Modelling habitat suitability for occurrence of human tick-borne encephalitis (TBE) cases in Finland

Ruut Uusitalo, Mika Siljander, Timothée Dub, Jussi Sane, Jani J. Sormunen, Petri Pellikka, Olli Vapalahti

1. Introduction

High-latitude regions (> 60°) in the Northern Hemisphere are undergoing rapid changes associated with climate warming. Climate change interacts with the global change through atmospheric circulation and biogeophysical and biogeochemical effects; such changes are particularly affected by rising temperatures at high latitudes (Groisman and Soja, 2007). As a result of these climatic changes, distribution of invasive species is expanding to new regions and environmental conditions have become more suitable for the circulation of several vector-borne viruses (Koch et al., 2017; Hobbs, 2000; Jore et al., 2014; Sutherst et al., 2000). Ensemble modelling and geographical...
information systems (GIS) have been used to understand the connections between vectors, their habitats, and vector-borne diseases (Honig et al., 2019; Stefanoff et al., 2018; Gama et al., 2017; Sun et al., 2017; Deka and Marsh, 2018). The distribution of vector-borne diseases remains understudied in many regions. GIS and ensemble modelling approaches used for identification of influential environmental factors and estimation of vector-borne disease risks can improve knowledge on disease prevention.

Tick-borne encephalitis (TBE) is a zoonotic disease caused by the TBE virus (TBEV). TBEV infection typically induces an influenza-like illness. In one third of cases, the initial illness may be followed by fever, meningitis, or meningoencephalitis. Neurological sequelae, including paralytic poliomyelitis, may occur. Death occurs in 1–2% of cases (World Health Organization (WHO), 2019). TBEV is found in endemic areas across large regions of the temperate and boreal forest regions of Europe and Asia (European Centre for Disease Prevention and Control (ECDC, 2015; Charrel et al., 2004; Woolhouse et al., 2001). The annual number of TBE cases in the European Union and the European Free Trade Association varies between 2000 and 3500 cases; most cases occur between June and September ( Beauté et al., 2018; European Centre for Disease Prevention and Control (ECDC, 2015). The Finnish Government has defined TBE as one of the major risks to public health in changing climate and has called for preparedness and risk assessment (Tuomenvirta et al., 2018).

Although ticks are both the vectors and the main reservoir for the virus, vertebrate species can also be infected by TBEV. Small mammals can harbor the virus and transmit it to their offspring. TBEV can be transmitted directly from an infected vertebrate host to its progeny or occasionally transovarially (Alekseev and Chunikhin, 1990). Transmission can also occur by co-feeding on continuously close to infected ticks (Labuda et al., 1993; Randolph et al., 1996). TBEV is transmitted by Ixodes spp. ticks from Western Europe through Russia to the Far East ( Lindquist and Vapalahi, 2008). Finland lies in the zone where the two tick species (Ixodes ricinus and I. persulcatus) overlap and they can transmit both the European and Siberian virus subtypes (Süss, 2011; Tonteri et al., 2015; Öhman, 1961; Jääskeläinen et al., 2006, 2010; Kuivanen et al., 2018). Ixodes ricinus is predominant in Southern Finland while I. persulcatus prevails in Northern Finland (Laaksonen et al., 2017). In Finland and in the neighboring countries Sweden and Russia, the host tick species and the distribution of TBEV have moved to more northern latitudes (Tokarevich et al., 2011; Jaenson et al., 2012; Laaksonen et al., 2017).

The natural cycles of TBEV are sensitive to various environmental and ecological factors, such as climate (Daniel et al., 2015; Semenza and Menne, 2009; Brabec et al., 2017), microclimate ( Haider et al., 2017; Randolph et al., 2001), and population density of tick hosts (Heyman et al., 2010; Brugger et al., 2017). Changes in weather conditions influence the distribution of vector ticks and host and reservoir animals and therefore affects pathogen transmission and incidence of human disease cases (Alikshie et al., 2017; Lindgren, 1998a). The likelihood of tick co-feeding (and consequently co-feeding transmission of TBEV), tick host densities, and human-tick encounters are all influenced by climatic associations. Seasonal tick activity starts when the daily temperature rises above 5 °C but ceases in temperatures above 25 °C (Daniel et al., 2015, 2018). Daylight length and relative humidity (RH), particularly long periods of low RH, affect tick activity (Gray et al., 1998; Daniel et al., 2015). In addition, RH also affects tick survival (Gray et al., 1998). The presence of suitable host animals also influences tick abundance. Adult I. ricinus and I. persulcatus normally feed on medium-sized and large-sized animals such as deer and hares, whereas nymphal ticks mainly feed on small to medium-sized animals such as rodents, birds and hares, but also on large-sized animals (Gray et al., 2016; Brugger et al., 2017). TBE has been found to correlate with the abundance of deer and hares in Sweden (Jaenson et al., 2018), and in Finland with white-tailed deer density (Dub et al., unpublished results).

