitude 1	46 VCC1115MX: EVI maximum
3 VCC1103A2: Middle infra-red amplitude 2	47 VCC1115P1: EVI phase 1
4 VCC1103A3: Middle infra-red amplitude 3	48 VCC1115P2: EVI phase 2
5 VCC1103MN: Middle infra-red minimum	49 VCC1115P3: EVI phase 3
6 VCC1103MX: Middle infra-red maximum	50 VCC1115VR: EVI variance
7 VCC1103P1: Middle infra-red phase 1	51 VCMI30GRDP1K DEM (Elevation) +1000
8 VCC1103P2: Middle infra-red phase 2	52 VC1920A0: ERA5 Precipitation mean
9 VCC1103P3: Middle infra-red phase 3	53 VC1920A1: ERA5 Precipitation amplitude 1
10 VCC1103VR: Middle infra-red variance	54 VC1920A2: ERA5 Precipitation amplitude 2
11 VCC1107A0: Daytime LST mean	55 VC1920A3: ERA5 Precipitation amplitude 3
12 VCC1107A1: Daytime LST amplitude 1	56 VC1920MN: ERA5 Precipitation minimum
13 VCC1107A2: Daytime LST amplitude 2	57 VC1920MX: ERA5 Precipitation maximum
14 VCC1107A3: Daytime LST amplitude 3	58 VC1920P1: ERA5 Precipitation phase 1
15 VCC1107MN: Daytime LST minimum	59 VC1920P2: ERA5 Precipitation phase 2
16 VCC1107MX: Daytime LST maximum	60 VC1920P3: ERA5 Precipitation phase 3
17 VCC1107P1: Daytime LST phase 1	61 VC1920VR: ERA5 Precipitation variance
18 VCC1107P2: Daytime LST phase 2	62 VCWPOPPPP: Worldpop Human Population density 2020
19 VCC1107P3: Daytime LST phase 3	63 VCV59EL500: GNTED Elevation + 500
20 VCC1107VR: Daytime LST variance	64 VCEELCBARE: consensus % bare ground
21 VCC1108A0: Nighttime LST mean	65 VCEELCDCBD3: consensus % deciduous broadleaved forest
22 VCC1108A1: Nighttime LST amplitude 1	
23 VCC1108A2: Nighttime LST amplitude 2	66 VCEELCEVGBD2: consensus % evergreen broadleaved forest
24 VCC1108A3: Nighttime LST amplitude 3	
25 VCC1108MN: Nighttime LST minimum	67 VCEELCEVGND1: consensus % evergreen needleleaved forest
26 VCC1108MX: Nighttime LST maximum	
27 VCC1108P1: Nighttime LST phase 1	68 VCEELCFLOOD8: consensus % flooded
28 VCC1108P2: Nighttime LST phase 2	69 VCEELCHVCB: consensus % herbaceous cover
29 VCC1108P3: Nighttime LST phase 3	70 VCEELCMANAG7: consensus % managed land
30 VCC1108VR: Nighttime LST variance	71 VCEELCOTHTR4: consensus % other land cover
31 VCC1114A0: NDVI mean	72 VCEELCSHRUB5: consensus % shrub cover
32 VCC1114A1: NDVI amplitude 1	73 VCEELCURB9: consensus % urban
33 VCC1114A2: NDVI amplitude 2	74 VCEELCWATER12: consensus % water
34 VCC1114A3: NDVI amplitude 3	75 VC82094A0: Relative Humidity mean
35 VCC1114MN: NDVI minimum	76 VC82094A1: Relative Humidity amplitude 1
36 VCC1114MX: NDVI maximum	77 VC82094A2: Relative Humidity amplitude 2
37 VCC1114P1: NDVI phase 1	78 VC82094A3: Relative Humidity amplitude 3
38 VCC1114P2: NDVI phase 2	79 VC82094MN: Relative Humidity minimum
39 VCC1114P3: NDVI phase 3	80 VC82094MX: Relative Humidity maximum
40 VCC1114VR: NDVI variance	81 VC82094P1: Relative Humidity phase 1
41 VCC1115A0: EVI mean	82 VC82094P2: Relative Humidity phase 2
42 VCC1115A1: EVI amplitude 1	83 VC82094P3: Relative Humidity phase 3
43 VCC1115A2: EVI amplitude 2	84VC82094VR: Relative Humidity variance
44 VCC1115A3: EVI amplitude 3	

