scientific reports

Check for updates

OPEN

Early warning system of the seasonal west nile virus infection risk in humans in northern greece, 2020–2024

Anastasia Angelou[1], Areti Pappa[1], Peter V. Markov[2,3], Sandra Gewehr[4], Nikolaos I. Stilianakis[2,5] & Ioannis Kioutsioukis[1]✉

This study introduces a novel forecasting tool for West Nile Virus (WNV) risk at the municipal level, based on a previously developed climate-dependent epidemiological model coupled with a data-driven model and field data in an ensemble framework. The modelling system was evaluated at the municipal level in the regions of Central Macedonia (RCM) and Thessaly (RTH) in Greece for the period 2020–2024 and classifies the municipalities according to the risk level of occurrence of a WNV human case (WNVhc) within each forecast season. The modelling system produces seasonal forecasts updated monthly throughout the entire mosquito breeding season (April-September). From the first forecast of the year, WNVhc were correctly predicted at 71% of the infected municipalities in RCM and 60% of the infected municipalities in RTH, percentages that increased by at least 20% after July when surveillance data were incorporated. The observed number of infected humans was within the predicted range, highlighting the part of the region with an expected outbreak. The approach shows that the coupling of a climate-dependent epidemiological model at local level with a data-driven model, incorporating entomological data to predict the risk of WNVhc, can be a useful tool in planning control strategies.

Keywords Early warning, West nile virus, Spatiotemporal epidemiological model, *Culex* mosquitoes, Risk map, Public health surveillance

West Nile fever (WNF) is a zoonosis caused by the mosquito-borne West Nile virus (WNV), part of the Flaviviridae family[1,2]. The transmission cycle involves birds as main hosts and mosquitoes, mainly of the *Culex pipiens*species, as vectors[3]. Bites by infected mosquitoes may inject the virus into humans and animals[4]. Most humans infected with WNV remain asymptomatic, however, old and immunocompromised people are at higher risk of developing West Nile Neuroinvasive Disease (WNND) with symptoms in the Central Nervous System (CNS)[5]. Currently, treatment of patients with WNND is only supportive and no human vaccine is available[6].

The first WNVhc in Greece was detected in 2010. During the period 2010–2024, 2068 WNVhc were confirmed in the country, with 49% and 13% of them reported at the RCM and the RTH respectively. Just under half of the total WNVhc occurred in three non-consecutive years, namely 2010 (262 cases, 35 deaths), 2018 (316 cases, 50 deaths) and 2022 (286 cases, 32 deaths)[7]. Compared to other European countries, Greece reported a high number of WNVhc in the investigated 15-year period, with most cases occurring in RCM and in RTH. Specifically in these regions, excluding the period 2014–2017 when a WNV hiatus occurred, from 13% (2012) to 99% (2010) of the cases of the entire country were recorded[8].

A wide range of predictors and modelling approaches can be identified from the literature to forecast epidemics of vector-borne infections like WNV[9,10]. Over the past years, several models for early warning for WNV transmission have been developed. Mechanistic and data driven models represent the basic approaches. Data-driven models use statistical techniques and historical data to forecast WNV risk and outbreaks, relying on empirical data and often employing machine learning or statistical methods[11–13]. Many of these models incorporate real-time climatic variables—such as temperature, precipitation, vegetation indices, and hydrological data—given the strong correlations between these factors and mosquito infection rates, which subsequently influence human WNV risk[14–16]. Recent advancements in modelling have further refined predictive accuracy by

[1]Department of Physics, University of Patras, Rio 26504, Greece. [2]Joint Research Centre (JRC), European Commission, Via E. Fermi 2749, Ispra 21027, VA, Italy. [3]London School of Hygiene and Tropical Medicine, London, UK. [4]Ecodevelopment S.A, Thessaloniki, Greece. [5]Department of Biometry and Epidemiology, University of Erlangen-Nuremberg, Waldstraße 6, 91054 Erlangen, Germany. ✉email: kioutio@upatras.gr

integrating interaction terms for environmental variables and validating against extensive historical datasets[15,16]. These models, spanning both the United States of America and Europe, underscore the significance of regional climatic and ecological factors in understanding WNV transmission dynamics[17,18]. Frameworks that integrate mosquito infection indices with human incidence data have demonstrated considerable efficacy in producing early warning indicators, achieving high accuracy in projecting seasonal peaks and assessing outbreak probabilities[13,16,17]. More recent approaches apply machine learning to assess the predictive influence of multiple transmission drivers in Europe, while spatially resolved ensemble models have been developed for specific regions, such as California, to refine forecasts of mosquito infection rates[19,20].

Mechanistic models capture the biological and environmental processes that drive WNV transmission, integrating complex ecological and meteorological factors to simulate outbreak dynamics and assess transmission risks. These models commonly focus on host-vector interactions, vector geographic distribution abundance, and climatic influences, which are essential for predicting spillover risk from animal hosts to humans[21–23]. Studies have shown that bird community composition, mosquito population dynamics, and seasonal climate variations are key drivers of WNV spread, with high-competence bird species and favorable environmental conditions significantly elevating human infection risks[21,24,25]. Moreover, regional models in both the U.S. and Europe underscore the importance of localized climate data, revealing temperature, precipitation, and vapor pressure as critical predictors of infection rates and human cases[23,26,27]. These mechanistic approaches also highlight the potential of targeted interventions; for example, models have explored the impact of ivermectin-treated bird feeders on reducing mosquito infection rates, suggesting viable control measures for WNV prevention[28].

The aim of our work was to develop a reliable and scalable early warning system for the prediction of the risk of WNV infection in humans to prevent outbreaks and minimize their impact at the local, regional, and country levels, and help efforts to minimize virus prevalence. We investigated the forecasting potential of our previously developed climate-dependent spatial epidemiological model for the transmission risk of WNV to humans[29–31]. The last version of the model[31] used here describes the dynamics of a WNV epidemic in population health states of mosquitoes, birds and humans employing mosquito-specific characteristics of transmission, such as host selection and temperature-dependent transmission of the virus. Due to the challenge of capturing the typically little known initial epidemiological conditions such as the density and distribution of birds and mosquitoes, we also developed a data-driven model to infer the initial infected mosquito populations. The forecast skill of the coupled models (mechanistic and data-driven) was evaluated at all municipalities in RCM and RTH in Greece for each year of the test period 2020–2024, utilizing prior calibration of the modelling system in the historic period extending from 2010 to the year ending before each forecast year. We examine the year-to-year variability in the predictive capacity of the model, as well as the short-term and long-term early warning potential. Moreover, we predict the annual range of expected WNVhc and highlight the part of the province with the highest number of expected cases.

