The Rise of West Nile Virus in Europe: investigating the combined effects of climate, land use and economic changes

Matthew J. Watts[1,*], **Panagiota Kotsila**[1, 3], **P. Graham Mortyn**[1,4], **and Victor Sarto i Monteys**[1,2]

[1]Institute of Environmental Science and Technology (ICTA) , Autonomous University of Barcelona (UAB), Bellaterra, Spain
[2]Departament d'Agricultura, Ramaderia, Pesca, Alimentació i Medi Natural, Generalitat de Catalunya, Avinguda Meridiana, Barcelona, Spain
[3]Barcelona Laboratory for Urban Environmental Justice and Sustainability (BCNEJ), Institute of Environmental Science and Technology (ICTA) , Autonomous University of Barcelona (UAB), Bellaterra, Spain
[4]Department of Geography, Autonomous University of Barcelona (UAB), Bellaterra, Spain
[*]corresponding.author@matthewjohnwatts@googlemail.com

ABSTRACT

West Nile Virus (WNV) has recently emerged as a major public health concern in Europe, its recent expansion also coincided with some remarkable socio-economic and environmental changes, including an economic crisis and some of the warmest temperatures on record. Here we empirically investigate the drivers for this phenomenon at a European wide scale by constructing and analysing a unique spatial-temporal data-set, that includes data on climate, land-use, the economy, and government spending in key sectors such as health and the environment. Drivers and risk factors of WNV were identified by building a conceptual framework, and relationships were tested using a Generalized Additive Model (GAM), which could capture complex non-linear relationships and also account for the spatial and temporal auto-correlation. As expected, land use factors (higher percentage of wetlands and agricultural land) and climate factors (higher summer rainfall and higher summer temperatures) were positive predictors of WNV infections. However, we found that a mean winter temperature of between 2 °C and 6 °C was the best predictor of annual WNV infections, suggesting that successful overwintering of infected adult mosquitoes (likely *Culex pipiens*) was key to the intensity of outbreaks for a given year. This result, taken in the context of recent winter warming due to climate change, demonstrates how the virus has become so prevalent in Europe. Furthermore, lower GDP growth, lower spending growth on environmental factors such as agriculture, forest, fisheries, and waste water management were also strong predictors of WNV infections, suggesting that austerity and cuts to key sectors benefited vector species and the virus during this crucial period. Waste water management spending may be a reliable causative factor, due to, for instance, the neglect of hazard prevention efforts, such as spending on flood defences, essential works like sanitation and upkeep and of infrastructure, that can lead to the creation of mosquito breeding habitats, e.g. potholes, blocked drain. This study should act as a cautionary tale; central governments looking to embark on austerity measures should exercise extreme caution given the long term trade-offs and unintended consequences they can have on the environment and human health, especially when considering we are facing numerous threats brought about by global warming and other anthropogenic induced changes, that can benefit emerging diseases i.e. global trade in wild animals, intensive agriculture animal rearing and land use conversion.

1 Introduction

Over the past few decades, new health risks have been emerging in Europe, particularly with the recent appearance of neglected tropical diseases (NTDs) such as Chikungunya, West Nile Virus (WNV), Dengue (DENV-1) and Crimean-Congo haemorrhagic fever [1, 2, 3]. Rising temperatures are likely increasing the transmission potential of NTDs in Europe, by affecting the geographic spread, abundance, survival and feeding activity of vector species and benefiting pathogen development in infected vectors [4, 5, 6, 7, 8, 9]. This, combined with other factors such as human population growth, trade in wild animals and derived products, intensive animal rearing, global commerce, air travel, urbanization and land-use changes, is boosting the potential for new diseases to emerge in Europe [10, 11, 12, 13].

Although the study of physical factors has been key in understanding disease transmission and distribution, few articles have included an examination of anthropogenic factors, such as political decision making and policy making [14, 15], which can have wider and unintended effects on natural ecosystems and, eventually, on disease [16].

Furthermore, at the time of writing, government debt is soaring in Europe because of the Covid-19 crisis [17], and there is a danger that European countries will revert back to austerity measures such as those imposed during the previous economic crisis, which will weaken public health systems' ability to combat diseases. [14, 18, 19, 20, 21, 22].

In this study, we focus on WNV, a single-stranded RNA *Flavivirus* closely related to other *Flaviviridae* pathogens such as dengue, Japanese encephalitis and yellow fever viruses [23]. While the largest proportion of people infected with WNV will not develop any symptoms, around 20% of them will get West Nile fever, a flu like illness, or severe West Nile disease. The most severe manifestations of the disease include acute aseptic meningitis or encephalitis, anterior myelitis, hepatosplenomegaly, hepatitis, pancreatitis, and myocarditis. On rare occasions, infection can lead to Guillain–Barré syndrome and other demyelinating neuropathies [23, 24]. Although WNV is a zoonotic pathogen, infecting mammals, particularly humans and horses, the transmission cycle only involves mosquitoes and birds, as mammals are considered dead end hosts, i.e. they cannot transmit the disease back to mosquitoes [24]. In Europe, the main vectors are *Culex pipiens*, *Culex modestus*, and *Coquillettidia richiardii*, although *Aedes* species can also transmit the disease [25] (see figures S16-S18 in additional information for vector distribution maps). Vector abundance and diversity is modulated by physical environmental factors, such as temperature, rainfall, water resources [23, 25, 26]. As well as vectors, crucial for transmission of WNV is the presence of suitable hosts, such as terrestrial and wetland birds. These hosts can occupy a diverse range of areas; indeed, spring migration of birds from infected regions of Sub-Saharan Africa to temperate regions of Europe is considered to be one of the main drivers of the disease in Europe [25, 27]. Although WNV infections are more common of late, sporadic West Nile virus outbreaks have occurred in humans and equines in southern and eastern European countries over the last century; the majority of which have occurred in wetland areas and densely inhabited urban areas [25]. In 2010, a major outbreak of West Nile WNV hit a number of countries in the European region: Greece, Hungary, Romania, and Turkey. Since then, outbreaks have occurred annually in multiple regions, including more northerly regions that had not previously reported cases (in Germany, Slovakia, Slovenia, and the Czech Republic) (see fig 1). This culminated in another major outbreak in 2018 which hit almost all susceptible regions, that is, areas in which local transmission occurred at least once over the study period. We should also bear in mind that the actual number of cases in Europe is likely to be much higher than reported, since most WNV infections in humans are asymptomatic, and we can therefore assume that only serious cases are reported.

To explain the recent rise of WNV infections in Europe in more depth, we empirically investigate the combined impact of three sets of factors: (1) Climate-related factors, including temperature, rainfall and surface water; (2) Geographic factors which include continuous/discontinuous urban fabric, and regional coverage of wetlands and arable land; (3) Economic and socio-economic factors which capture the impact of the economic crisis and are proxied by GDP growth, spending on agriculture, forestry and fisheries, waste water management, and health (for a more detailed explanation of how all factors interact with WNV, see the conceptual framework).

Our analysis focuses on regions in the seven European countries hardest hit by WNV outbreaks - Austria, Bulgaria, Croatia, Greece, Hungary Italy, and Romania, our time series data set captures the time period before and after the economic crisis.

2 Results

To carry out our empirical analysis, we compiled a spatial temporal data-set that captures the main drivers and risk factors of WNV infections in Europe (see Table 1 and supplementary information - conceptual framework).

Figure 1 characterizes the climate in the study regions. As we can observe, WNV infections occurred in regions with climates that can be described as 'Hot-summer Mediterranean,' 'Humid subtropical,' Temperature oceanic' and ' Warm-summer humid continental' or 'Temperate oceanic.' Figure 2 shows the WNV infection incidence rates over the study period. From 2007 to 2009, very few regions were affected by WNV, however, 2010 saw an outbreak that spread far and wide. Since 2010, the number of regions that have been affected by WNV increased. In 2018, a massive outbreak affected almost all of the regions in our study.