Habitat suitability refers to condition, when the combination of abiotic environmental variables at the site is included in the environmental conditions that a species needs to grow and maintain viable populations (Hutchinson, 1992). This is the basis for habitat suitability modelling approach which can improve knowledge on the distribution of pathogens, vector-borne diseases, vectors and hosts in the changing climate ( Guisan and Zimmermann, 2000; Guisan et al., 2017). In this study, we use the definition of habitat suitability to refer to the ability of a habitat to promote the occurrence of TBE, and to the probability of TBE occurrence. Recently, ensemble modelling has been widely used to predict distribution of vectors (Miller et al., 2018; Chalghaf et al., 2018; Uusitalo et al., 2019) and vector-borne diseases (Deribe et al., 2018; Deka and Marsh, 2018; Gama et al., 2017; Eneanya et al., 2018). This tool also has the advantage of producing more robust decision making in the face of uncertainty in comparison with single-model forecasts (Araújo and New, 2007). Modelling studies of vector-borne diseases and their associations with climate change have been conducted since the 1990s ( Rogers and Packer, 1993; Haines et al., 2000). The Southern Finland coastline has been predicted to be suitable for TBEV circulation a few decades ago ( Randolf and Rogers, 2001), and several distinct foci have since been found in the region ( Jääskeläinen et al., 2016; Smura et al., 2019). However, recent spatial modelling studies on TBE in Finland are missing.

The aims of this study are to predict the distribution of TBE in the identified foci (Tonteri et al., 2015) in Finland based on known factors that affect TBE incidence; to discover which environmental, anthropogenic, and ecological factors best explain the probability of TBE disease case occurrence; and to analyze the effect of predicted climate change on TBE distribution based on bioclimatic variables. To reach these aims, we chose to apply biomod2, an ensemble platform for species distribution modelling (Thuiller et al., 2016).

2. Material and methods

2.1. TBE occurrence data and study datasets

Finland (59°50’N, 20°38’E, 70°09’N, 31°30’E) is located in Northern Europe between Sweden and Russia ( Fig. 1.). TBE occurrence data include serologically confirmed human TBE cases by municipality from 2007 to 2017 obtained from the National Infectious Diseases Register (NIDR). In total, 488 TBE cases were reported in 2007–2017. Patients infected abroad and whose location of exposure or date of onset were unknown, were excluded from this study (n = 158) (Fig. 1.).

Overall, the number of TBE cases has increased annually and the distribution of exposures has spread to new areas from 2007 to 2017 (Figs. 1 and 2). The geographical distribution of TBE cases is mainly focused in coastal and Southern Finland (including the Åland Islands). In the past 6 years, the distribution of TBE cases has also expanded to Central, Eastern, and Northern Finland, excluding northernmost Finland. The annual number of TBE cases has almost quadrupled in 11 years from 2007 ( n = 20) to 2017 ( n = 72).

TBE datasets were built based on NIDR data. Data were split into two different datasets (2007–2011 and 2012–2017), to identify differences in predictors and TBE risk between the two time periods. The first dataset includes an area of 35,344 km2 and consists of 24 municipalities with 86 TBE cases and 24 control municipalities without TBE cases (Fig. 3.). The second dataset includes an area of 83,720 km2, and consists of 51 (presence) municipalities with 244 TBE cases and 51 control (absence) municipalities (Fig. 3.). Absence data were randomly selected from municipalities located in the vicinity of TBE municipalities.

2.2. Environmental, ecological and anthropogenic data

The selection of explanatory data were based on those introduced in the existing literature. Environmental data for Finland were obtained directly from satellite imagery, or were derived from the satellite imagery in ESRI ArcGIS version (10.3.1). Ecological data included tick
data and hunting data of selected game animals. Anthropogenic data consisted of data on human population density and people working in primary production. Altogether, explanatory data included 50 predictors (excluding separate variables for each month e.g. temperature) before modelling analysis (Table 1.; Table A.1).

Bioclimatic variables \((N = 19)\) were derived from the WorldClim datasets (Hijmans et al., 2005; Fick and Hijmans, 2017). Precipitation, temperature, land surface temperature, snow cover, and vegetation data were derived from satellite imagery. Due to the lack of real game animal density data in Finland, we used hunting data of known tick host game animals as game animal densities, as hunting data correlates with animal densities (Cattadori et al., 2003; Jore et al., 2014). The data consists of hunted European hares \((Lepus europaeus)\), mountain hares \((Lepus timidus)\), moose \((Alces alces)\), red foxes \((Vulpes vulpes)\), roe deer \((Capreolus capreolus)\), white-tailed deer \((Odocoileus virginianus)\), raccoon dogs \((Nyctereutes procyonoides)\), martens \((Martes spp.)\), minks \((Neovison vison)\), fowls, waterfowls, and other wildfowls. For the Åland Islands, only annual hunting data of moose, white-tailed deer, and roe deer was available.

2.3. Data analysis and habitat suitability modelling

All the geospatial datasets including environmental and other attributes, were set to the same spatial extent, geographic coordinate system (EUREF FIN TM35FIN), and resolutions \((1000 \text{ m} \times 1000 \text{ m})\). Mean values of explanatory data per each municipality were calculated. The compiled dataset consisted of the presence-absence data of TBE and environmental, ecological, and anthropogenic data. Prior to modelling analysis, multicollinearity of the variables was tested using Variance Inflation Factors (VIFs) in R (Besley et al., 1980). The VIFs of the suite of environmental variables were calculated and correlated variables were excluded in a stepwise procedure at a generally accepted threshold value of 5 (Dormann et al., 2013; Ringle et al., 2015). Final explanatory variables used in the modelling consisted of 14 variables in the first dataset and 13 variables in the second dataset (Table 2.).