LST = Land Surface Temperature. NDVI Normalised Difference vegetation Index; EVI Enhanced Vegetation Index. DEM Digital Elevation. All files starting with VCC11, VC19 and VC82 are Fourier processed MODIS Satellite Imagery produced by the Environmental Research Group Oxford according to the methods set out in Scharlemann *et. al.* (2008) (23).
The Relative Humidity layers were produced as described in n Kraemer *et. al.* (2019)(37) and then Fourier Processed as described above.
The Elevation layer was extracted from the GMTED datasets (https://topotools.cr.usgs.gov/gmted_viewer/gmted2010_global_grids.php) and the negative values removed by adding 1000.
Population layers derived from layers produced by worldpop (https:// www.worldpop.org/datacatalog/)
All Files with VCEELC in file name were derived from the Earthenv consensus land cover data product (https://www.earthenv.org/landcover)
All layers extracted and standardised by ERGO for MOOD Horizon 2020 project N° 874850 (https://mood-h202.eu)

Supplementary Figure 10:VectorNet data locations

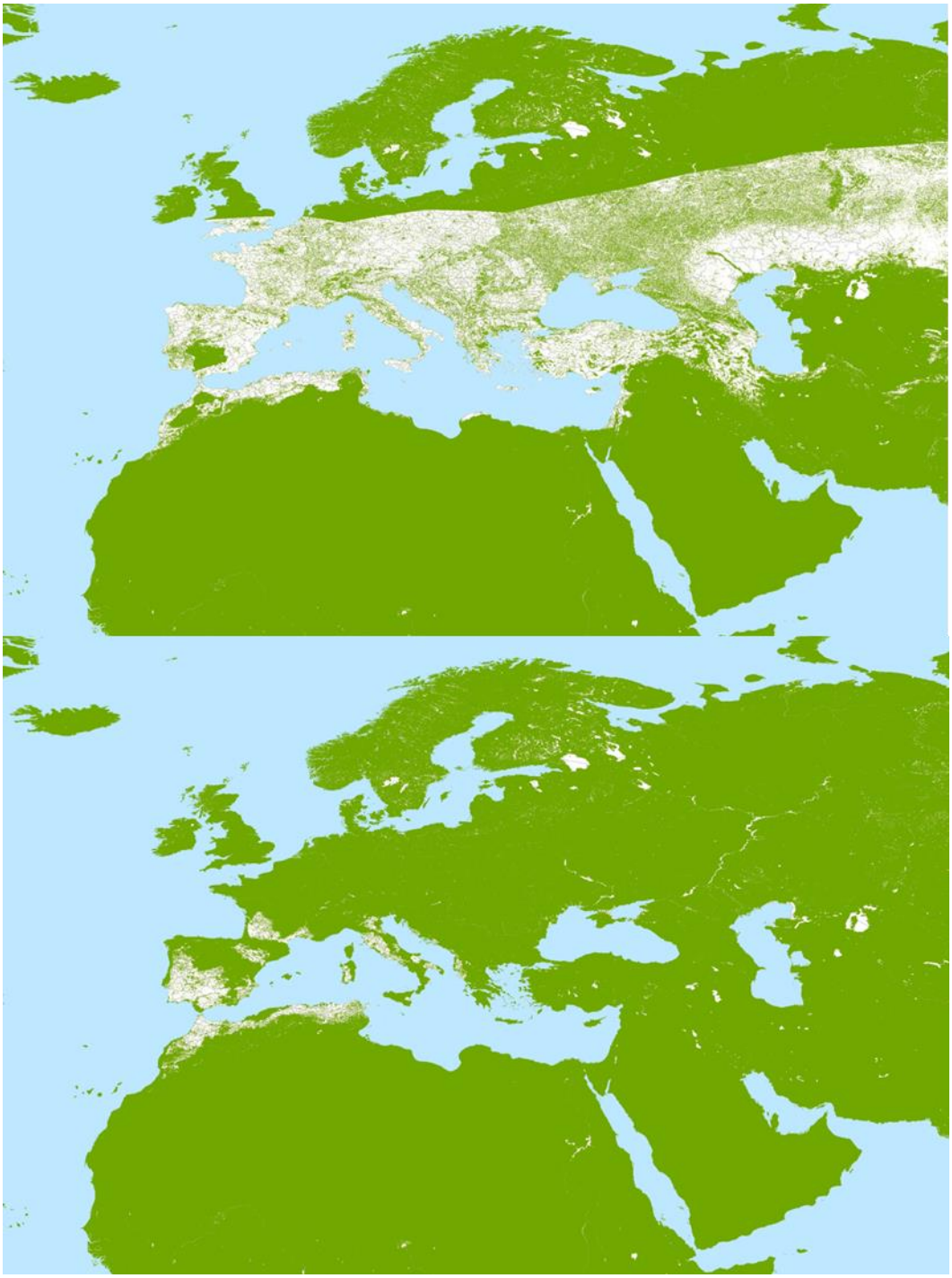
Top *H. marginatum*, Bottom *H. lusitanicum.*