The relationship between the incidence of WNV infections and the climate, land-use, and economic factors was modelled via a Generalized Additive Model (GAM), which also accounted for the spatial and temporal auto-correlation. Table 2 show the results of our statistical analysis. We built the statistical model in a stepwise fashion using the lowest Akaike Information Criterion (AIC) to help us assess the different specifications. We selected relevant variables in each specification according to their category, i.e. climate, land-use and economic. We included all variables in our final specification to ascertain the contribution of each driver, all else equal. Table 2 also summarises the relevant statistics (AIC, Deviance, Adjusted R squared and so on) to compare the different specifications. We find that our final model (Full model) has the best fit in terms of the AIC, followed by the economic model, the climate model and land use model, as shown in Table 2 (for model diagnostics see additional information figures

S8-15 and Table S1). Note that as we are not estimating a standard regression model, the figures reported should not be read as coefficients, but degrees of freedom of the smooth terms. Given that we cannot interpret the coefficients to infer the sign and magnitude of the relationship, we visualise it by plot. Figures (3-5) plot the partial effects—the relationship between a change in each of the covariates and a change in the fitted values in the full model. If we don't control for other factors, the climate variables are highly statistically significant. Mean summer temperature (°C) is associated with an increase in the incidence of human annual WNV infections; WNV infections are also associated with the amount of rain days per summer (Jun-Aug). Both variables have a linear relationship with the outcome variable. Mean winter temperature has a quadratic relationship with WNV. Number of days of rain in the winter is also positively correlated with annual WNV infections. The impact of the climate variables remains highly statistically significant in specification 4 (full model), with the exception of mean summer temperature, whose impact becomes substantially weaker. By plotting the partial effects, we can also see that the WNV cases do not occur in regions when average summer temperatures are below 22°C, while they steadily rise in temperatures above this point (figure 3, top left). In addition, winter temperatures below 2 °C reduce WNV infections and so do temperatures above 6°C. We also examined how the extent of regional surface water in the summer affect the incidence of the infection. As the partial effects plotted in figure 4 (top left) show, the relationship is linear and negatively correlated with WNV infections. As expected, the percentage of wetlands and arable land in a region (Land-use model) are positively correlated with the incidence of WNV and highly significant. However, the percentage of discontinuous urban fabric, which represents built up areas (residential suburbs, villages) only becomes statistically significant in our final specification, which is positively associated with WNV infections. The economic indicators are also highly significant and have very strong negative relationships with WNV infections. These figures represent a 2007 baseline GDP / government spending growth for each country. Generally, regions in counties which tended to have lower annual GDP growth and lower central government spending on areas of the environment, such as agriculture, forest fisheries spending and waste water management, have much higher WNV infection rates. On the contrary, central government spending on health is positively correlated with WNV infections, a finding which we will expand on in the next section.

3 Discussion

A range of regional environmental risk factors associated with human WNV infections, along with a set of economic indicators capturing the impact of the economic crisis, was used to explain the rise in WNV outbreaks in Europe over the last 13 years (2007-2019). This allowed for the identification of changes, in both environmental and economic factors that possibly led to the recent rise in infections. This is the first study we know of that considers this phenomenon in such detail, at a European wide scale.

Over the past 70 years, the countries analysed in this study have been experiencing increasingly warmer temperatures throughout the year (see figs 6-9). According to our initial analysis, the last decade has been the warmest. The results of our final model show that average summer temperatures above 22 °C are positively associated with an increase in WNV infections. It is likely that rising temperatures have created more favourable conditions for the virus and this is why it is occurring in more northerly regions than in previous decades. Higher temperatures, especially in summer months, is a key driver of WNV transmission, since warmer temperatures are known to influence the hatching rate and development time of mosquitoes, and shorten the extrinsic incubation period (EIP) of WNV and related viruses [28, 29, 30, 31, 32, 33]. Our analysis of the mean winter temperature (Dec-Feb) reveals a quadratic relationship with WNV and that it is the strongest predictor of annual WNV infections. Temperatures below 2 °C and above 6 °C have a negative impact on WNV infections. Our results are consistent with findings by [34]. The authors looked at *Culex pipiens* overwintering potential in mosquitoes in sheds and houses. Temperatures in sheds ranged from 2.7–17.0°C, and in houses 8.8–23.7°C. Survival rates were much better in sheds that had cooler temperatures, and also more humidity, which is also consistent with our results. Diapausing *Culex pipiens* mosquitoes do not necessarily do better under warmer conditions and there is a temperature range in which they can successfully diapause. Although our results may be slightly less accurate than results derived through an experimental study (since we analysed infection data at a regional scale using mean monthly temperatures), as with the study by Koenraadt et al. [34], the results do adhere to the same quadratic relationship. WNV infections for the months following winter are higher, if the specific temperature range of between 2 °C and to 6 °C is met. This is likely because WNV can overwinter with the mosquitoes when conditions are right, and then can be transmitted early on in the year, and throughout the year. This result, taken in the context of recent winter warming due to climate change, demonstrates how the virus has become so prevalent in Europe. Given current trends, we can also expect to see regions that have previously been too cold for *Culex pipiens* to survive over winter become viable locations and cause further havoc in regions that are currently experiencing just a few annual cases. On the contrary,

regions which currently have optimal conditions for overwintering mosquitoes may become too warm. It is also important to note that it is not currently clear at what temperatures *Culex modestus* and *Coquillettidia richiardii* overwinter as adults, if at all.

The number of rain days per summer is also positively correlated with WNV infections. This result is consistent with the literature, i.e. a steady flow of aquatic resources for mosquitoes has a positive impact on their abundance, and therefore an increase in disease transmission; [35, 36]. Rainfall patterns have also been shifting since the 1950s (see figs 6-9) although unlike climate, there is no a clear trend and results are more difficult to interpret. In general, Austria, Croatia and Italy are seeing less intense rainfall in the summer months, but higher in Autumn, whereas Bulgaria, Greece, Hungary and Romania are receiving more rainfall in summer months. Additional figures (S4-S7) also capture this information in a long term-time series. Our results also show that higher summer surface water extent for a given year in a given region, is negatively correlated with WNV. This was not an expected finding, since we would have expected higher levels of surface water would mean more breeding grounds for mosquitoes. However, the result is consistent with part of the literature as sometimes desiccation of water resources can bring mosquito and bird hosts closer together and can increase transmission potential [23, 25]. These results may also suggest that regions with higher surface water extent for a given year may be experiencing flooding and fast water movement, which may wash away mosquito eggs and larvae, and also may inhibit contact between birds, mosquitoes and humans [23, 25, 37]. It is important to note that this variable probably does not capture the creation of short-term water resources created by rainfall (e.g. pools, puddles), which can be used as breeding habitat by mosquitoes. It rather captures long term and large water surface such as deltas, lakes and flood plains.

As for the land-use variables, as expected, regions with a larger proportion of arable land (particularly rice paddies and irrigated agriculture) are more likely to see outbreaks of WNV. This is consistent with the literature as such land-use patterns attract bird and mosquito species capable of transmitting WNV, that can then have a knock-on effect on human populations. The percentage of discontinuous urban fabric, which represents populated area of low to medium density such as residential suburbs, villages etc. [see 38] is positively associated with WNV, also consistent with the literature. However, this variable only becomes weakly statistically significant in our final specification. This could be interpreted as locations with such land-use patterns are particularly at risk when climate conditions permit and/or during periods of economic decline. According to our additional descriptive analysis (figures S2-S3), we are seeing increases in land use types associated with human WNV infections, which may have also contributed to the rise in cases over the past decade. In terms of economic factors influencing WNV, in general, regions in countries which were experiencing lower GDP growth, lower spending growth on environmental factors such as agriculture, forest, fisheries, and waste water management had a higher number of cases over the study period. Waste water management spending may be a reliable causative factor, due to, for instance, the neglect of hazard prevention efforts, such as spending on flood defences, essential works like sanitation and upkeep of infrastructure, that can lead to the creation of mosquito breeding habitats, e.g. potholes, blocked drains [23, 25, 39]. We find that government spending on the agriculture, forest + fisheries sectors, which represents investment or subsidies that can go into improving environmental care, landscapes and biodiversity in rural locations, as well as investments in agricultural technology that improve worker safety or mechanise agriculture, had negative effects on WNV infections. A drop in government spending on these three economic sectors can be considered as a proxy for austerity. Our results suggest that economic problems and spending reductions in key sectors have clearly had knock on effects. On the contrary, our statistical analysis shows that central government spending on health had a positive, albeit concave, impact on WNV infections, which is consistent with the literature, since better funded health systems tend to do better at disease diagnosis and detection, but clearly once those are improved, more health spending does not lead to more cases, hence the concavity [40, 41, 42, 43]. Additional figure S1 provides a multiple time-series of the economic situation in the 7 countries in terms of GDP and government spending.