To model the habitat suitability of TBE in Finland, we used ensemble modelling in the biomod2 platform (version 3.3–7) in RStudio computing software (version 1.2.5033; RStudio Inc.). We fitted our data using the following eight predictive modelling techniques: generalized linear models (GLM), generalized additive models (GAM), classification tree analysis (CTA), artificial neural networks (ANNs), multivariate adaptive regression splines (MARS), generalized boosting models (GBM), random forest (RF), and maximum entropy (MAXENT). Flexible discriminant analysis (FDA) and surface range envelope (SRE) were excluded due to methodological weaknesses (Zhao and Gao, 2015; Elith et al., 2006; Pearson et al., 2006). Models were processed mainly using the default settings of the biomod2 with the following exception: we used the function bam mgcv package in GAM model in R due to its advantages regarding use of random term smoothers, specification of weights, offset, and ability to handle large datasets with good estimates (Wood, 2017).

We performed an iteration of 10 model runs for each of the eight algorithms, (80 model runs total). Model calibration was performed using randomly divided model training (70%) set, and model evaluation was performed using the remaining 30% of the data over the 10 model replicate runs (Guisan and Zimmermann, 2000). The area under the curve of a receiver operating characteristic (ROC) value (Fielding and Bell, 1997) and true skill statistics (TSS) (Peirce, 1884) were produced based on each model to estimate the predictive power of the
model. Ensemble model was generated using the best-performing model algorithms ($0.7 < \text{ROC} > 1.0$) (Drew et al., 2010). Sensitivity (the proportion of observed presences) and specificity (the proportion of observed absences) were calculated to quantify the omission errors (Fielding and Bell, 1997). The contribution of variables to the model was calculated by correlating the fitted values of the full models with those from the model in which the values of the predictor have been randomly permuted. The relative importance of variables was averaged over the eight model algorithms and compared with the variable contributions in the models to define the most powerful variables and their relative magnitude. All habitat suitability maps were created using ArcGIS software. Spatial autocorrelation (SAC) of the predictor variables was measured using Moran’s Index (Moran, 1950).

3. Results

3.1. Ensemble distribution model performance

The model performance of eight model algorithms with mean ROC and TSS values of 10 model runs in both datasets is presented in Fig. 4. Altogether, six out of the eight models provided reliable estimates for TBE distributions resulting in area under the ROC value of 0.74 at minimum (ROC = 0.74–0.91) in the first dataset. In the second dataset, seven out of the eight model algorithms resulted in an ROC value of 0.76 at minimum, suggesting good to excellent predictive power. RF and GBM models were the best performing model algorithms (ROC = 0.89–0.96) in both datasets.

Ensemble model performance in both datasets was classified as excellent, based on the ensemble median TSS value (0.988) in the first dataset and (0.956) in the second dataset. The ensemble model correctly predicted 98.8% of TBE presences (i.e. sensitivity) and 100% of TBE absences (i.e. specificity) in the first dataset. In the second dataset, sensitivity was 97.5% and specificity was 98.0%. SAC of the predictor variables for both datasets indicated that several covariates were highly spatially autocorrelated ($0.5 \leq \text{Moran’s I} \geq 0.9$) for short distances ($p < 0.05$) but not at longer distances. (Figs. A.1. and A.2.).

3.2. The relative importance of predictors

The relative contribution of influential variables (%) in both datasets showed that several factors resulted in values close to or greater than the mean importance value; these are thus important predictors (Fig. 5.). Proportion of built-up areas, forest, and people working in the primary sector, mean precipitation in April and July, and white-tailed deer and European hare density were the most important predictors (19–41%) in the first dataset. In the second dataset, the most important predictors (12–27%) were proportion of built-up areas, field, human population density, people working in the primary sector, snow-covered land in November, and raccoon dog density.

3.3. TBE risk based on suitability maps

The produced habitat suitability map of the first dataset suggests that the majority of coastal municipalities, municipalities in the Åland Islands, and several municipalities in Central and Southwestern Finland are suitable for TBE transmission (Fig. 6.). The habitat suitability map of TBE in the second dataset, strengthens the suggestion that coastal municipalities widely reflect the geographic distribution of the actual TBE risk and already notified TBE cases (Fig. 1.). Furthermore, the habitat suitability map from 2012 to 2017 period suggests that there is a considerable TBE transmission risk in several municipalities in Central Finland and in Eastern municipalities close to the Russian border. Inland municipalities in Northern Finland are estimated to have a lower risk of TBE transmission in both datasets.

3.4. Predicted current and future distribution of TBE based on the bioclimatic data

Our results showed an influence of a suite of environmental, anthropogenic, and ecological factors on probability of TBE disease case occurrence in Finland during two time periods. The habitat suitability maps indicated certain hotspot areas for TBE distribution. However, based on these results we cannot reliably estimate how the distribution of TBE may change in future climate in Finland. Therefore, we decided to use current and future global climate data including 19 bioclimatic variables obtained by WorldClim 2.0 and 1.4 (Table A.1.; Fick and...
Table 1
Description and source of all explanatory data.