Results

Model evaluation: detection of WNV events

Surveillance-network benchmarking

Prior to investigating the fMIMESIS early warning skill, we examine the skill embedded in the observational datasets. The investigated 'models', at municipality scale, are:

- M1 (persistence): detection of a WNVhc in year *y* given a WNVhc observed in year *y-1*.
- M2 (IM): detection of a WNVhc in year *y* given an infected mosquito sample found in year *y*.

M1 yields predictions with lag at seasonal scale while M2 typically provides 2–4 weeks warnings.

The geographical variability of the M1 and M2 models for the forecast period 2020–2024 is presented in Fig. 1 with the aid of the Critical Success Index (CSI) (see Model verification measures section) calculated for each municipality. M1 has an average CSI of 0.50 in RCM and 0.37 in RTH; in other words, at half (about one third) of the municipalities with detected WNVhc in a year in RCM (RTH), WNVhc will also appear in the next year. The smaller value in RTH denote a weaker endemicity compared to RCM. CSI however has a significant multiannual variability (Table 1) ranging in RCM from 0.28 (2021) to 0.77 (2023) and like RTH from 0 (2021, 2022) to 0.54 (2024), limiting its forecast capacity. M2 demonstrates an average CSI that is not superior compared to the one of M1, both in RTH (0.39 vs. 0.37) and RCM (0.40 vs. 0.50). Therefore, the finding of infectivity in mosquitoes can serve as an indicator for the transmission of the virus to humans, but these municipalities are very few in comparison to the total number of municipalities with WNVhc of each year. The year-to-year CSI can range in RCM from 0.33 (2020) to 0.56 (2021) and in RTH from 0 (2020, 2021) to 0.67 (2022). The CSI skill of M2 does not generally contain false alarms as can be deducted from the generally similar values between Probability of Detection (POD) and CSI (in all years except 2021, when there were several municipalities in which, although at least one positive mosquito sampling was found, no human case was recorded). This implies that its value can be increased by eliminating the misses, through increased sample tests. Unlike M2, the hits in the CSI of M1 are surrounded with both false alarms and misses.

Hence, the presented analysis demonstrated (a) the skill threshold derived from the data, (b) the diversity of the year 2021, where a major number of false alarms was observed across all investigated regions in Greece, as well as in Italy (not shown).

Evaluation of the model's predictive capacity to detect events

The fMIMESIS produced seasonal forecasts issued monthly from April to September. The variability of the modelled CSI (solid line) and POD (dotted line) for the municipalities at risk are shown in Fig. 2 for each forecast month. The lines have small positive slopes in the first three months (April-June) that in most cases

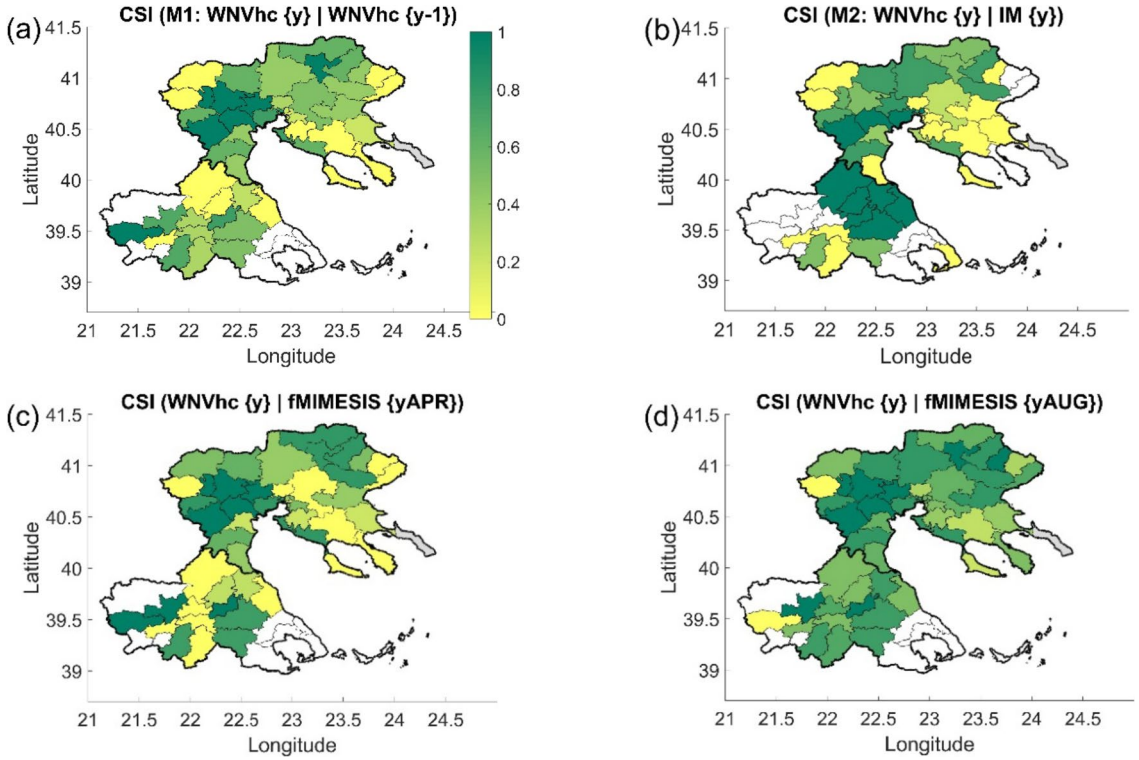

Fig. 1. Maps of CSI for different models (a) M1, (b) M2, (c) fMIMESIS forecast in April, (d) fMIMESIS forecast in August. White colour denotes municipalities with no detected WNVhc (and no surveillance in subplot b) (see model verification measures section). The maps have been created using the Mapping Toolbox of MATLAB R2024a (Natick, Massachusetts, United States: The MathWorks Inc, 2024, https://www.mathworks.com).