Although not all the variables used in this study represent disease transmission mechanisms directly, understanding the relative impact of the environmental, climate and economic factors on disease outcomes can help risk assessors predict where diseases are likely to occur in the future, by identifying locations with vulnerabilities in public health systems and/or by identifying impoverished areas that tend to be susceptible to disease. One important factor we did not analyse in this study, which may have also contributed to the upsurge in infections over the past 10 years, is the prevalence of WNV in migratory birds. It may be the case that infection levels in birds in endemic regions are now much higher, which has increased the chance of wider outbreaks in Europe. It is important to note that, with any analysis dealing with regional data, results should be taken with caution because of issues of scale and uncertainty introduced by the aggregation procedure; therefore, our findings should be evaluated with results from more localised studies. Nevertheless, this study should act as a cautionary tale; central governments looking to embark on austerity measures should exercise extreme caution given the long term trade-offs and unintended consequences they can have

on the environment and human health, especially when considering we are facing multiple threats brought about by global warming and other anthropogenic induced changes that can benefit emerging diseases, i.e. global trade in wild animals, intensive agriculture / animal rearing and land use conversion. It is vital for governments to invest in environmental infrastructure and health services to prevent future epidemics, rather than helping to instigate them.

4 Data collection and methods

In this study, we compiled a spatial temporal data-set that captures the main drivers and risk factors of WNV infections in Europe. We compiled the data-set based on findings from a literature review (see supplementary information - conceptual framework).

4.1 Aggregation

All data were aggregated annually to produce the final yearly panel data-set and aggregated at the NUTS 3 country subdivision level [see 44], apart from economic data which was sourced at the country level (central government).

4.2 WNF Case Data

WNV case data were provided at request by the European Centre for Disease Prevention and Control (www.ecdc.europa.eu). Cases data are collected weekly by EU member states and affiliates. Data were aggregated at NUTS 3 country subdivisions [44]. Positive cases were confirmed using WNV-specific IgM in blood. All cases were aggregated yearly to create the annual panel data set.

4.3 Economic, Socio-Economic and Demographic Data

Economic data were extracted from the Eurostat database (https://ec.europa.eu/eurostat/data/database), which provides comparable statistics and indicators and is presented in yearly time series. To capture factors determining the economic crisis, austerity and cuts to public spending we selected country level "Gross domestic product", "Agriculture, forestry", fisheries spending",”Waste water management spending" , and "Health spending". The "Agriculture, forestry, fisheries spending" variable captures spending in rural areas on factors that would help to improve the environment and agricultural development which would benefit agricultural workers and/or mechanise production [see 45]. In order to represent spending before and after the economic crisis, we created a baseline index for each variable set at 2007 levels, which represented negative or positive growth from the point just before the economic crisis hit Europe.

Population count data to predict the number of persons at risk in a region were sourced from the Socio-economic Data and Applications Center's Gridded Population of the World data set [46]. This data-set estimates population count for the years 2000, 2005, 2010, 2015, and 2020, consistent with national censuses and population registers. Data were extracted from areas where vector presence was predicted. R's Zoo package was used to replace values for missing years, by implementing a linear interpolation method that would predict trends between years. This way, increases or decreases in human population were controlled for in the final model.

4.4 Climate Data

Climate data were sourced from the E-OBS gridded data-set [see 47]. The data-set was created using a series of daily climate observations at meteorological stations throughout Europe and the Mediterranean. Files were processed in R with the Tidyverse, NetCDF and Raster packages to create regional seasonal variables: "mean temp winter (°C)", "mean temp spring (°C)", "mean temp summer (°C)", "winter precipitation (total mm per region),"spring precipitation (total mm per region",”summer precipitation (total mm per region)". Although, to standardize the data across regions we created variables which counted rain occurrence per season i.e. ”days of rain in winter",”days of rain in spring" and "days of rain in summer". Winter was designated as December to March, Spring as March to June, and summer June to September. We did not include "mean temp spring (°C)" in our final data set as it was correlated with winter and summer temperature variables, we concluded that we would capture more variation using the winter and summer variables which were not highly correlated.

4.5 Land Use

Land use statistics were captured at NUTS3 level using the CORINE Land Cover (CLC) 2006, 2012 and 2018 databases. The database provides information on the biophysical characteristics of the Earth's surface [see 48]. We calculated the percentage cover of each land use type to give us an indication of the land-use per region. R's Zoo package was used to replace values for missing years, by implementing a linear interpolation method that would predict trends between years, apart from 2019 where 2018 values were used. This way increases or decreases in land

use changes were controlled for in the final models. To incorporate some of the main land-use WNV risk factors into our model, we selected variables: "continuous urban fabric", "discontinuous urban fabric", "wetlands (fresh water)" and "arable land".

4.6 Surface Water

Regional surface water data was sourced using the JRC Monthly Water History, v1.2 data set (https://developers.google.com/earth-engine/datasets/catalog/JRC_GSW1_2_MonthlyHistory) via Google Earth Engine using a 30 metre pixel resolution. This data set contains maps of the location and temporal distribution of surface water from 1984 to 2019 and provides statistics on the extent and change of those water surfaces. Data were generated using scenes from Landsat 5, 7, and 8. Each pixel was individually classified into water / non-water using an expert system and the results were collated into a monthly history [see 49] . In order to standardize this data across regions, we grouped surface water observations by region and season. We then converted the surface water observations to Z-scores to give us the standard deviation of the water extent. This would help us determine if the water extent was average, below the mean (low), or above the mean (high).

4.7 Statistical Methods

4.7.1 General additive regression model to assess impact of independent variables on WNV case data at regional level

One of the main issues with our data-set is that it did not meet some basic assumptions for statistical inference, and specifically the data are not independent and identically distributed random variables (iid). More specifically, the data-set captured repeated measurements over the same regions, and observations were not independent because of spill over effects from neighbouring regions, therefore we needed to implement an appropriate statistical design to control for both temporal and spatial pseudo replication (lack of independence). We could deal with this in two ways, 1) either using a generalized linear mixed model (GLMM) approach, relaxing the assumption of independence and estimating the spatial/temporal correlation between residuals, or 2) model the spatial and temporal dependence in the systematic part of the model [50]. We opted to use a Generalized Additive Model (GAM) using R's Mgcv statistical package because of its versatility and ability to fit complex models that would converge even with low numbers of observations and could capture potential complex non-linear relationships. One of the advantages of GAMs is that we do not need to determine the functional form of the relationship beforehand. In general, such models transform the mean response to an additive form so that additive components are smooth functions (e.g., splines) of the covariates, in which functions themselves are expressed as basis-function expansions. The spatial auto-correlation in the GAM model was approximated by a Markov random field (MRF) smoother, defined by the geographic areas and their neighbourhood structure. We used R's Spdep package to create a queen neighbours list (adjacency matrix) based on regions with contiguous boundaries i.e. those sharing one or more boundary point. We used a full rank MRF, which represented roughly one coefficient for each area. The local Markov property assumes that a region is conditionally independent of all other regions unless regions share a boundary. This feature allowed us to model the correlation between geographical neighbours and smooth over contiguous spatial areas, summarizing the trend of the response variable as a function of the predictors [see 51, section 5.4.2]. In order to account for variation in the response variable over time, not attributed to the other explanatory variables in our model, we used a saturated time effect for years, where a separate effect per time point is estimated.

We first tried to fit our model using a Poisson distribution. However, the mean of our dependent variable (dengue cases by region and year) was lower than its variance - E(Y) <Var(Y), suggesting that the data are over-dispersed. We also tried to fit our models using the negative binomial, quasi-Poisson and Tweedie distribution, all particularly suited when the variance is much larger than the mean. After several tests, we concluded that the Tweedie distribution worked well with our data and allowed us to model the incident rate, although results were comparative across all distributions (note that WNV infection count data, offset by a log of population at risk was used for the neg bin and quasi-Poisson models). Analysis of model diagnostic tests did not reveal any major issues; in general residuals appeared to be randomly distributed (see additional information - figures S8-S15 and Table S1 for diagnostics). Tweedie distributions are defined as subfamily of (reproductive) exponential dispersion models (ED), with a special mean-variance relationship . A random variable Y is Tweedie distributed if:

$TW_p(\mu,\sigma^2)$ if Y $ED(\mu,\sigma^2)$, with mean $= \mu = E(Y)$, positive dispersion parameter σ^2 and $Var(Y) = \mu\sigma^2$. The empirical model can then be written as:

$$E(Y) = f_1(\mathrm{X}_{it}) + f_n(\mathrm{Year}_t) + f_m(\mathrm{Region}_i)$$

Where the $f(.)$ stands for smooth functions; $E(Y)_{it}$ is equal to the WNV infection incidence in region i at time t,

which we assume to be Tweedie distributed; Xit - is a vector of economic, demographic, environmental and climate variables. $Year_t$ is a function of the time intercept and $Region_i$ represents neighbourhood structure of region.