| Data source                                      | Data layer(s)                                                                                                                                                                                                 | Spatial resolution |
|--------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------|
| Finnish Meteorological Institute (FMI)           | Mean precipitation (mm) and mean temperature (°C) in April, May, June, July, August and October in 2007–2011 and 2012–2017 were calculated based on daily means inputs in ArcGIS. Daily relative humidity and daily global radiation were obtained from FMI. Temperature and relative humidity measurement height was 2 m (FMI). | 1000 m/10,000 m    |
| Natural Resources Institute Finland (LUKE)       | Mountain hare, European hare, white-tailed deer, roe deer, moose, fox, martens, mink, fowl, raccoon dog, waterfowl and other wildlife density data mainland Finland. Average hunting data in 2007–2011 and 2012–2017 were calculated using the hunting data of each game management area (GMA). To obtain the average hunting data for each municipality, we used the value of game management area to which the municipality belonged. Game animal density data were generated for both study periods TBE 2007–2011 and TBE 2012–2017 based on annual hunting data. | Game management area |
| NASA Earthdata                                  | Derived from MODIS satellite imagery. Mean land surface temperature (°C) for future climate conditions and from WorldClim version 1.4 (Hijmans et al., 2005) for future climate conditions. Climate data were extracted to correspond the two datasets in ArcGIS. | 5600 m/1000 m/5600 m |
| Finnish Institute for Health and Welfare (THL)   | Patient TBE case data obtained from the National Infectious Diseases Register (NIDR) maintained by THL, and the locality of TBEV transmission based on patient interviews. | Municipality        |
| National Land Survey of Finland                 | Slope, flow direction, flow accumulation and elevation were derived from 25 m digital elevation model (DEM) (National Land Survey of Finland (NLS of Finland, 2000). Topographic wetness index (TWI) by Beven and Kirkby, 1979, was created based on slope and flow accumulation in ArcGIS. This was calculated by the following equation: TWI = ln (a/tan β), where a is the specific catchment area (SCA), the local upslope area draining through a certain point per unit contour length, and β is the local slope. | 25 m               |
| Statistics of Finland                           | People working in primary production (%) created by using the data of people working in different sectors compared to all labor. Human population density (people/km²) created by using population data per municipality from 2016 and calculating human population density per each municipality in ArcGIS. | Municipality        |
| Finnish Environment Institute (SYKE), European Environment Agency (EEA) | Proportion (%) of field, water, wetlands, forest and built-up areas were derived from CORINE land cover 2012, which was produced by the Finnish Environment Institute (SYKE), based on automated interpretation of satellite images and data integration with existing digital map data (European Environment Agency, EEA). Proportion (%) of field, water, wetlands, forest and built-up areas were obtained by calculating percentage of each land cover type for each municipality. | 20 m               |
| University of Turku                             | Tick species distribution data included nearly 20,000 collected ticks identified by species based on nationwide crowd sourcing study in 2015 (Laaksonen et al., 2017). | Location (Longitude, Latitude) |
| WorldClim-Global climate data                   | Annual Mean Temperature, Mean Diurnal Range (Mean of monthly (max temp - min temp)), Isothermality (BIO2/BIO7) (* 100), Temperature Seasonality (standard deviation *100), Max Temperature of Warmest Month, Min Temperature of Coldest Month, Temperature Annual Range (BIO5-BIO6), Mean Temperature of Wettest Quarter, Mean Temperature of Driest Quarter, Mean Temperature of Warmest Quarter, Mean Temperature of Coldest Quarter, Annual Precipitation, Precipitation of Wettest Month, Precipitation of Driest Month, Precipitation Seasonality (Coefficient of Variation), Precipitation of Wettest Quarter, Precipitation of Driest Quarter, Precipitation of Warmest Quarter and Precipitation of Coldest Quarter obtained from WorldClim version 2.0 (Fick and Hijmans, 2017) for current climate conditions and from WorldClim version 1.4 (Hijmans et al., 2005) for future climate conditions. Climate data were extracted to correspond the two datasets in ArcGIS. | –1000 m |
| Government of The Åland Islands                 | Hunting data of white-tailed deer, roe deer and moose per each municipality. Average hunting data of each municipality in 2007–2011 and 2012–2017 was calculated in ArcGIS and used as game animal density data. | Municipality        |

Table 2
Final environmental and other TBE predictor data used in the modelling analysis of TBE 2007–2011 and TBE 2012–2017 with value ranges. GMA = game management area.