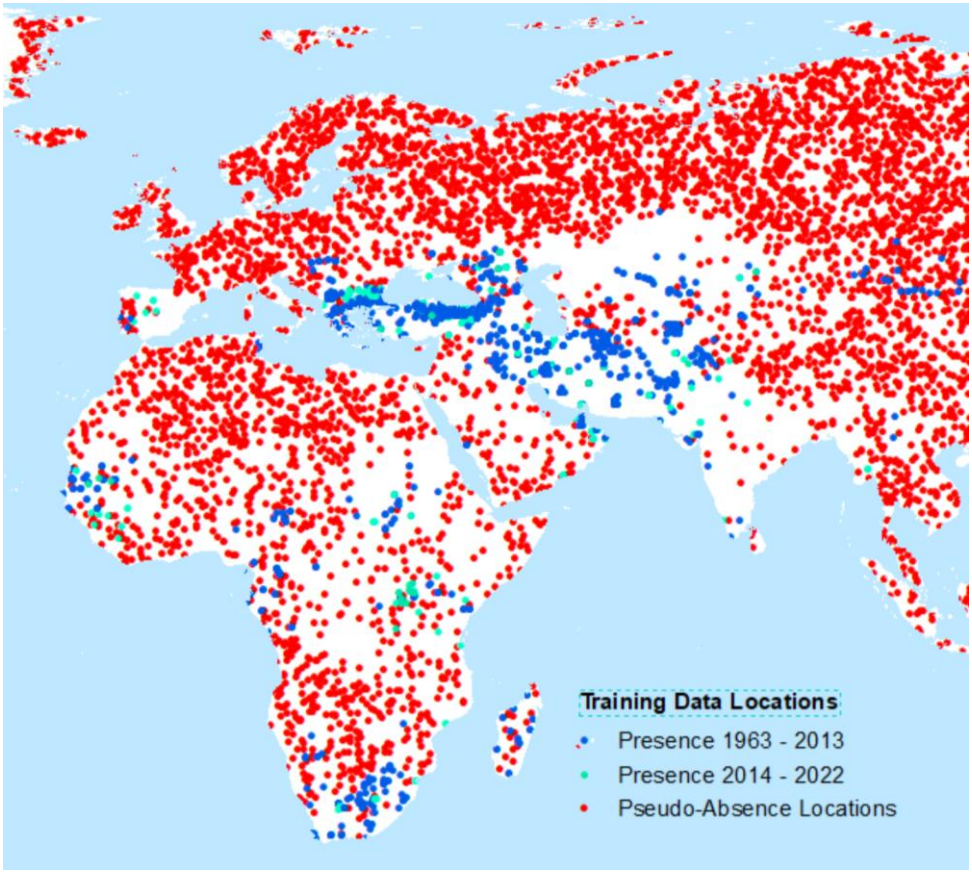


Supplementary Figure 11: Habitat suitability for each vector.