Model	Region	POD/CSI (Year)					
		2020	2021	2022	2023	2024	Average
M1: WNVhc {y} | WNVhc {y-1}	RCM	0.62/0.55	0.73/0.28	0.31/0.31	1.00/0.77	0.69/0.56	0.64/0.50
	RTH	0.83/0.38	-/0	0/0	0.44/0.36	0.64/0.54	0.50/0.37
M2: WNVhc {y} | IM {y}	RCM	0.33/0.33	0.91/0.56	0.40/0.40	0.56/0.52	0.38/0.34	0.42/0.40
	RTH	-/0	-/0	0.80/0.67	0.57/0.57	0.30/0.27	0.42/0.39
WNVhc {y} | fMIMESIS {yAPR}	RCM	0.69/0.56	1.00/0.39	0.63/0.63	0.85/0.72	0.59/0.49	0.71/0.56
	RTH	0.83/0.38	-/0	0.67/0.44	0.56/0.45	0.45/0.42	0.60/0.42
WNVhc {y} | fMIMESIS {yAUG}	RCM	0.88/0.72	1.00/0.35	0.86/0.86	0.96/0.76	0.93/0.77	0.91/0.70
	RTH	0.67/0.31	-/0	1.00/0.75	1.00/0.82	1.00/0.92	0.94/0.68

Table 1. POD and CSI values applying different models.

experience a sharp increase in July, when field data are assimilated. For RCM, the multiannual CSI (POD) value for the month of April (Table 1) equals 0.56 (0.71), while for August increased to 0.70 (0.91), i.e. have in both months values above 0.5. For example, in August, in 91% of the municipalities that the model characterized as "municipalities at risk" a human case was observed (the other 9% of observed WNVhc are 'missed') while considering also the modelled 'false-alarms' the success rate falls to 0.70. In RTH, the multiannual CSI and POD for April (August) equal 0.42 (0.68) and 0.60 (0.94) respectively. The August forecast is improved compared to April, with the CSI in both regions equal to approximately 0.70. The scores, although larger compared to M1 and M2, shows an increased number of false-alarms in RTH relative to RCM. Moreover, the diversity seen earlier for M1 and M2 in 2021 is also found in the fMIMESIS scores. The geographical variability of the CSI score in April and August for the forecast period 2020–2024 is presented in Fig. 1, together with M1 and M2. fMIMESIS was able to early detect few municipalities with zero skill in both data models M1 and M2.

Figure 3 presents the CSI as a function of Risk Levels (RL) for WNV transmission to humans in the first forecast of the year. In both RCM and RTH, the CSI increases significantly as levels rise, with the highest values observed at the upper thresholds (RL = 4 or RL = 5). In the RCM, the CSI remains diminished at the lower RL end improves steadily, peaking sharply at the highest thresholds, particularly in 2023 and 2024. A similar pattern is

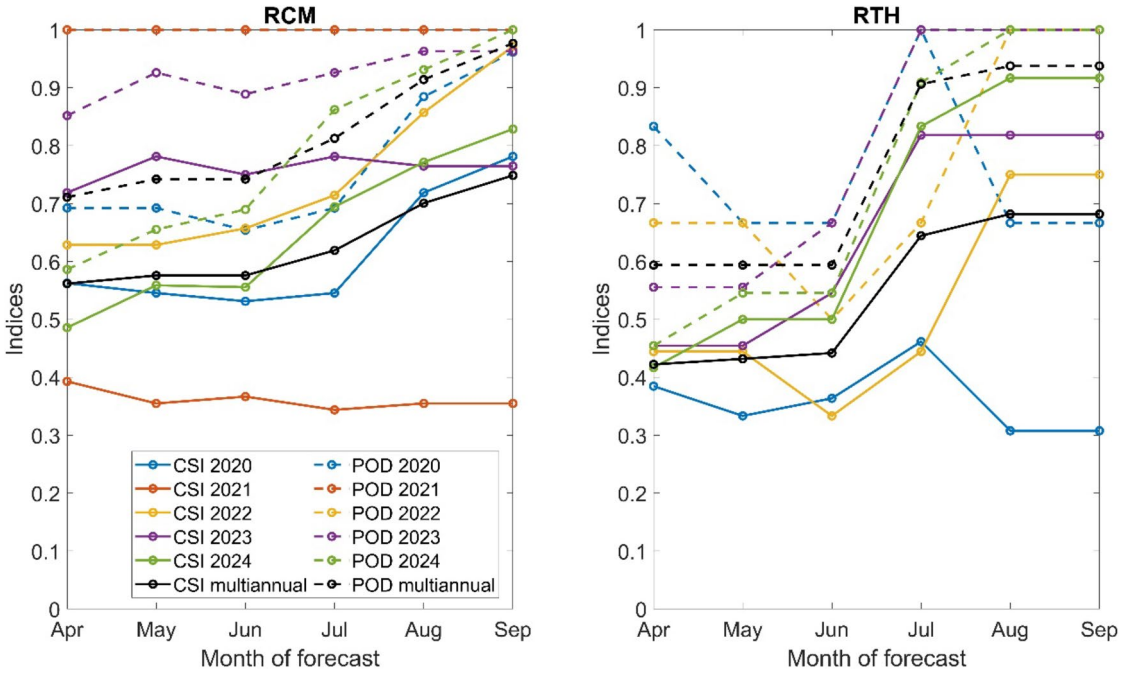

Fig. 2. POD and CSI for municipalities at risk in Central Macedonia (left) and Thessaly (right) as a function of month of production of the forecast (April-September). The different years are represented in different colours, as shown in the legend. POD and CSI lines for RCM in 2022 overlap, while 2021 for RTH has been excluded due to the absence of recorded human cases in the region.

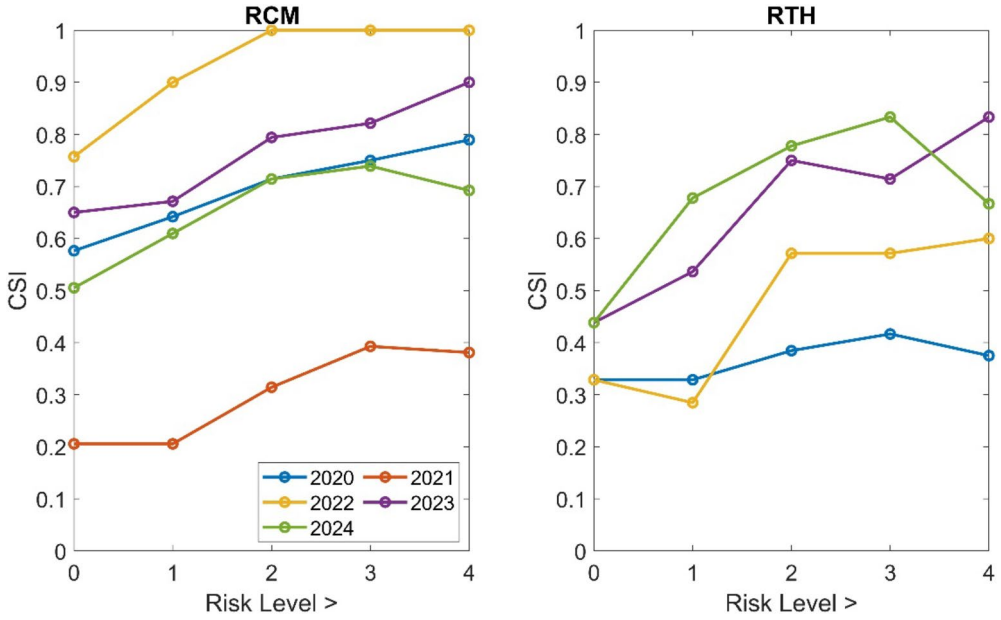

Fig. 3. CSI according to the Risk Level of occurrence of human cases in Central Macedonia (left) and Thessaly (right) in the April forecast. The different years are represented in different colours, as shown in the legend.

observed in the RTH, although the variability between years is more pronounced. While improvements are less consistent at lower levels, the top thresholds show the highest predictive performance across all years. To address this trade-off, the decision was made to classify "municipalities at risk" as those with RL of 4 or 5. This approach ensures that the maximum CSI is achieved while preserving a sufficient number of municipalities identified as being "at risk".