| Dataset of 2007–2011 | Min. | Max. | Avg. | Dataset of 2012–2017 | Min. | Max. | Avg. |
|----------------------|------|------|------|----------------------|------|------|------|
| European hare density (hunted animals/GMA) | 0    | 15,840 | 5205 | Fox density (hunted animals/GMA) | 0    | 9366 | 3791 |
| White-tailed deer density (hunted animals/GMA) | 0    | 7957  | 3102 | Raccoon dog density (hunted animals/GMA) | 0    | 28,800 | 14,338 |
| People working in primary sector (%) | 0.10 | 22.90 | 6.29 | People working in primary sector (%) | 0.10 | 29.70 | 6.02 |
| Built-up areas (%) | 0.09 | 38.55 | 3.75 | Built-up areas (%) | 0.21 | 38.55 | 4.52 |
| Field (%) | 0.37 | 28.55 | 5.72 | Field (%) | 0.20 | 34.18 | 6.33 |
| Forest (%) | 2.31 | 83.63 | 32.01 | Wetlands (%) | 0.14 | 19.91 | 2.20 |
| Mean precipitation in April (mm) | 21.43 | 34.82 | 26.16 | Human population density (persons/km²) | 1    | 2863 | 152 |
| Mean precipitation in May (mm) | 33.00 | 55.94 | 41.09 | Snow-covered land in November (%) | 0.00 | 92.29 | 63.67 |
| Mean precipitation in July (mm) | 42.44 | 96.39 | 68.56 | Snow-covered land in December (%) | 1.84 | 100.00 | 37.70 |
| Mean precipitation in August (mm) | 69.14 | 100.54 | 79.14 | Mean precipitation in August (mm) | 58.42 | 99.34 | 75.60 |
| Mean precipitation in September (mm) | 50.51 | 81.78 | 65.40 | Mean precipitation in September (mm) | 42.22 | 76.70 | 57.41 |
| Mean precipitation in October (mm) | 55.03 | 86.59 | 70.44 | Mean precipitation in October (mm) | 49.18 | 87.97 | 64.18 |
| Normalized Difference Vegetation Index (NDVI) | 0.59 | 0.81 | 0.76 | Topographic wetness index (TWI) | 9.68 | 11.90 | 10.78 |
| Topographic wetness index (TWI) | 9.89 | 11.59 | 10.80 | |

5
Hijmans, 2017; Hijmans et al., 2005), to explore potential changes in the distribution of TBE in the future within the study area of the second dataset. In this substudy, we ignored all the other variables included in the major study as future prediction data were not available for them.

For future climate scenarios, for reasons of feasibility only one global climate model (GCM) called IPSL-CM5 (IPSL Earth system model for the 5th Intergovernmental Panel on Climate Change (IPCC) Assessment (2014)) was used (Dufresne et al., 2013). For greenhouse gas scenarios, we chose representative concentration pathways (RCPs) of both medium (RCP 4.5) and high (RCP 8.5) change for the years 2041–2060 and 2061–2080 to reveal possible climate change influences in two different time periods. Temperature in RCP 4.5 scenario ranges from 2.0° to 4.5° and in RCP 8.5 scenario from 3.5° to 4.5° (van Vuuren et al., 2011; Nazarenko et al., 2015). After multicollinearity of bioclimatic variables was run with VIF, six variables for modelling analyses were included (Table A.2). Modelling analyses were performed by using the same workflow as described in Section 2.3. The performance of eight models was similar within all climate conditions (Table A.3). RF and GBM performed best in current and future climate conditions with medium and high change scenarios and had predictive power from good to excellent (ROC = 0.76–0.94). A composition of influential factors in the second dataset is presented in Fig. 7.A–E.

The habitat suitability map of TBE in current climate conditions indicates a high TBE risk in the Åland Islands, in Southern and Southeastern Finland, and in Ostrobothnia and Northern municipalities (Fig. 8.). In ensemble forecasts for 2041–2060 and 2061–2080 with medium and high change scenarios, there is a slightly wider geographical extent of TBEV transmission risk in Northern Finland in several inland municipalities compared to current climate conditions. Southern Finland is predicted to be at high risk for TBE in all climatic conditions. No increase in TBE risk is predicted for inland municipalities in Central and Eastern Finland.

**Fig. 4.** Model performance comparison of eight model algorithms for TBE 2007–2011 (A) and TBE 2012–2017 (B) by area under the receiver operating characteristic curve (ROC) and true skill statistic (TSS) values of 10 model runs. The points represent the mean estimates and the solid lines represent the 95% confidence intervals. ANN = artificial neural networks; CTA = classification tree analysis; GAM = generalized additive models; GBM = generalized boosted models; GLM = generalized linear models; MARS = multivariate additive regression splines; MAXENT.Phillips = maximum entropy models; RF = random forest model.
Fig. 5. Mean relative importance of the explanatory variables (%) used to predict the distribution of TBE during 2007–2011 (A) and 2012–2017 (B). Bars represent the mean value of relative importance obtained from eight modelling algorithms with corresponding standard deviation values. The last bar on the right represents the mean relative importance value across the predictors. GMA = game management area.

Fig. 6. Mean predicted habitat suitability of TBE in the two datasets in Finland. The probability of TBE in both datasets was highest in coastal municipalities in Southern, Western and Northern Finland and several municipalities in Central and Eastern Finland.
Fig. 7. Relative contribution of influential variables (%) predicting the distribution of human TBE disease cases in current climate conditions (A), in 2041–2060 climate conditions with representative concentration pathways (RCP) 4.5 (B), in 2041–2060 climate conditions with RCP 8.5 (C), in 2061–2080 climate conditions with RCP 4.5 (D), and in 2061–2080 climate conditions with RCP 8.5 (E).