Top *H. marginatum*, Bottom *H. lusitanicum.* Unsuitable = Green, Suitable=white



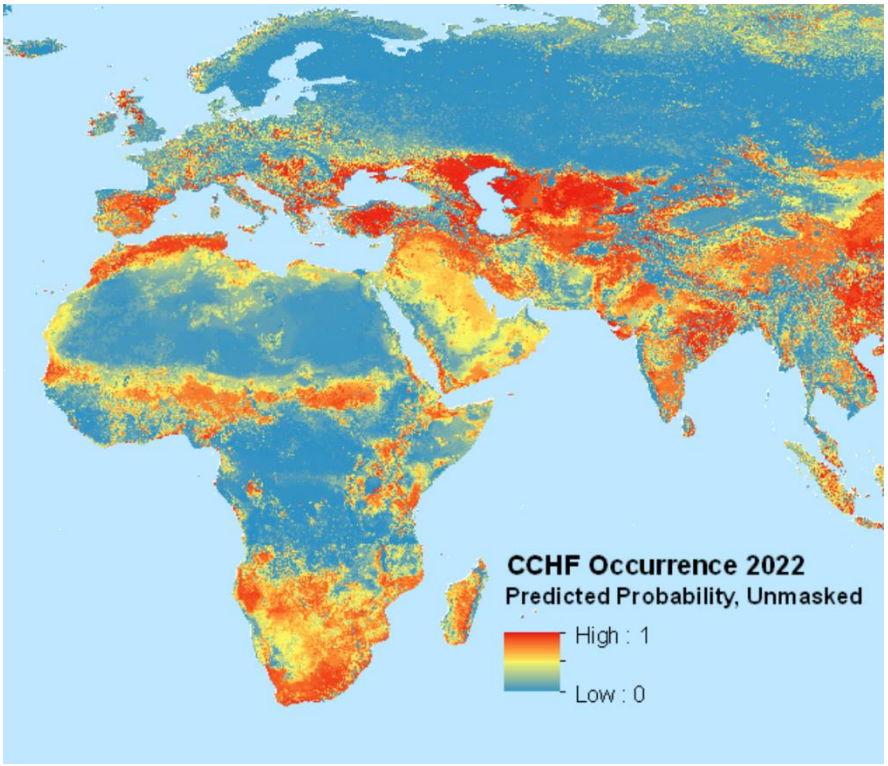
Supplementary Figure 12: Full extent of CCHF pseudo-absence and occurrence locations



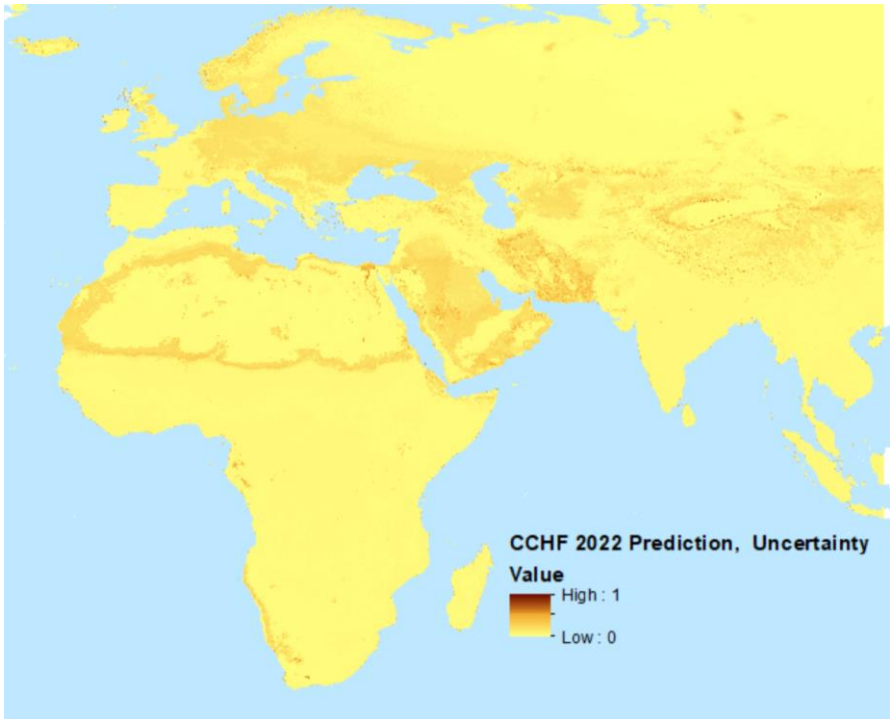
H. lusitanicum				*H. marginatum*			
RF		**BRT**		**RF**		**BRT**	
Predictor	**Metric**	**Predictor**	**Metric**	**Predictor**	**Metric**	**Predictor**	**Metric**
Rainfall Amplitude Component 1	20.44	Rainfall Amplitude Component 1	2.57	NDVI Minimum	9.84	Day Temperature Mean	1.66
Day Temperature Phase Component 2	17.2	Rainfall Amplitude Component 2	2.45	Rainfall Amplitude Component 2	9.8	NDVI Minimum	1.62
Rainfall Minimum	11.68	Day Temperature Phase Component 2	2.17	Day Temperature Mean	7.3	Rainfall Amplitude Component 2	1.55
Night Temperature Amplitude Component 1	7.1	Infra Red Amplitude Component 1	1.92	Rainfall Phase Component 3	4.36	Rainfall Amplitude Component 1	1.46
Night Temperature Minimum	4.12	EVI Phase Component 1	1.62	Rainfall Phase Component 1	4.34	Night Temperature Phase Component 2	1.41
Rainfall Amplitude Component 2	3.8	Infra Red Phase Component 1	1.52	Bare Ground Proportion	3.48	Night Temperature Maximum	1.35
Rainfall Phase Component 2	2.91	Night Temperature Amplitude Component 1	1.52	Night Temperature Phase Component 2	2.95	Rainfall Maximum	1.29
Infra Red Phase Component 1	2.59	Day Temperature Mean	1.4	Relative Humidity minimum	2.85	Day Temperature Phase Component 1	1.29
Night Temperature Mean	2.12	NDVI Phase Component 1	1.3	Relative Humidity Phase Component 3	2.8	NDVI Amplitude Component 2	1.26
Relative Humidity Amplitude Component 2	1.91	Infra Red Phase Component 2	1.28	Rainfall Phase Component 2	2.77	Day Temperature Maximum	1.25