4.7.2 Climate modelling

In order to model long term climate trends and seasonal or within-year variation, we fit a GAM model using the following equation:

$$y = \beta_0 + f_{\text{seasonal}}(x_1) + f_{\text{trend}}(x_2) + \varepsilon$$

where B_0 is the intercept, $f_{seasonal}$ and f_{trend} are smooth functions for the seasonal and trend features. Month is represented by x_1 and x_2 is the series of years in the entire time period i.e. within-year and between year (see figures S4-S7)

To model any variation in, or interaction between, the trend and seasonal features of the data. We also fit a slightly different GAM model to allow for an interaction between $f_{seasonal}$ and f_{trend} using the following equation (see figures 6-9):

$$y = \beta_0 + f(x_1, x_2) + \varepsilon$$

To create the models, we used either monthly mean regional temperatures (°C) or regional sum precipitation (mm) as mentioned in the previous section. Temperature models were fit using the Gaussian distribution and precipitation models fit using the Tweedie distribution.

References

1. Calleri, G., Angheben, A. & Albonico, M. Neglected tropical diseases in europe: rare diseases and orphan drugs? *Infection* **47**, 3–5, DOI: 10.1007/s15010-018-1241-2 (2018).

2. Hotez, P. J. Southern europe's coming plagues: Vector-borne neglected tropical diseases. *PLoS Negl Trop Dis* **10**, e0004243, DOI: 10.1371/journal.pntd.0004243 (2016).

3. Olesen, O. & Ackermann, M. Increasing european support for neglected infectious disease research. *Comput. Struct. Biotechnol. J.* **15**, 180–184, DOI: 10.1016/j.csbj.2017.01.007 (2017).

4. Morrison, A. *et al.* Impact of the dengue vector control system (dvcs) on aedes aegypti populations in iquitos, peru 2004-2005. *Am. J. Trop. Medicine Hyg.* **73**, 326–326 (2005).

5. Patz, J. A., Epstein, P. R., Burke, T. A. & Balbus, J. M. Global climate change and emerging infectious diseases. *JAMA* **275**, DOI: 10.1001/jama.1996.03530270057032 (1996).

6. Bouzid, M., Colón-González, F. J., Lung, T., Lake, I. R. & Hunter, P. R. Climate change and the emergence of vector-borne diseases in europe: case study of dengue fever. *BMC Public Heal.* **14**, 781, DOI: 10.1186/1471-2458-14-781 (2014).

7. Tabachnick, W. J. *Climate Change and the Arboviruses: Lessons from the Evolution of the Dengue and Yellow Fever Viruses*, vol. 3 of *Annual Review of Virology*, 125–145 (2016).

8. Mweya, C. N., Kimera, S. I., Stanley, G., Misinzo, G. & Mboera, L. E. G. Climate change influences potential distribution of infected aedes aegypti co-occurrence with dengue epidemics risk areas in tanzania. *Plos One* **11**, DOI: 10.1371/journal.pone.0162649 (2016).

9. Shope, R. Global climate change and infectious diseases. *Environ Heal. Perspect* **96**, DOI: 10.1289/ehp.9196171 (1991).

10. Magori, K. & Drake, J. M. *The population dynamics of vector-borne diseases*, vol. 4 (2013).

11. Patz, J. A., Martens, W. J., Focks, D. A. & Jetten, T. H. Dengue fever epidemic potential as projected by general circulation models of global climate change. *Environ Heal. Perspect* **106**, DOI: 10.1289/ehp.98106147 (1998).

12. Naish, S. *et al.* Climate change and dengue: a critical and systematic review of quantitative modelling approaches. *Bmc Infect. Dis.* **14**, DOI: 10.1186/1471-2334-14-167 (2014).

13. Gubler, D. J. Dengue, urbanization and globalization: The unholy trinity of the 21(st) century. *Trop. Medicine Heal.* **39**, 3–11, DOI: 10.2149/tmh.2011-S05 (2011).

14. Karanikolos, M. *et al.* Financial crisis, austerity, and health in europe. *Lancet* **381**, 1323–31, DOI: 10.1016/s0140-6736(13)60102-6 (2013).

15. Baumbach, A. & Gulis, G. Impact of financial crisis on selected health outcomes in europe. *The Eur. J. Public Heal.* **24**, 399–403, DOI: 10.1093/eurpub/cku042 (2014).

16. Quinn, S. C. & Kumar, S. Health inequalities and infectious disease epidemics: A challenge for global health security. *Biosecurity Bioterrorism: Biodefense Strateg. Pract. Sci.* **12**, 263–273, DOI: 10.1089/bsp.2014.0032 (2014). PMID: 25254915.

17. Wyplosz, C. The euro area after covid-19. Report, European Parliament (2020).

18. Quaglio, G., Karapiperis, T., Van Woensel, L., Arnold, E. & McDaid, D. Austerity and health in europe. *Heal. Policy* **113**, 13–19, DOI: https://doi.org/10.1016/j.healthpol.2013.09.005 (2013).

19. Suhrcke, M. *et al.* The impact of economic crises on communicable disease transmission and control: a systematic review of the evidence. *PLoS One* **6**, e20724, DOI: 10.1371/journal.pone.0020724 (2011).

20. Stuckler, D., Reeves, A., Loopstra, R., Karanikolos, M. & McKee, M. Austerity and health: the impact in the uk and europe. *Eur. J. Public Heal.* **27**, 18–21, DOI: 10.1093/eurpub/ckx167 (2017).

21. Kentikelenis, A. *et al.* How do economic crises affect migrants' risk of infectious disease? a systematic-narrative review. *Eur J Public Heal.* **25**, 937–44, DOI: 10.1093/eurpub/ckv151 (2015).

22. Kentikelenis, A. *et al.* Health effects of financial crisis: omens of a greek tragedy. *The Lancet* **378**, 1457–1458, DOI: 10.1016/s0140-6736(11)61556-0 (2011).

23. Paz, S. & Semenza, J. C. Environmental drivers of west nile fever epidemiology in europe and western asia–a review. *Int. journal environmental research public health* **10**, 3543–3562, DOI: 10.3390/ijerph10083543 (2013).

24. ECDC. Factsheet about west nile virus infection. (2018).

25. Hubálek, Z. & Halouzka, J. West nile fever–a reemerging mosquito-borne viral disease in europe. *Emerg. infectious diseases* **5**, 643–650, DOI: 10.3201/eid0505.990505 (1999).

26. ECDC. West nile virus infection. (2020).

27. Landesman, W. J., Allan, B. F., Langerhans, R. B., Knight, T. M. & Chase, J. M. Inter-annual associations between precipitation and human incidence of west nile virus in the united states. *Vector-Borne Zoonotic Dis.* **7**, 337–343, DOI: 10.1089/vbz.2006.0590 (2007).

28. Reisen, W. K., Fang, Y. & Martinez, V. M. Effects of temperature on the transmission of west nile virus by culex tarsalis (diptera: Culicidae). *J Med Entomol* **43**, 309–17, DOI: 10.1603/0022-2585(2006)043[0309:Eotott]2.0.Co;2 (2006).

29. Chen, S. *et al.* Estimating west nile virus transmission period in pennsylvania using an optimized degree-day model. *Vector-Borne Zoonotic Dis.* **13**, 489–497, DOI: 10.1089/vbz.2012.1094 (2013).

30. Kamiya, T. *et al.* Temperature-dependent variation in the extrinsic incubation period elevates the risk of vector-borne disease emergence. *Epidemics* **30**, 100382, DOI: 10.1016/j.epidem.2019.100382 (2020).

31. Mohammed, A. & Chadee, D. D. Effects of different temperature regimens on the development of aedes aegypti (l.)(diptera: Culicidae) mosquitoes. *Acta tropica* **119**, 38–43 (2011).

32. Tun-Lin, W. *et al.* Reducing costs and operational constraints of dengue vector control by targeting productive breeding places: a multi-country non-inferiority cluster randomized trial. *Trop. Medicine Int. Heal.* **14**, 1143–1153, DOI: 10.1111/j.1365-3156.2009.02341.x (2009).

33. Watts, D. M., Burke, D. S., Harrison, B. A., Whitmire, R. E. & Nisalak, A. Effect of temperature on the vector efficiency of aedes aegypti for dengue 2 virus. *The Am. journal tropical medicine hygiene* **36**, 143–152 (1987).

34. Koenraadt, C. J. M., Möhlmann, T. W. R., Verhulst, N. O., Spitzen, J. & Vogels, C. B. F. Effect of overwintering on survival and vector competence of the west nile virus vector culex pipiens. *Parasites Vectors* **12**, 147, DOI: 10.1186/s13071-019-3400-4 (2019).

35. Soverow, J. E., Wellenius, G. A., Fisman, D. N. & Mittleman, M. A. Infectious disease in a warming world: how weather influenced west nile virus in the united states (2001-2005). *Environ Heal. Perspect* **117**, 1049–52, DOI: 10.1289/ehp.0800487 (2009).