Fig. 8. Mean predicted habitat suitability of TBE in current and future climate scenarios with representative concentration pathways (RCP) 4.5 and RCP 8.5 based on six bioclimatic variables.
4. Discussion

4.1. Validity of the study

In this study, we used the biomod2 ensemble platform in R to create ensemble models to identify areas with suitable habitat conditions for TBE under present-day and future climate in Finland. Ensemble forecasting yields more accurate estimates over single-model estimates (Aratjo and New, 2007) and is commonly used to estimate the potential distributions of species and vector-borne diseases under current and future climate conditions (Deka and Morshed, 2018; Eneanya et al., 2018). Two different methods of accuracy, ROC and TSS were used in order to give more comprehensive view of the model performance. The ROC uses all possible thresholds for classifying the scores into confusion matrices and obtains each matrix’ sensitivity and specificity; comparing sensitivity against the corresponding proportion of false positives (Fielding and Bell, 1997). The TSS is independent of prevalence and an intuitive method of performance measurement of SDMs when predictions are expressed as presence-absence maps (Allouche et al., 2006). Even though ROC is widely used in modelling species distributions, TSS might be more realistic and practical method (Shabani et al., 2018). The majority of the models used herein produced moderate to high ROC and TSS values in both datasets indicating reliable estimations. ROC scores range from 0 to 1, where models higher than 0.5 predict better than random draws, and TSS values range from –1 to 1, where models higher than zero indicate a performance better than random (Ruete and Leynaud, 2015). In this study, RF and GBM models had the highest predictive power of all models and MAXENT the lowest (Fig. 4.). GBM and RF have been widely used to predict disease distributions (Eneanya et al., 2018; Elith et al., 2008; Breiman, 2001a) and the use of them was found to lead to tick mortality and reduced late-season populations for Lyme scapularis and I. ricinus ticks (Berger et al., 2014; Perret et al., 2000). July is typically seen as a “hiatus” month in the activity of I. persulcatus adults (Gray et al., 2016; Kørenberg, 2000), and I. ricinus nymphs and adults, when ticks cease questing and withdraw to humid microhabitats, either to escape unfavorable ambient conditions or due to diapause (Cayol et al., 2017; Sormunen et al., in press).

Higher precipitation in July might not only allow for continued tick questing during this period, but also ensure that ticks in shelters survive to later activity periods in August–September (Sormunen et al., 2016). Snow cover acts as an insulating blanket over ground litter and tends to further insulate ticks from the frigid winter air temperatures (Lundkvist et al., 2011; Volland et al., 2017). The proportion of snow-covered land in November may positively affect tick activity and survival because snow cover protects ticks from exposure to freezing and frequent temperature shifts when the air temperature decreases (Daust et al., 2016).

Medium-sized and large animals such as deer, hares and raccoon dogs are potential hosts for I. ricinus nymphs and adults, and I. persulcatus adults (Gray et al., 2016; Klemola et al., 2019). The population sizes of white-tailed deer and raccoon dogs have significantly increased in Finland recently (Natural Resources Institute of Finland, 2018) and may possibly be connected with higher I. persulcatus and I. ricinus density. White-tailed deer density and hare density have earlier been confirmed to correlate with I. ricinus abundance and consequently TBE distribution (Brugger et al., 2017; Jaenson et al., 2018). The proportion of field and forest area in municipalities have indirect effects on TBE transmission. Forests are typical habitats for many important tick host animals, such as deer and hares, and higher amounts of forests therefore typically increase host animal and consequently tick abundance. On the other hand, the increasing proportion of field area often means that a more fragmented habitat mosaic is formed, wherein the amount of boundary areas between different habitats increases. These increasing edge effects allow for greater biodiversity and often higher animal densities (Tack et al., 2012; Czupryna et al., 2016; Nadolny and Gaff, 2018). High human population density and a high proportion of built-up areas are associated with the large number of population and naturally increase TBE risk. People who are working in the primary sector spend relatively more time outdoors than other sectors and consequently have a higher risk for getting tick bites (Randolph et al., 2008).

In our future forecasts, mean temperature of the warmest month
and wettest quarter and temperature seasonality were the most influential bioclimatic factors on TBE distribution. High spring and summer temperatures and mild winter temperatures are drivers of new tick establishment and higher TBE risk at high-latitudes in Northern Europe (Gray et al., 2009; Randolph and Rogers, 2000).

5. Conclusions

Our results confirm the influence of game animal densities and anthropogenic and environmental factors on distribution of human TBE disease cases in Finland. The proportion of built-up areas, field, forest, and snow-covered land, people working in the primary sector, human population density, mean precipitation in April and July, white-tailed deer density, and European hare and raccoon dog density were found correlated with the occurrence of TBE. In future forecasts, temperature variables were the most influential drivers for higher TBE risk. Based on habitat suitability maps of 2007–2011 and 2012–2017, high-risk areas of TBE transmission were estimated to be in the Åland Islands, the coastal regions of Southern, Western, and Northern Finland, and several municipalities in Central and Eastern Finland. In future forecasts for 2041–2060 and 2061–2080 climate, a slightly wider geographical extent of TBE risk was introduced in the Åland Islands and Southern, Western and Northern Finland, even though the risk itself was not significantly increased (Fig. 7). Identified risk areas were consistent with previous study results, in which Southern Finland (including the Åland Islands) and Southeastern Finland were estimated to be suitable for TBE transmission in the 2020 forecast, and risk areas were suggested to expand up to Central Finland in the 2080 forecast (Randolph and Rogers, 2000). Higher TBE risk in Northern regions will be reasonable as temperature increase is greater in the northermost latitudes (Trenberth et al., 2007), which makes the region more favorable for tick activity (Soucy et al., 2018; Medlock et al., 2013; Lindgren and Gustafson, 2001; Gray et al., 2009). Our findings provide insight into the identification of risk areas and influential factors for TBE in Finland and can be applied to other regions located at high latitudes in the Northern Hemisphere. Our goal in future studies is to combine larger and more detailed datasets of human TBE cases from Scandinavia and create predictions across Northern Europe.