Model evaluation: Estimation of WNVhc

We now evaluate the modelled WNVhc for the municipalities at risk. Utilizing the full ensemble, the range of human cases can be estimated. However, since the full ensemble can contain few simulations with extreme values, we used as modelled range of WNVhc for each municipality at risk the corresponding interquartile range.

The predicted WNVhc range in the municipalities at risk, alongside the observed number of WNVhc at those municipalities is presented in Fig. 4a for all forecast years. For the WNVhc it turns that the forecast issued in June gives the most reliable result in terms of both the RMSE being minimum and the range being the narrowest with the observed WNVhc falling within (Fig. 4a). The range of modelled WNVhc follows the annual trend of the observed human cases. The range of expected cases for all forecast months, as well as the RMSE are presented in the Supplementary Material (Table S1, Figure S1).

For the municipalities at risk where human cases were indeed ultimately observed, we calculate the modelled WNVhc as the median value of the expected WNVhc. Figure 4b maps the modelled WNVhc for the period 2020–2024 according to the June forecast alongside observed values. Maps of numbers of predicted and observed WNVhc showed similar patterns in each year, with more human cases in RCM, especially northeastern or southwestern - depending on the year, than in RTH. The model accurately predicted the municipalities where

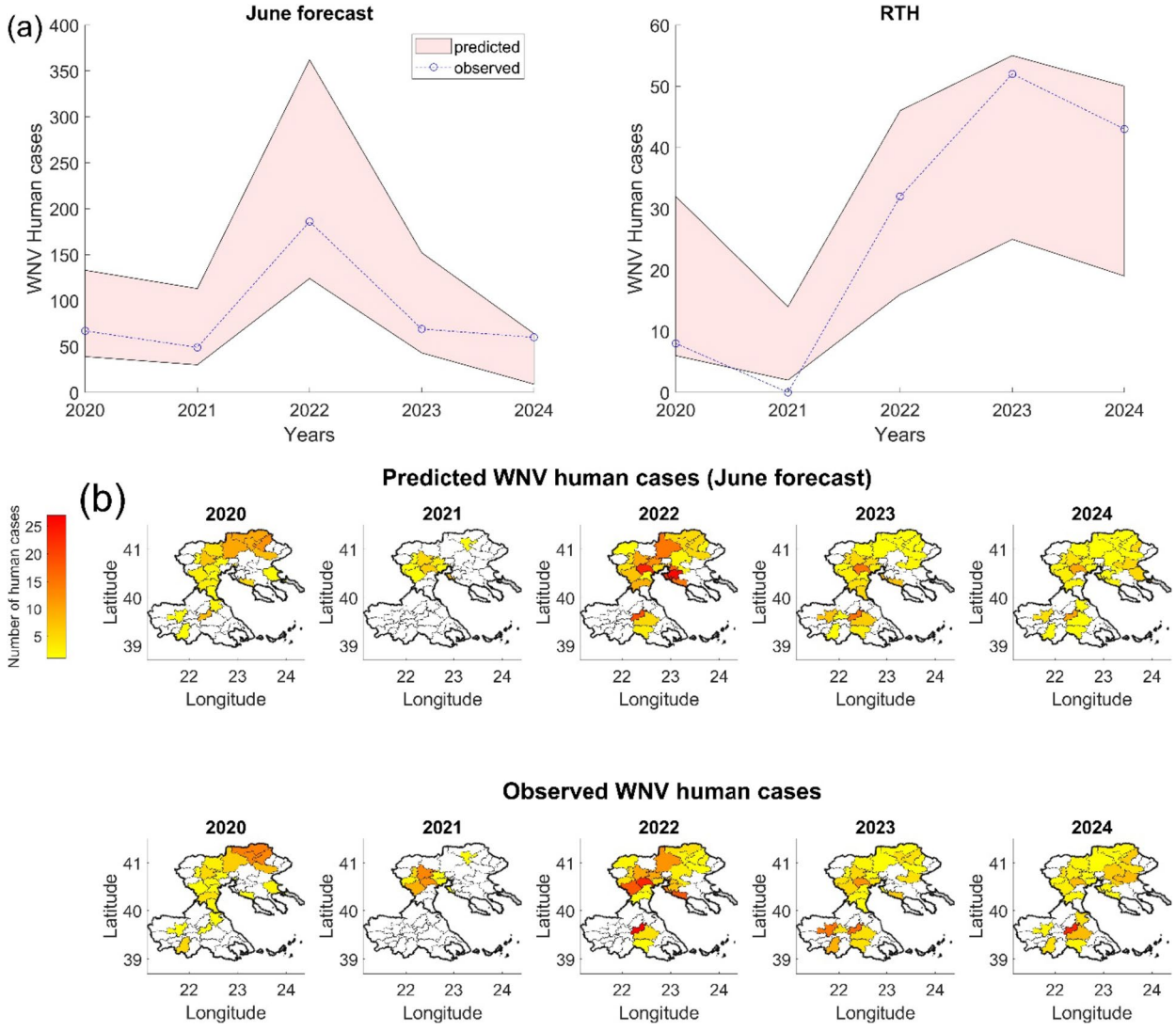

Fig. 4. (a) Range of predicted WNVhc in municipalities at risk, in the region of Central Macedonia (left) and Thessaly (right) according to June forecast (shaded area) for 2020–2024. The observed number of WNVhc at the municipalities at risk is plotted with the dotted line. (b) Maps with the number of predicted (top) and observed (bottom) cases for the municipalities where WNVhc has been both detected and modelled (i.e., WNVhc for the true positive events), for the period 2020–2024 derived from the June forecast. The modelled range at each municipality has been replaced with the corresponding median value. The maps have been created using the Mapping Toolbox of MATLAB R2024a (Natick, Massachusetts, United States: The MathWorks Inc, 2024, https://www.mathworks.com).

the most human cases were observed each year, even if it did not predict the exact number of cases. The parts of the province where the most human cases were recorded was therefore correctly predicted.