36. Takeda, T., Whitehouse, C. A., Brewer, M., Gettman, A. D. & Mather, T. N. Arbovirus surveillance in rhode island: assessing potential ecologic and climatic correlates. *J Am Mosq Control. Assoc* **19**, 179–89 (2003).

37. Cabrera, M. & Taylor, G. Modelling spatio-temporal data of dengue fever using generalized additive mixed models. *Spatial Spatio-temporal Epidemiol.* **28**, 1–13, DOI: https://doi.org/10.1016/j.sste.2018.11.006 (2019).

38. ETC/ULS. Updated clc illustrated nomenclature guidelines. Report, European Environment Agency (2019).

39. Wallace, R. *et al. Clear-cutting disease control: capital-led deforestation, public health austerity, and vector-borne infection* (Springer, 2018).

40. Samb, B. *et al.* Prevention and management of chronic disease: a litmus test for health-systems strengthening in low-income and middle-income countries. *The Lancet* **376**, 1785–1797 (2010).

41. Caliendo, A. M. *et al.* Better tests, better care: improved diagnostics for infectious diseases. *Clin. Infect. Dis.* **57**, S139–S170 (2013).

42. Kruk, M. E. & Freedman, L. P. Assessing health system performance in developing countries: a review of the literature. *Heal. policy* **85**, 263–276 (2008).

43. López, F. C., Martín, E. A., Maleras, R. G. & Ginestar, C. M. Relationship between research funding in the spanish national health system and the burden of disease. *Rev Esp Salud Publica* **83**, 137–151 (2009).

44. Eurostat. Nuts - nomenclature of territorial units for statistics. (2020).

45. Eurostat. Agriculture, forestry and fishery statistics. Report (2019). DOI: 10.2785/798761.

46. SEDAC. Gridded population of the world, version 4 (gpwv4): Population count, revision 11. (2018).

47. Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M. & Jones, P. D. An ensemble version of the e-obs temperature and precipitation data sets. *J. Geophys. Res. Atmospheres* **123**, 9391–9409, DOI: https://doi.org/10.1029/2017JD028200 (2018).

48. EEA. Copernicus land monitoring service. (2018).

49. Pekel, J.-F., Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* **540**, 418–422, DOI: 10.1038/nature20584 (2016).

50. Aswi, A., Cramb, S. M., Moraga, P. & Mengersen, K. Bayesian spatial and spatio-temporal approaches to modelling dengue fever: a systematic review. *Epidemiol. infection* **147**, 1–14, DOI: 10.1017/S0950268818002807 (2018).

51. Wood, S. N. *Generalized additive models: an introduction with R* (CRC press, 2017).

Acknowledgements (not compulsory)

This research was funded by ICTA's Maria de Maeztu Unit of Excellence, awarded by the Spanish Ministry of Economy and Competitiveness. The award is the highest institutional recognition of scientific research in Spain.Thanks also to Patrizia Ziveri and Pedro Manuel Gonzalez Hernandez for supporting the project.

Author contributions statement

MW led the work and was responsible for the conceptualization of the project, data collection, data processing, formulation of the methodology, statistical analysis, modelling, analysis and writing the original draft and interpreting the results. All authors made substantive contributions to developing the original paper outline. PK gave critical advice throughout the project. All contributors revised the manuscript and copy-edited the final submission version. All the authors were also involved in revising the manuscript critically for important intellectual content. All authors read and approved the final manuscript. PK, PGM VSM supervised the project and are responsible for formulating ideas for the umbrella project Impacts of Climate Change (CC) on Human Health (HH) at ICTA-UAB: Integrating socio-economic and policy studies with natural science studies to enhance consilience of climate policy science.

5 Tables

Table 1. Final dataset - 2007-2019

Statistic	N	Min	Max	Mean	St. Dev.
WNV cases	2,158	0	100	1.649	6.012
Human population	2,158	6,254.326	10,534,640.000	443,384.400	513,000.000
Mean temp winter (C)	2,158	−6.072	14.564	3.190	3.623
Mean temp spring (C)	2,158	4.092	18.684	12.433	2.157
Mean temp summer (C)	2,158	13.109	28.011	22.465	2.285
Days of rain in winter	2,158	0	68	30.156	12.546
Days of rain in spring	2,158	0	71	31.723	12.918
Days of rain in summer	2,158	0	65	26.021	14.374
Spring surface water extent Z-score (30m2)	2,158	−2.876	2.404	0.000	0.958
Summer surface water extent Z-score (30m2)	2,158	−3.301	3.080	−0.000	0.958
Continuous urban fabric % cover	2,158	0.000	45.056	1.336	6.754
Discontinuous_urban_fabric (% cover)	2,158	0.534	60.457	5.511	7.044
Wetlands (% cover)	2,158	0.000	25.460	0.569	2.026
Arable_land (% cover)	2,158	0.000	86.307	33.893	22.172
GDP (2007=100%)	2,158	74.878	188.752	108.992	24.171
Agri, forest + fish spending (2007=100%)	2,158	27	251	80.202	35.165
Waste water management spending (2007=100%)	2,158	5	352	93.476	53.933
Health spending (2007=100%)	2,158	60	212	114.802	33.307

Table 2. Final Regression Models: EDF value is reported as the coefficient, and DF is included in parentheses (not standard errors because a chi-square test is used for the smooth terms

	Clim model	Land-use model	Econ model	Full model
Intercept	-2.22^{***}	-2.13^{***}	-2.29^{***}	-2.44^{***}
	(0.48)	(0.45)	(0.34)	(0.50)
Mean temp summer (C)	1.00^{***}			1.00^{*}
	(1.00)			(1.00)
Mean temp winter (C)	1.96^{***}			1.92^{***}
	(1.99)			(1.98)
Days of rain in winter	1.89^{***}			1.91^{***}
	(1.98)			(1.98)
Days of rain in summer	1.07^{**}			1.37^{**}
	(1.13)			(1.58)
Summer surface water extent (30m2)	1.69^{***}			1.22^{**}
	(1.90)			(1.39)
Continuous urban fabric %		1.00		1.00
		(1.00)		(1.00)
Discontinuous_urban_fabric %		1.00		1.00^{*}
		(1.00)		(1.00)
Wetlands %		1.00^{**}		1.00^{*}
		(1.00)		(1.00)
Arable_land %		1.81^{***}		1.61^{***}
		(1.89)		(1.75)
GDP index (2007=100%, 1yr lag)			1.18^{**}	1.96^{***}
			(1.31)	(1.99)
Agri, forest + fish spending (2007=100%)			1.95^{***}	1.84^{***}
			(1.99)	(1.96)
Waste water management spending (2007=100%)			1.65^{***}	1.00^{***}
			(1.87)	(1.00)
Health spending (2007=100%)			1.00^{***}	1.96^{***}
			(1.00)	(1.99)
NUTS 3 regions	75.82^{***}	80.54^{***}	88.67^{***}	71.66^{***}
	(106.69)	(111.42)	(120.91)	(101.47)
Year	11.72^{***}	11.75^{***}	11.48^{***}	11.64^{***}
	(12.00)	(12.00)	(12.00)	(12.00)
AIC	3952.13	3992.59	3927.81	3873.14
BIC	4532.48	4569.26	4575.72	4509.05
Log Likelihood	−1873.83	−1894.71	−1849.78	−1824.55
Deviance	3728.81	3895.52	3557.78	3396.17
Deviance explained	0.63	0.62	0.65	0.66
Dispersion	2.84	2.95	2.74	2.65
R^2	0.29	0.36	0.31	0.27
GCV score	1898.91	1927.27	1884.48	1855.32
Num. obs.	2158	2158	2158	2158
Num. smooth terms	7	6	6	15