CRediT authorship contribution statement

Ruut Uusitalo: Conceptualization, Data curation, Formal analysis, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. Mikael Siljander: Methodology, Project administration, Supervision, Visualization, Writing - review & editing. Timothée Dub: Writing - review & editing. Jani J. Sormunen: Writing - review & editing. Petri Pellikka: Funding acquisition, Resources. Olli Vapalahti: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors have declared that no competing interests exist.

Acknowledgements

We are grateful for many researchers and institutions who provided data for this study. We thank Natural Resources Institute Finland and the Government of the Åland Islands (Ålands landskapsregering) for providing hunting data of game animals, and Niko Kulha from the University of Helsinki for delivering the presence data of tick species. This study was funded by the Tiina and Antti Herlin Foundation, the Jane and Aatos Erkko Foundation, and the Doctoral Programme in Interdisciplinary Environmental Sciences (DENVI) of the University of Helsinki. The authors would also like to acknowledge the funding from the Academy of Finland for the VECLIMIT project and the Ministry for Foreign Affairs of Finland for the TAITAGIS project supported by Board of Education of Finland (http://cimo.innofactor.com/programmes/hei_index/index/programmes/hei_index/projects/taitagis).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ttbdis.2020.101457.

References

Alekseev, A.N., Chunikhin, S.P., 1990. The exchange of the tick-borne encephalitis virus between ixodid ticks feeding jointly on animals with a subthreshold level of viremia. Med. Parazitol. (Mosk.) 48-50.
Aliksiie, A.A., Peterson, A.T., Samy, A.M., 2017. Climate change influences on the potential geographic distribution of the disease vector tick Ixodes ricinus. PLoS One 12, e0189092.
Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43 (1223-), 1222.
Aranjou, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22 (2), 42–47.
Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol. Evol. 3, 327–338.
Beaute, J., Spiteri, G., Warns-Petit, E., Zeller, H., 2018. Tick-borne encephalitis in Europe, 2012 to 2016. Euro Surveill. 23 (45).
Berger, K., Ginsberg, H., Dugas, K., Hamel, L., Mather, T., 2014. Adverse moisture events predict seasonal abundance of Lyme disease vector ticks (Ixodes scapularis). Parasit. Vectors 7, 181.
Besley, D., Kuh, E., Welsch, R., 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Wiley, New York.
Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology/1. A physically based model of soil water flow. Hydrolog. Sci. J. 24 (1), 43-69.
Bhatt, S., Gething, P.W., Brady, O.J., Fanello, P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., 2013. The global distribution and burden of dengue. Nature 496, 604.
Brabec, M., Daniel, M., Malý, M., Danielová, V., Krň, B., Kott, L., Beneš, Č., 2017. Analysis of meteorological effects on the incidence of tick-borne encephalitis in the Czech Republic over a thirty-year period. Virol. Res. Rev. 1, 2-8.
Breiman, L., 2001a. Random forests. Mach. Learn. 45, 5–32.
Bruggner, K., Walter, M., Chitimia-Dobler, D., Dobler, G., Rubel, F., 2017. Seasonal cycles of the TBE and Lyme borreliosis vector Ixodes ricinus modelled with time-lagged and interval-averaged prediction. Exp. Appl. Acarol. 73, 439–460.
Cattadori, I.M., Haydon, D.T., Thirgood, S.J., Hudson, P.J., 2003. Are indirect measures of the TBE and Lyme borreliosis vector Ixodes ricinus modelled with time-lagged and interval-averaged prediction. Exp. Appl. Acarol. 73, 439–460.
Cayol, C., Koskela, E., Mappes, T., Siukkola, A., Kallio, E.R., 2017. Temporal dynamics of Ixodes ricinus in northern Europe: epidemiological implications. Parasit. Vectors 10, 166.
Chalghaf, B., Chemkhi, J., Mayala, B., Harrabi, M., Benie, G.B., Michael, E., Samy, A.M., 2017. Climate change in the Republic over a thirty-year period. Virol. Res. Rev. 1, 2-8.
Charrel, R.N., Attoui, H., Butenko, A.M., Clegg, J.C., Deubel, V., Fialová, A., Heinz, F.X., Labuda, M., 2004. Tick-borne virus diseases of human interest in Europe. Clin. Microbiol. Infect. 10, 1040–1055.
Chunikhin, S.P., Alekseev, A.N., Alekseev, A.N., Chunikhin, S.P., 1990. The exchange of the tick-borne encephalitis virus between ixodid ticks feeding jointly on animals with a subthreshold level of viremia. Med. Parazitol. (Mosk.) 48-50.
Ciosek, E., Piekarski, J., 2015. Increased Borrelia burgdorferi infection in dogs from central Poland. Parasit. Vectors 8, 478.
Daniel, M., Danielová, V., Fialová, A., Malý, M., Šmídová, Z., 2014. Abiotic and bioclimatic predictors and tick-borne disease incidence. In: Ecology and Sources of Collinearity. Wiley, New York.
Deribe, K., Cano, J., Njouendou, A.J., Eyong, M.E., Beng, A.A., Giorgi, E., Pigott, D.M., 2019. Predicted distribution and burden of dengue. Nature 461.
Besley, D., Kuh, E., Welsch, R., 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Wiley, New York.
Kull, E., Antti, Herlin Foundation, the Jane and Aatos Erkko Foundation, and the Doctoral Programme in Interdisciplinary Environmental Sciences (DENVI) of the University of Helsinki for delivering the presence data of tick species. This study was funded by the Tiina and Antti Herlin Foundation, the Jane and Aatos Erkko Foundation, and the Doctoral Programme in Interdisciplinary Environmental Sciences (DENVI) of the University of Helsinki. The authors would also like to acknowledge the funding from the Academy of Finland for the VECLIMIT project and the Ministry for Foreign Affairs of Finland for the TAITAGIS project supported by Board of Education of Finland (http://cimo.innofactor.com/programmes/hei_index/index/programmes/hei_index/projects/taitagis).
Kulkarni, M.A., 2018. High-resolution ecological niche modeling of Ixodes scapularis ticks based on passive surveillance data at the Northern Frontier of Lyme disease emergence in North America. Vector-Borne Zoonotic Dis. 18, 235–242.