Potential extension of the model to different areas

Our model was also successfully applied in the region of Veneto in Italy at the province level, due to the lower spatial resolution of the epidemiological and entomological data. For the Veneto region, highlighting the provinces at risk in April the multiannual value of the POD was 0.88, indicating that the model correctly predicts 88% of the infected provinces of the region. The annual POD varied between 0.50 (2021) and 1.00 (2020, 2023, 2024). The long-term value of the CSI is 0.65 in April and 0.70 from August onwards, with yearly values varying from 0.20 (2021) to 1.00 (2023, 2024). In 2021 in Veneto, in 75% of provinces (3 out of 4) an infected mosquito was recorded but no human cases were reported, as seen earlier in the examined Greek regions. Detailed information on the model's performance in the Veneto region is presented in the Supplementary Material (Table S2, Figure S2).

Discussion

The aim of this modelling work was to assess the forecast skill and its spatiotemporal reliability of our previously developed climate-dependent spatial epidemiological model MIMESIS. The model was applied for the period 2020–2024 in the municipalities of the regions of Central Macedonia and Thessaly, in Greece. It describes the transmission cycle of the WNV from birds and mosquitoes to humans and includes fixed parameters, parameters with geographical dependence, and biological parameters with seasonal and climatic dependence. The key challenges to develop and fit the model in forecasting mode were the unknown meteorological and epidemiological conditions. To address the meteorological uncertainty, we powered the model by temperature data from 51 ensemble members of seasonal forecasts released monthly by the ECMWF. To cope with the unknown initial epidemiological conditions, the model incorporates three different scenarios describing the initial population of infected mosquitoes, a parameter determined by the overwintering of the WNV, which varies spatially and annually. One epidemiological scenario was based on the epidemiological situation of each municipality in the previous year, and further two scenarios were based on Analog Ensemble methods using infections surveillance data and climatic indices. Ultimately, the coupled-model (mechanistic and data-driven) predictions were derived by combining the ensemble of simulations in an uncertainty probabilistic framework. By integrating mechanistic and data-driven models, the forecasting framework benefits from the robustness of mechanistic principles and the adaptability of data-driven approaches, resulting in more accurate, generalizable predictions.

According to the surveillance data, if we assumed that at least one human case would be reported in the areas where infected mosquito samples was found, on average, we would have anticipated by a couple of weeks only approximately 40% of the municipalities with reported infection in RCM and RTH, setting the average threshold of spillover infection to humans. Furthermore, mosquito control results were usually available from July, so it is not possible to produce forecasts earlier based on this assumption. Especially in 2021, infected mosquitoes were found in several municipalities, where no human cases were reported. The pattern of lack of human cases reported in municipalities where infected mosquitoes were found in the same year could potentially be explained by a not very successful enzootic cycle resulting in a weak spillover to humans or an under-reporting of human cases. Furthermore, in 2021 Greece reported the highest number of deaths from covid-19 (15,823 deaths[32]), compared to other years. The elderly and immunocompromised people were at higher risk of both death from covid-19 and of severe disease or death from WNV infection. Therefore, the absence of WNV human cases that year might be due to population screening mainly for Corona Virus Disease 2019 (COVID-19) and the masking of the WNF cases by the significant mortality due to covid-19.

Our results were presented at high-resolution (municipal) level for RCM and RTH. In the first annual forecast issued in April, the model correctly identified 71% of the municipalities reporting human infections of RCM and 60% of those in RTH (POD). The inclusion of the false alarms reduced the skill scores to 56% and 42% for RCM and RTH, respectively (CSI). Even so, the skill scores provide an improvement over persistence thanks to the assessment of overwintering in the epidemiological model via the AnEn module. This considerable variation in the proportion of municipalities reporting infected humans between the two regions may be due to the better-quantity data on the circulation of the virus in RCM compared to that in RTH. RCM however objectively reports more human cases of WNV on average, and significantly more municipalities reported infections there during the period 2010–2024. After assimilating the mosquito positivity data of the current season around July, the number of highlighted municipalities changed, the indices improved, and their values were typically higher than the corresponding ones from any single epidemiological scenario. The forecast effectiveness increased by more than 20%, demonstrating the critical role and importance of the epidemiological uncertainty over the climatic uncertainty, suggesting the need for more intensive sampling for WNV.

For municipalities at risk, the model predicted a range of expected annual WNV human cases. The range of cases from the forecast of June was found to most accurately represent reality, as (in all years except 2021) the observed cases fell within the modelled prediction range. The fMIMESIS skill improvement after July did not lead however to a dramatically better estimation of the number of human WNV cases due to the qualitative nature of the virus circulation provided. This implies that the model was well calibrated and could produce reliable forecasts of the risk even without the use of field data from the current season. Obviously a more sophisticated filter for WNVhc could be utilized, considering the temporal evolution of the uncertainties at each month. Moreover, the results demonstrated the need for systematic analyses of the field data to infer true positivity estimations.The predicted number of annual cases mirrors the observed pattern, indicating that this model was valuable as a WNV incidence prediction tool. In addition, the model pinpointed the parts of the province with the highest expected number of WNV human cases.

In forecasting the municipalities at risk and the number of humans affected by WNF, our model outperforms considerably potential approaches based on oversimplified assumptions, i.e. relying exclusively on historical data on local numbers of infection of WNV in humans and mosquitos or such based only on historical weather data. We demonstrated that our model could identify with high accuracy the municipalities, which are most likely to report human cases in both regions covered by our study. Notably our approach can pinpoint these high-risk municipalities as early as April, before, or at the very start of the WNV season, and then perform this risk forecasting 'live', as the season unfolds. This is a crucial feature of our model forecasting, as this early warning gives sufficient head time to municipalities identified as high-risk to take preventive steps to mitigate the risk of WNV case on their territory.

Besides identifying high risk municipalities, our model further reliably estimates a range of expected numbers of WNV case in the municipalities identified to be of high risk. This output of our model provides a two-fold benefit. The number of expected cases in a municipality is a further metric for the degree of risk of WNV at municipality level, promoting informed decisions on the prioritization of public spending and allocation of resources to the prevention of WNF cases. These accurate projections further allow for *a posteriori* evaluation of the effectiveness of public health measures to prevent WNF cases at the end of each season, by comparing the number of cases projected by the model to the number of cases actually reported in municipalities where preventive measures were introduced.