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

6 Figures

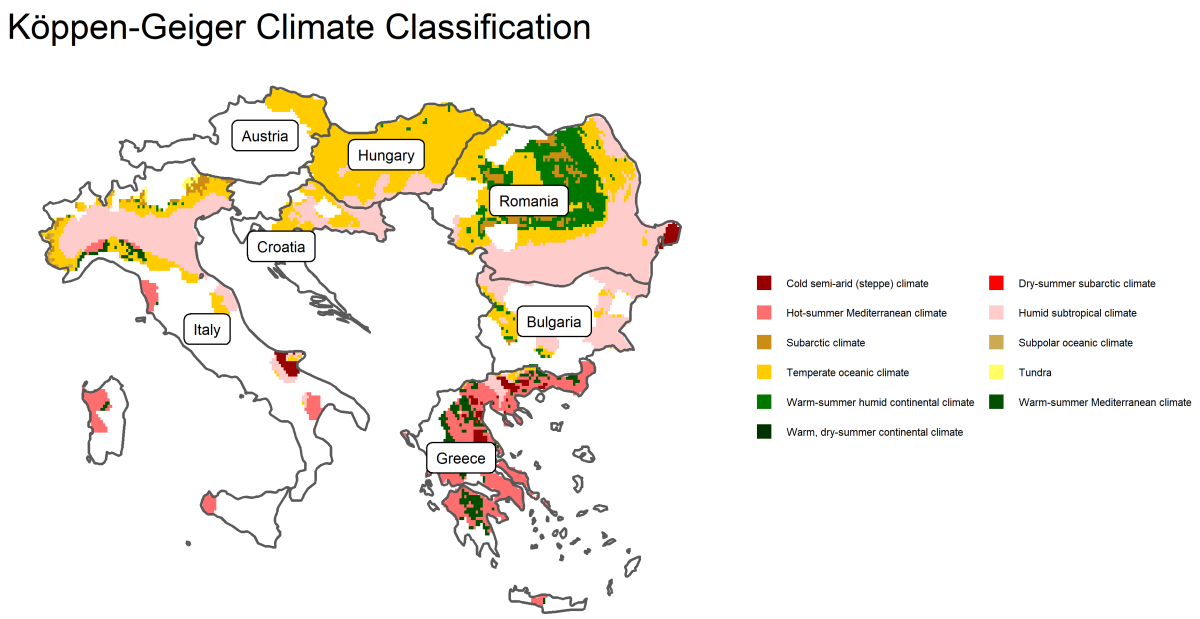

Figure 1. Koppen-Geiger Climate Classification in Study Regions (Source: koeppen-geiger.vu-wien.ac.at)

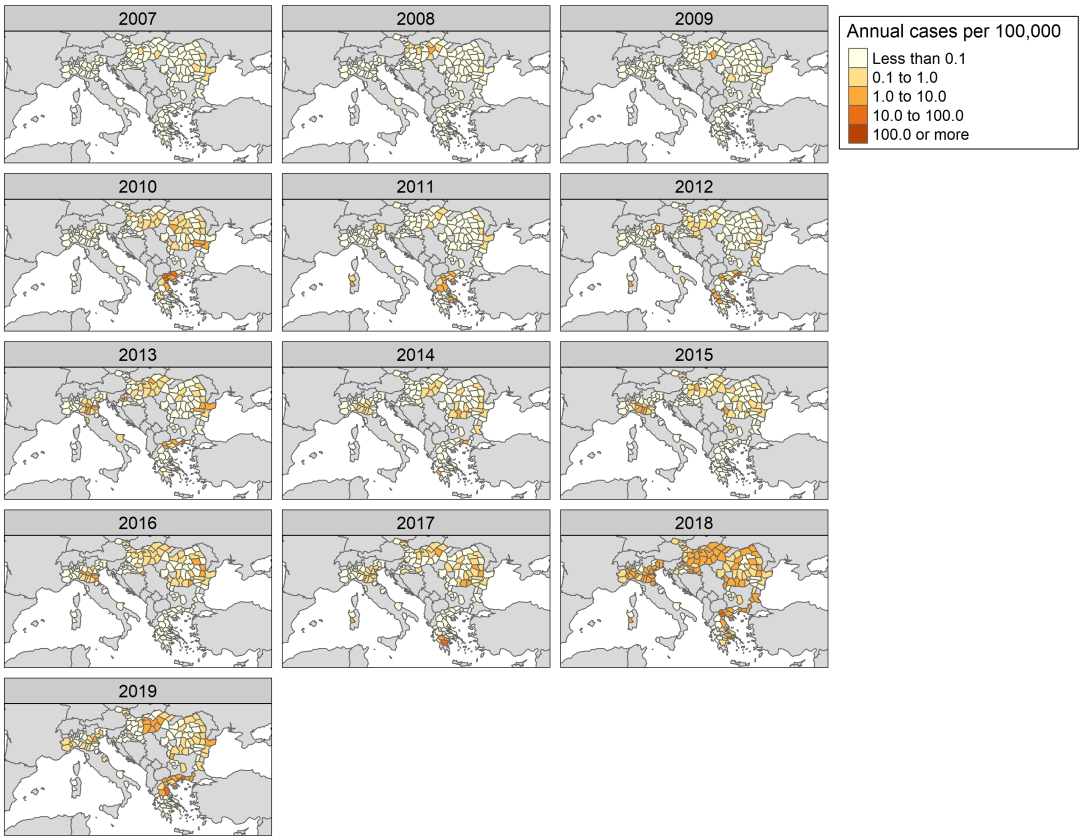

Figure 2. European Nuts 3 Distribution of human WNV infections 2006 to 2019.

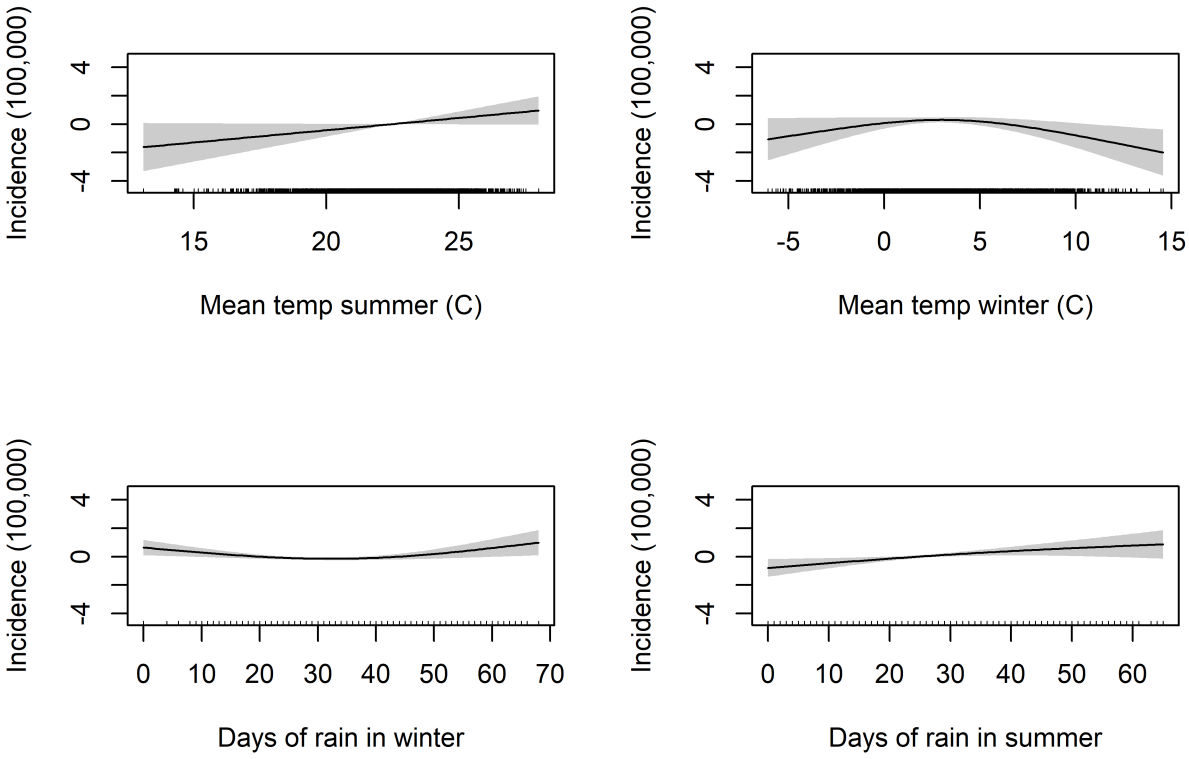

Figure 3. Partial effects main model (1).