Note: for both RF and BRT the higher the metric within each method the more important the predictor. The Metrics are different for each method, and should not be compared

Supplementary Figure 13: Full extent of modelled CCHF suitability map

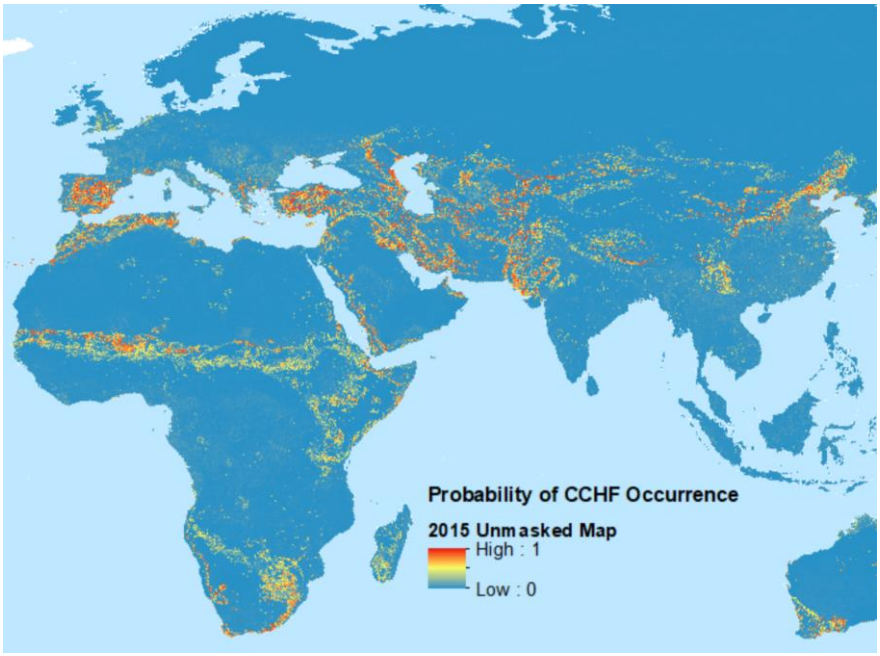


Supplementary Figure 14: Uncertainty estimates for CCHF suitability estimates (probability of occurrence)