Our findings align with prior modeling efforts aimed at forecasting WNV dynamics. Mechanistic models have demonstrated the critical role of temperature in WNV transmission and the advantages of integrating data assimilation to enhance forecast accuracy[33,34]. Our climate-driven model advances these approaches by incorporating seasonal climate forecasts and ensemble methodologies to achieve refined, municipal-scale predictions. Consistent with studies that identified key parameters influencing WNV dynamics[35], our results emphasize the value of field data in improving model precision. Furthermore, data-driven models linking mosquito infection patterns to weather variables[11,16]corroborate our findings that temperature-driven forecasts substantially enhance predictive skill. Compared to simple statistical models[13,18], our approach offers higher spatial resolution and actionable insights, facilitating targeted interventions and bolstering public health preparedness. Prevention measures can be orders of magnitude more cost effective than response, mitigation and control measures[36]. Unlike other risk assessment approaches[16,23,37,38], which are *early detection models*, i.e. integrate surveillance data to assign risk level to an area, our model is *early-warning model*, in that it estimates the level of risk in advance and provides a forecast, and therefore an opportunity for preparedness and prevention measure. Furthermore, the fine forecasting resolution of our model, at the municipality level, in comparison to other studies, suggests more precisely targeted prevention measures, rendering them more efficient and at the same time more cost effective.

The climate-dependent epidemiological model presented here at municipal level can be applied as a forecasting tool for the risk of WNV infection in humans. It is the only predictive tool at this scale that combines a dynamic model with Analog Ensemble techniques, while also using field data to improve the model's predictive power. Based on the predictions of this model, the municipal authorities will be able to organize campaigns to inform residents of municipalities at risk about mosquito protection measures, and could carry out spraying in potential mosquito breeding grounds (marshes, canals, etc.). Limitations to our modelling approach mostly stem from uncertainties or biases resulting from data collection or validity. These include the variable and often unknown time lag between mosquito sampling and the reporting of results, the unknown number of actual cases as opposed to reported cases, as well as the impact of intervention efforts for the control of WNV circulation in specific regions and municipalities with a history of WNV, which distorts the true epidemiological relationships in our study system. Improving the quantity and consistency of surveillance data for humans, mosquitoes, and sentinel chickens will improve the accuracy of our model's predictions. In our study we did not consider some environmental factors, such as relative humidity, soil moisture and wind speed, and some socio-economic factors, like the age or income of the residents, which are related to the risk of WNV infection or disease. To improve the accuracy of forecasting, future research efforts may include studying potential control strategies via elimination of larvae and adult mosquitoes, as well as understanding human and bird population movements and changes. The inclusion of additional field data updating more frequently (i.e. sub monthly) could also result in a reduction of the epidemiological uncertainty in the period when the number of human cases grow (e.g. July) impacting positively the forecast skill. Our model could be geographically extended by inputting testing data from other municipalities of Greece or in other countries with available entomological and epidemiological data series.

Conclusion

The fMIMESIS system presents a novel approach for forecasting West Nile Virus (WNV) risk at the municipal level, integrating dynamic modelling with Analog Ensemble techniques, seasonal climatic forecasts and field data. This model effectively predicts both the spatial distribution and expected number of WNV human cases, providing early warnings that enable policy makers at the municipal level to implement timely interventions. While the model demonstrates strong forecasting capabilities, future improvements, such as incorporating more frequent field data updates and additional environmental and socio-economic factors, could enhance accuracy. fMIMESIS is a valuable tool for public health management, contributing to more targeted and cost-effective WNV prevention strategies.

Methods

Study area

The study area consists of the RCM and RTH in Greece. This research is conducted at the municipal level. The RCM is divided into 38 municipalities, while the RTH into 25 municipalities. Data from the period 2010–2019 are included for training the models while the period 2020–2024 is used for evaluating their forecast skill.

Data sources

Demographic data

Data on the local administrative units (LAU) in Greece, the permanent population, and the area (km^2) of municipalities were obtained from the corresponding National Statistical Authority. Geospatial coordinates of each municipality were taken from Google maps.

Climate data

At the end of each month the seasonal forecasts of air temperature (2 m-height) were obtained from the European Centre for Medium-Range Weather Forecast (ECMWF) for four hours each day for a seven-month time period, considering 51 ensemble members. From these data, the daily average temperature of each day was estimated. To calculate monthly Consecutive Wet Days (CWD), daily precipitation data for the same areas and period were also obtained from ECMWF. The CWD index was the largest number of consecutive "wet days" (where precipitation > 1.5 mm) during the period of interest. The threshold value was found after comparing the total precipitation data from the 51 ensemble members with the corresponding data from ERA 5 (fifth generation ECMWF atmospheric reanalysis of the global climate). ECMWF updates the seasonal climate forecasts every month, hence we also replace the climate data monthly.

WNVhc

Data of the reported WNVhc, including individuals with and without symptoms in the CNS, in Greece for the period 2020–2024 were obtained from the National Public Health Organisation (EODY) at the municipal level. In this work, we calibrated our model towards reported cases (i.e. detected), even though seroprevalence studies in Greece indicate that the number of infections may be two orders of magnitude higher[39]. The geographical distribution of the WNVhc and their variability during the forecast period 2020–2024 are presented in Fig. 5. Detailed information on the annual WNVhc is shown in the Supplementary Material (Figures S3- S4, Tables S3-S4).

Mosquito surveillance

Mosquito samples were collected every two weeks in selected human settlements in RCM and RTH using CO_2 and light baited mosquito traps for the period 2019–2024. The selected locations were protected from direct sunlight and wind, either near a settlement or inside a house yard[40]. A selection of samples were tested for WNV infection at the laboratory[41]. For the species identification dichotomous identification keys were used[42]. Detection of mosquitoes infected with WNV was performed by an RT-nested PCR in reference laboratories.

The spatial analysis of entomological data was conducted bimonthly at the settlement level, while this research was carried out at the municipal level. A pre-processing of entomological data was necessary to transform the bimonthly settlement-based data to monthly municipality-based numbers of Culex mosquitoes per sample per month per municipality. Only qualitative infection data was used due to the absence of data on pool size and fraction of pools tested.

WNV infections in chickens may be another useful indicator of the virus circulation in an area[43], however due to the sparse spatial data in the examined area it was not considered.