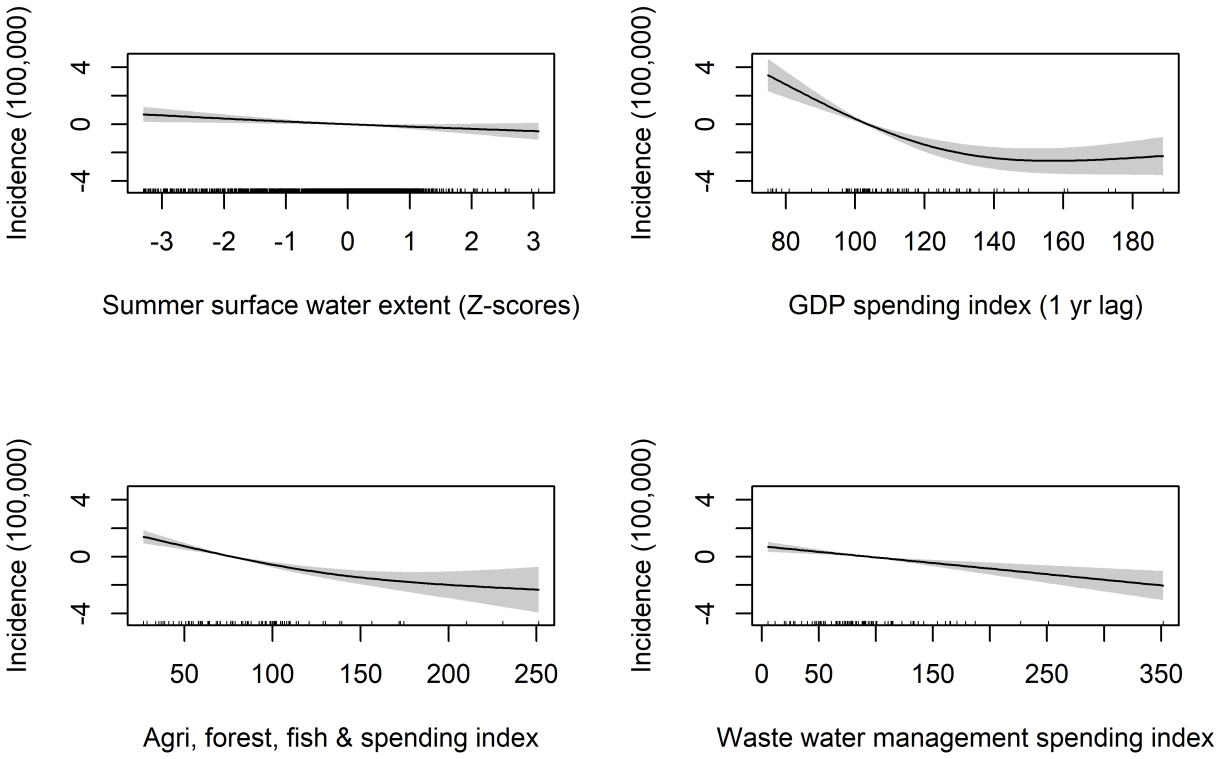

Figure 4. Partial effects main model (2).

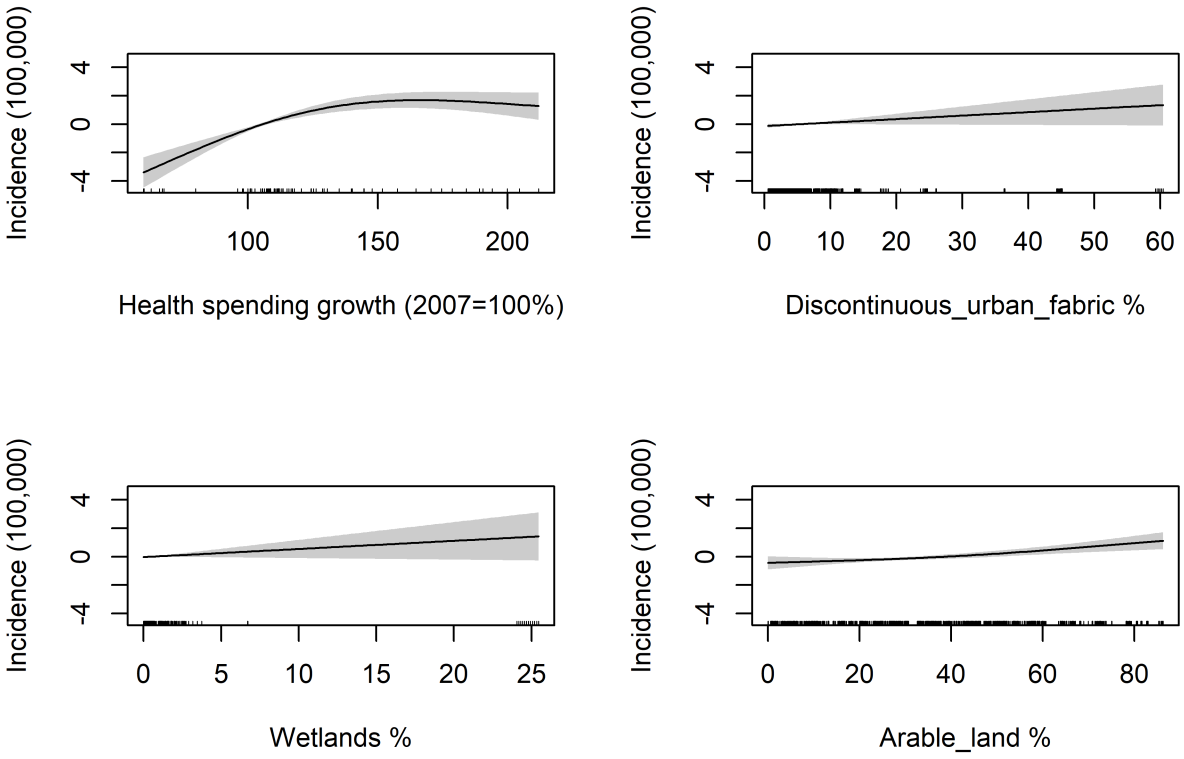

Figure 5. Partial effects main model (3).

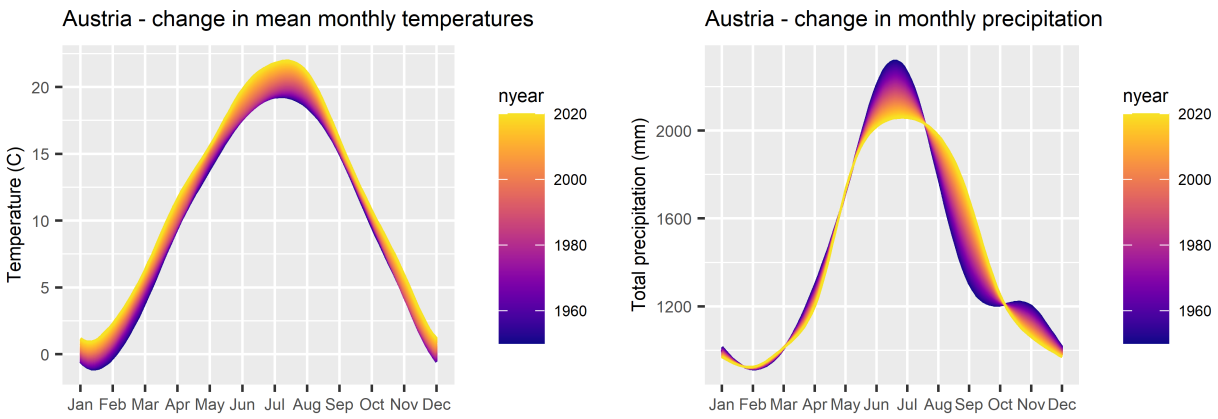

Figure 6. Seasonal climate trends (Data source: E-OBS version 22.0e).