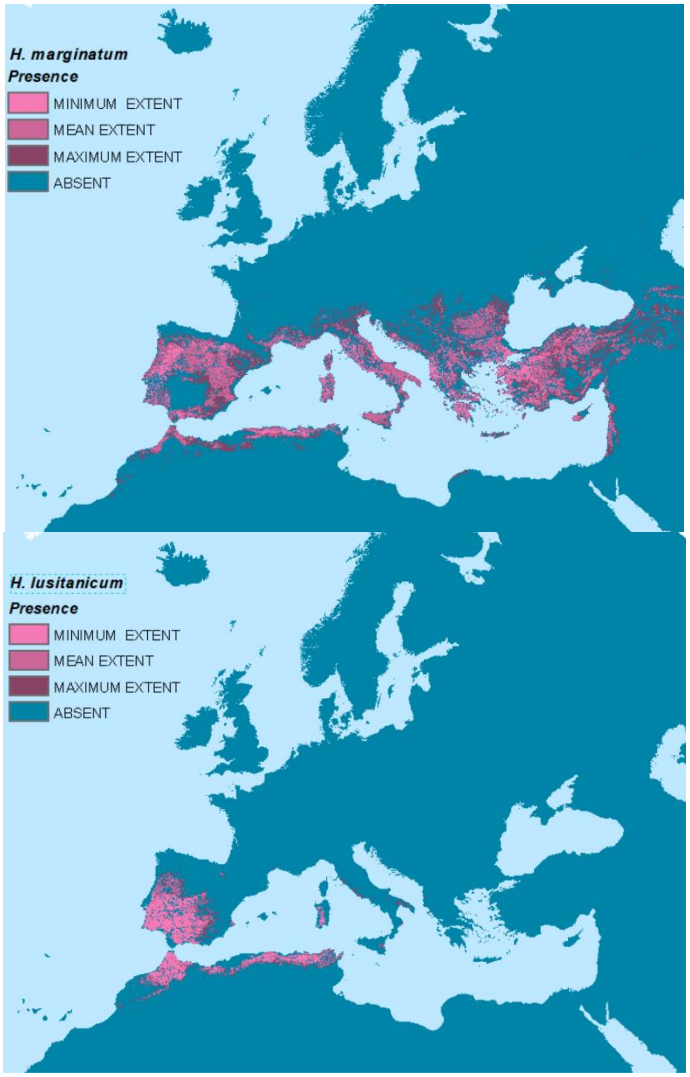


Uncertainty ranges from 0 to 1, with higher values indicating greater ranges in the estimated probability of occurrence.

Supplementary Figure 15: Unmasked 2015 CCHF prediction from Messina et al (2015).



Supplementary Figure 16: Areas with > 0.75 probability of vector presence



Terms of Reference

The main deliverable is a technical report geographically mapping the potential risk of Crimean-Congo haemorrhagic fever (CCHF) in the VectorNet geographic area, based on ecological predictions of CCHF vector tick (*Hyalomma marginatum* and *H. lusitanicum*) present distributions in the western Palearctic area combined with information from global CCHF incidence and prevalence studies and environmental drivers.

Messina and colleagues (2015) [https://doi.org/10.1093/trstmh/trv050] mapped CCHF environmental suitability globally, masking out areas without known CCHF occurrence at the time, such as Spain, where CCHF was detected more recently. Okely and colleagues (2020) [https://doi.org/10.1016/j.actatropica.2019.105319] built on this work, updating the global CCHF risk map and comparing it (unmasked) with ecological niche maps of seven ticks species in the Old World, but predicting (unrealistically) high CCHF risk and *H. marginatum* presence in eastern Germany, Poland and Hungary.

The report should document updating the global CCHF database and risk map with masked and unmasked versions; a methodology to further inform CCHF risk estimates in western Palearctic area with (realistic) *H. marginatum* and *H. lusitanicum* ticks distributions, and resulting modified maps of CCHF.

The production of the needed *H. marginatum* and *H. lusitanicum* ecological occurrence maps (based on VectorNet and other available data) is not covered by this contract but are expected to be provided through a collaboration with the MOOD project at no cost to ECDC. The ecological tick occurrence maps should feed into the CCHF risk maps directly as a covariate or indirectly as graduated (based on habitat suitability / probability) or Boolean mask.

An interim report is required seven months after signature, and the final report is required nine months after contract signature.

Together with the report, the contractor should provide:

1. The updated datasets of CCHF and covariates
2. The input and output map raster files.
3. The analysis scripts
4. The references library in a format compatible with EndNote X7.