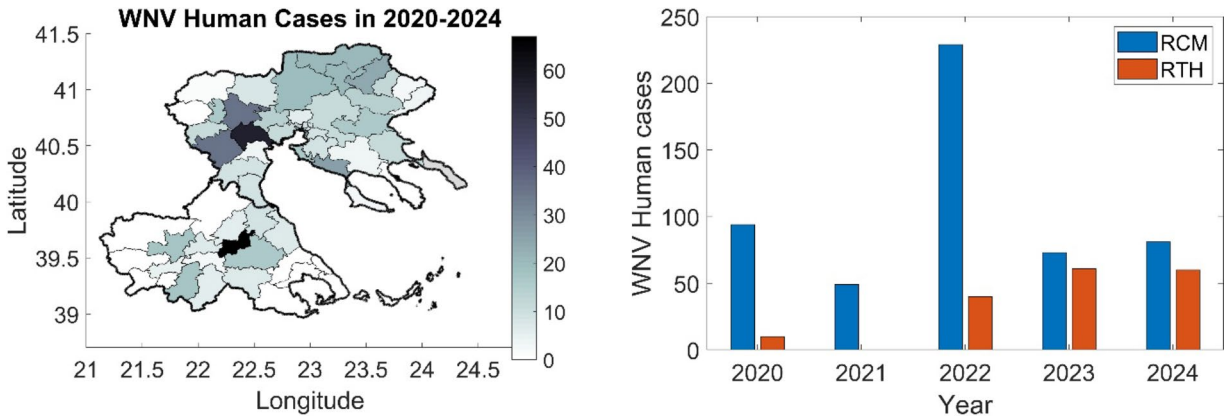

Fig. 5. WNVhc in the investigated regions during 2020–2024. White colour denotes municipalities where no human cases were observed (left). Multiannual variability of the WNVhc in the investigated regions (right). The map has been created using the Mapping Toolbox of MATLAB R2024a (Natick, Massachusetts, United States: The MathWorks Inc, 2024, https://www.mathworks.com).

Epidemiological model

The MIMESIS-2 (spatial dynaMIcal Model for wESt nIle viruS) is a dynamical model which simulates the transmission cycle of WNV between mosquitoes, birds, and humans[31]. The model segregates the populations into 14 health or epidemiological states. The basic health states of the populations are the susceptible, exposed, infected, recovered and dead individuals. The interface between the 14 compartments is governed by several parameters dependent on climate, demography, geography or season. Mathematically, the model is described by 14 differential equations solved numerically with a time-step of one day. Spatially, the study was conducted at the municipal level. All the required input data (demographical, geographical, environmental epidemiological, entomological) was formatted at this scale.

Data-driven model

Data-driven modelling is rapidly being applied across various fields, including medicine, powering innovations in areas such as medical imaging, disease diagnosis, and personalized treatment plans. Analog Ensemble (AnEn) is a statistical post-processing method that utilizes forecasts in tandem with historical predictions and observations to (a) identify a number of past states (analogues) that closely resemble the present state and (b) utilize them to generate the AnEn forecast. This technique was proposed by Delle Monache et al.[44]and has been successfully applied across diverse fields including meteorology[45], renewable energy[46], and air pollution[47,48]. The optimal number of analog members for the AnEn forecast and the most suitable combination of auxiliary variables are determined during the training period 2010–2019 by adopting an error minimization approach. Multiple epidemiological and meteorological variables were tested, with the identified predictors ultimately including monthly temperature, monthly CWD and the number of infected humans in the previous year. The AnEn prediction for each year in 2020–2024 and for each municipality utilizes the intersection of optimal combinations across all years in the training period. Given the ensemble seasonal weather forecasts, a statistical distribution is hence provided for initial infected mosquito population ($IM_{0)}$ at each municipality every month.

Forecasting methodology

MIMESIS-2 can operate in forecast mode to provide predictions of key WNV-related parameters, such as, infected mosquitoes, infected birds, infected humans given appropriate initial conditions and climate projections. Regarding climate, we fed the model with all 51 ensemble members of 7-month ahead seasonal forecasts released each month from ECMWF. In terms of the unknown initial epidemiological conditions (i.e., the initial number of infected mosquitoes, IM_0), we developed a module per se which employs the AnEn algorithm, a technique based on weighted analogue past forecasts and observations.

Overall, the following epidemiological scenarios were investigated:

- S1: $IM_0$is equal to the initial number of infected mosquitoes in the last year, as found by calibrating the MIMESIS-2 in the historic period from 2010 up to the calendar year before the forecast is generated. Due to the endemicity of WNV in Greece[49], the hypothesis behind this scenario is persistence, i.e. WNVhc will probably occur in municipalities where human cases were recorded in the previous year.
- S2: IM_0 was calculated from the mean value of AnEn forecasts falling within the interquartile range. Avoiding extreme percentile values, this scenario was considered a likely reasonable assumption of the number of initial infected mosquitoes given the historic epidemiological and environmental conditions and the projected seasonal forecasts.
- S3: IM_0 equals the value of the 95th percentile in the AnEn forecast. This scenario will highlight the potential for extreme values at the different municipalities arising from the environmental and epidemiological conditions.

In this work, first we calibrated the mechanistic model (MIMESIS2) and the data-driven model (AnEn) over a historic period extending from 2010 to the year ending before each forecast year. Next, we utilised both models with the seasonal climate forecasts to generate the epidemiological scenarios. Then, we run MIMESIS2 under different climatic and epidemiological conditions and we combined the ensemble of the 153 individual forecasts (51 seasonal X 3 epidemiological) in a probabilistic framework to generate the fMIMESIS forecast, which started in April and was updated on a monthly basis until September, providing each time predictions which covered the entire WNV season. The contribution from S1 is proportional to the measure of the endemicity of the virus in each municipality, while the contribution from S2 and S3 were given equal weights. Field data were also assimilated if/when they became available (typically after July). We used during the period April-June, the entomological data of the previous year, while in forecasts issued from July onwards the epidemiological data of the same year were used. Specifically, when an infected mosquito was found in a municipality, IM0 in S2 and S3 increased by half a class (20 mosquitoes). Figure 6 describes the conceptual framework followed to produce the model predictions.

The probability of occurrence of a human case is equal to the proportion of the ensemble members in which at least one human case was predicted. To use the MIMESIS-2 as a risk assessment tool, we defined five possible levels of risk (level 1 – level 5) for transmission of WNV to humans (Table 2). Depending on the identified risk level for an area, additional surveillance activities can be implemented, as well as further preparedness, response or control measures. All the modelled skill scores presented correspond to the identified municipalities as 'High risk level' (RL4 and RL5).