Stefano, P., Rubikowska, B., Bratkowski, J., Ustrnul, Z., Vanwanmeke, S., Rosinska, M., 2018. A predictive model has identified tick-borne encephalitis high-risk areas in regions where no cases were reported previously, Poland, 1999–2012. Int. J. Environ. Res. 15, 677.

Sun, R., Lai, S., Yang, Y., Li, X., Liu, K., Yao, H., Zhou, H., Li, Y., Wang, L., Mu, D., 2017. Mapping the distribution of tick-borne encephalitis in mainland China. Ticks Tick. Dis. 8, 631–639.

Süss, J., 2011. Tick-borne encephalitis 2010: epidemiology, risk areas, and virus strains in Europe and Asia—an overview. Ticks Tick. Dis. 2, 2–15.

Sutherst, R.W., Collyer, B.S., Yonow, T., 2000. The vulnerability of Australian horticulture to the Queensland fruit fly, Bactrocera (Dacus) tryoni, under climate change. Aust. J. Agric. Res. 51, 467–480.

Tack, W., Madder, M., Baeten, L., Vrenen, P., Verheyen, K., 2012. The abundance of Ixodes ricinus ticks depends on tree species composition and shrub cover. Parasitol. 139, 1273–1281.

Thuiller, W., Georges, D., Engler, R., Breiner, F., Georges, M.D., Thuiller, C.W., 2016. Package ‘biomod2’. Species Distribution Modeling Within an Ensemble Forecasting Framework. biomod2. https://CRAN.R-project.org/package=biomod2.

Tokarevich, N.K., Tronin, A.A., Blinova, O.V., Buzinov, R.V., Boltenkov, V.P., Yurasova, E.D., Nurse, J., 2011. The impact of climate change on the expansion of Ixodes persulcatus habitat and the incidence of tick-borne encephalitis in the north of European Russia. Glob. Health Action 4, 8448.

Tonteri, E., Kurkela, S., Timonen, S., Manni, T., Vuorinen, T., Kuusi, M., Vapalahti, O., 2015. Surveillance of endemic foci of tick-borne encephalitis in Finland 1995-2013. Euro Surveill. 20 (37), 30020.

Trenberth, K.E., Jones, P.D., Ambenje, P., Bejiariu, R., Easterling, D., Klein Tank, A., Parker, D., Rahimzadeh, F., Renwick, J.A., Rusticucci, M., 2007. Observations: surface and atmospheric climate change. Chapter 3. Clim. Change 235–336.

Tuomenvirta, H., Haavisto, R., Hildén, M., Lanki, T., Lohtala, S., Meriläinen, P., Mäkinen, K., Parjanne, A., Peltonen-Sainio, P., Pihl-Sivola, K., 2018. Säät-ja Ilmastoriskit Suomessa-kansallinen Arvio. https://julkaisut.valtioneuvosto.fi/handle/10024/161015.

Uusitalo, R., Siljander, M., Culverwell, C.L., Mutai, N.C., Forbes, K.M., Vapalahti, O., Pellikka, P.K.E., 2019. Predictive mapping of mosquito distribution based on environmental and anthropogenic factors in Taita Hills, Kenya. Int. J. Appl. Earth Obs. 76, 84–92.

Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hutt, G.C., Kram, T., Key, V., Lamerque, J., 2011. The representative concentration pathways: an overview. Clim. Change 109, 5.

Vollack, K., Souodhi, S., Névir, P., Müller, K., Richter, D., 2017. Influence of meteorological parameters during the preceding fall and winter on the questing activity of nymphal Ixodes ricinus ticks. Int. J. Biometeorol. 61, 1787–1795.

Wan, Z., Hook, S., Hulley, G., 2015. MOD11C3 MODIS/Terra Land Surface Temperature/Emissivity Monthly L3 Global 0.05 Deg CMG V006. NASA EOSDIS LP DAAC.

Wood, S.N., 2017. Generalized Additive Models: an Introduction With R. Chapman and Hall/CRC, Florida.

World Health Organization (WHO), 2019. Tick-borne Encephalitis. Accessed 27th June 2019. https://www.who.int/ith/diseases/tbe/en/.

Zhao, C., Gao, F., 2015. A nested-loop Fisher discriminant analysis algorithm. Chemometr. Intell. Lab. Syst. 146, 396–406.