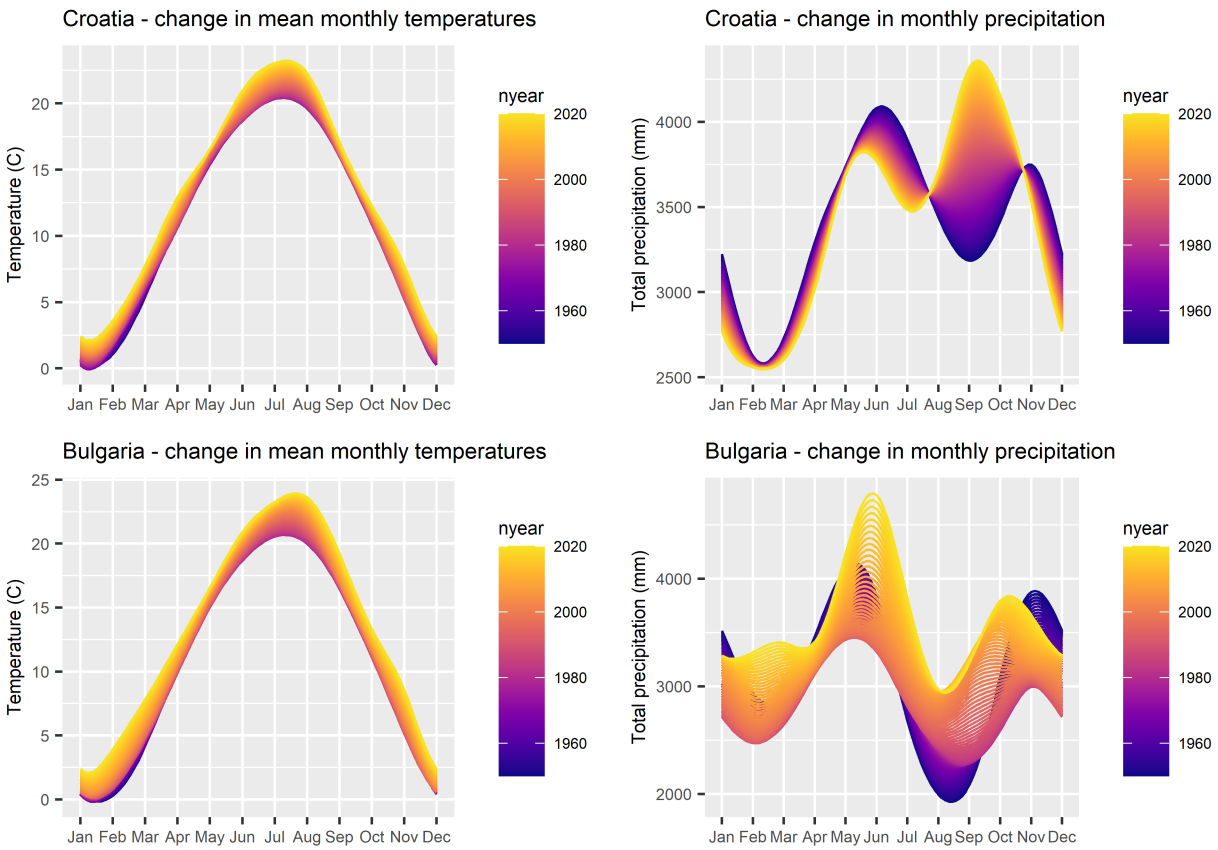

Figure 7. Seasonal climate trends (Data source: E-OBS version 22.0e).

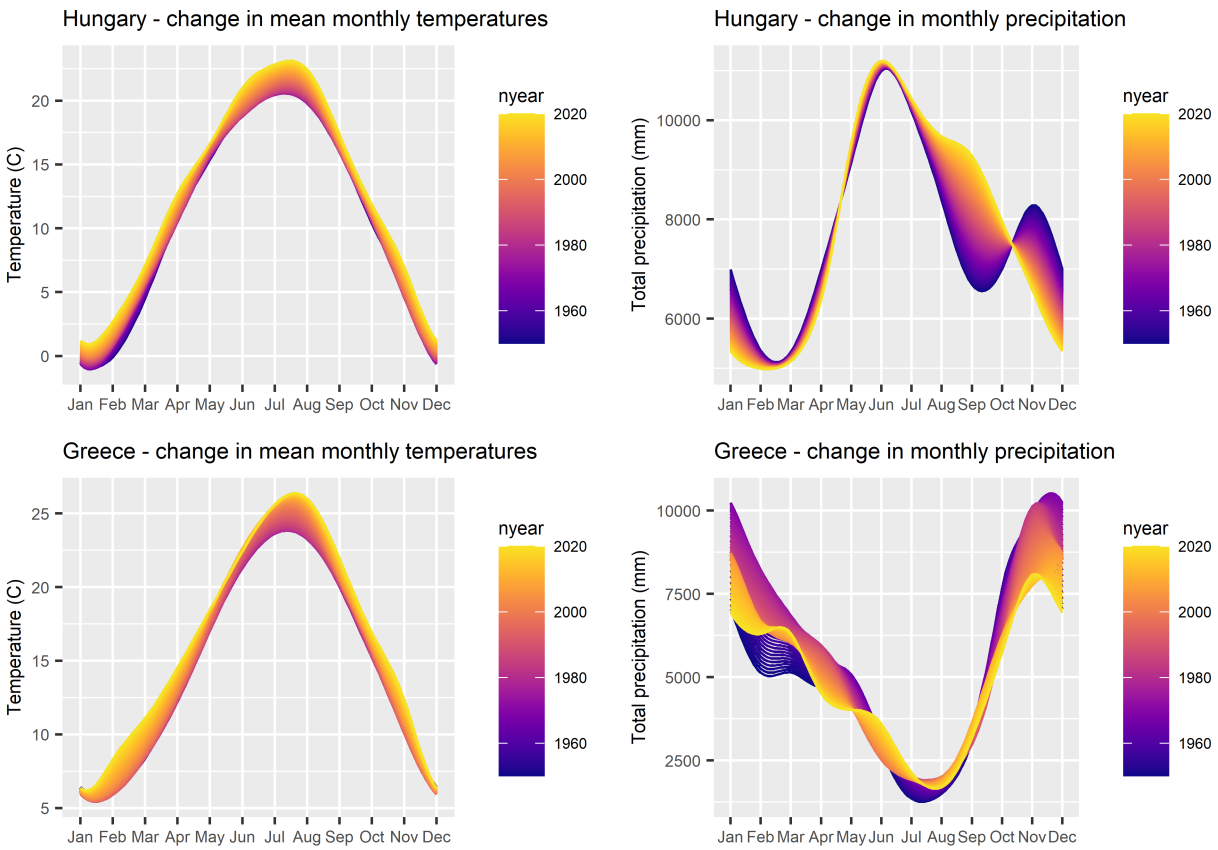

Figure 8. Seasonal climate trends (Data source: E-OBS version 22.0e).

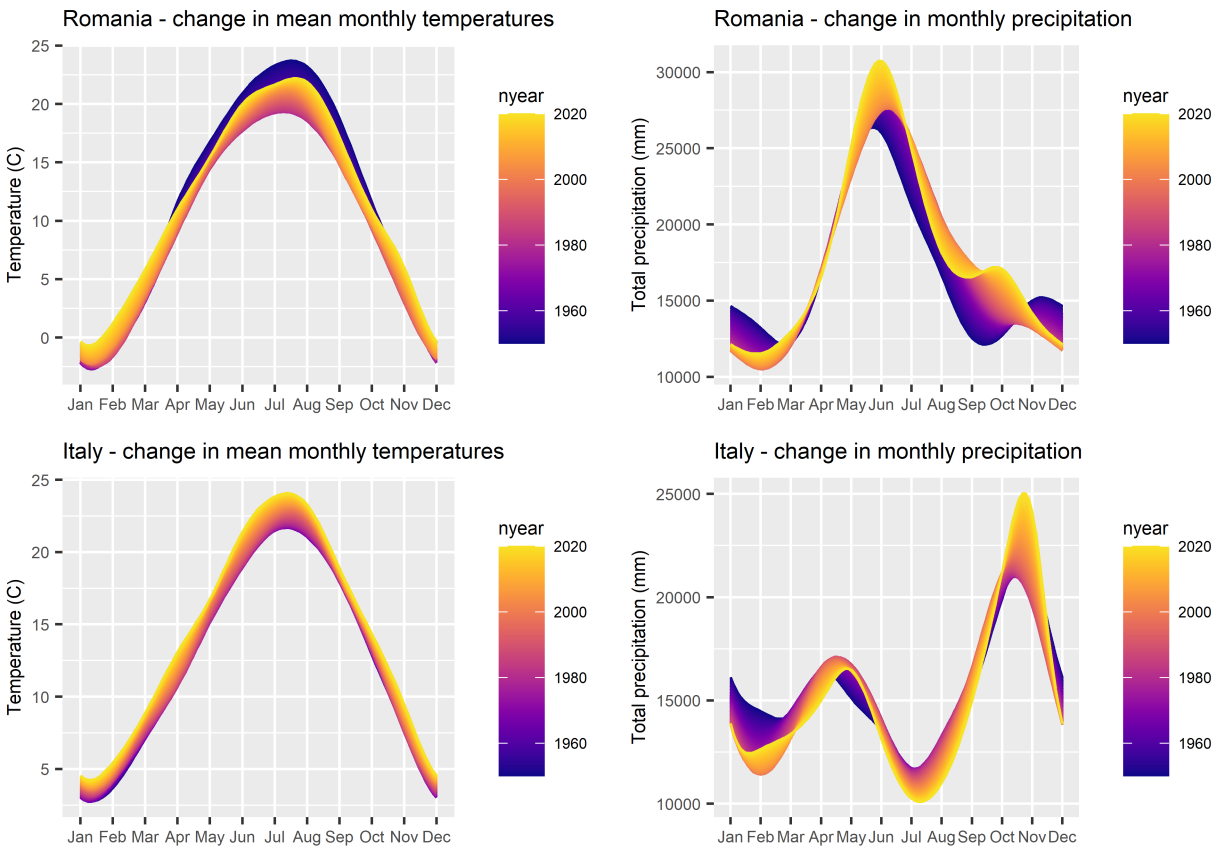

Figure 9. Seasonal climate trends (Data source: E-OBS version 22.0e).