Model verification measures

The performance of the examined model for capturing events (e.g. occurrence or not of WNVhc in a specific municipality and year) is assessed with the verification indices (POD, CSI). Specifically:

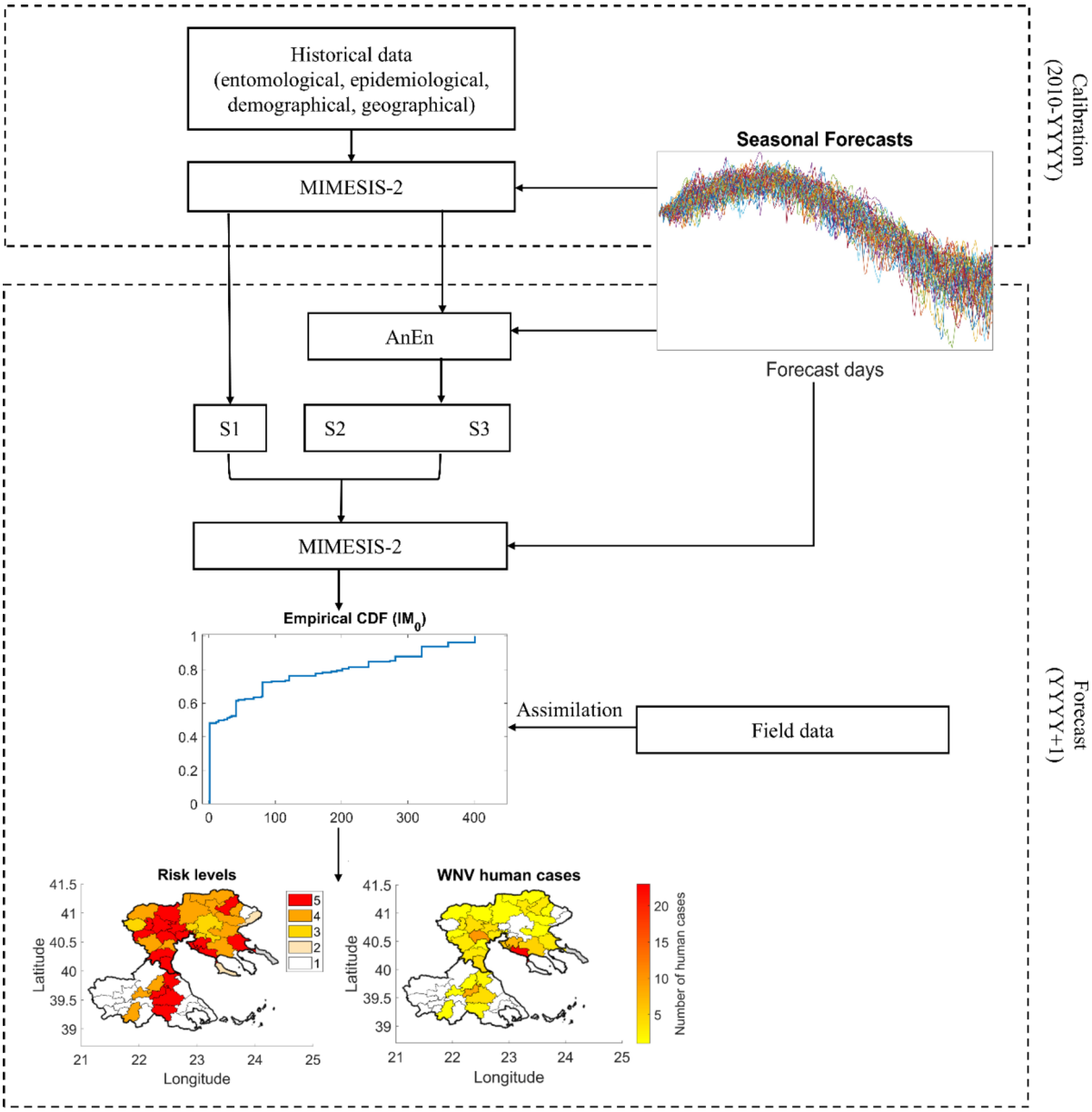

Fig. 6. Conceptual methodology framework to generate seasonal forecasts of WNV risk with fMIMESIS. The map has been created using the Mapping Toolbox of MATLAB R2024a (Natick, Massachusetts, United States: The MathWorks Inc, 2024, https://www.mathworks.com).

Probability of WNVhc occurring	Risk Level (RL)	Characterization
1–20	1	Low
21–40	2	Moderate
41–60	3	
61–80	4	High
81–100	5	

Table 2. Seasonal risk levels of WNV transmission to humans with the corresponding characterization.

The Probability of Detection (POD) is calculated as $POD = a/(a + c)$, while Critical Success Index (CSI) is estimated as $CSI = a/(a + b + c)$,where a, b, c symbolizes the true positives, false positives and false negatives respectively. POD measures the ability of a model to correctly identify all actual human cases, indicating the proportion of true positives among all actual cases. CSI provides a single number that reflects how

well the model predicts the human cases, balancing the importance of different types of errors (false alarms and misses). The multiannual value of POD (identically for CSI) is calculated as follows:

$$POD_{2020-2024} = \frac{\sum_{i=2020}^{2024} a_i}{\sum_{i=2020}^{2024} a_i + \sum_{i=2020}^{2024} c_i}$$

Data availability

The data on mosquito abundance in traps are available from EcoDevelopmet S.A. but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of EcoDevelopmet S.A. The other data used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 30 July 2024; Accepted: 24 February 2025
Published online: 28 February 2025

References

1. Benjelloun, A., El Harrak, M. & Belkadi, B. West nile disease epidemiology in North-West Africa: bibliographical review. *Transbound. Emerg. Dis.* **63**, e153–e159. https://doi.org/10.1111/tbed.12341 (2016).
2. Selim, A. & Abdelhady, A. The first detection of anti-West nile virus antibody in domestic ruminants in Egypt. *Trop. Anim. Health Prod.* **52**, 3147–3151. https://doi.org/10.1007/s11250-020-02339-x (2020).
3. Bunning, M. L. et al. Experimental infection of horses with West nile virus and their potential to infect mosquitoes and serve as amplifying hosts. *Ann. N Y Acad. Sci.* **951**, 338–339. https://doi.org/10.3201/eid0804.010239 (2001).
4. Calistri, P. et al. Epidemiology of West nile in Europe and in the mediterranean basin. *Open. Virol. J.* **4**, 29–37. https://doi.org/10.2174/1874357901004010029 (2010).
5. Petersen, L., Brault, A. & Nasci, R. West nile virus: review of the literature. *JAMA* **310**, 308–315. https://doi.org/10.1001/jama.2013.8042 (2013).
6. World Health Organization & Virus, W. N. Available online: https://www.who.int/en/news-room/fact-sheets/detail/west-nile-virus (accessed on March 2024).
7. Angelou, A., Schuh, L., Stilianakis, N. I., Mourelatos, S. & Kioutsioukis, I. Unveiling Spatial patterns of West nile virus emergence in Northern Greece, 2010–2023. *One Health (Amsterdam Netherlands)*. **19**, 100888. https://doi.org/10.1016/j.onehlt.2024.100888 (2024).
8. National Public Health Organisation (EODY). Annual reports about West Nile Virus (WNV) surveillance, Available online, (2024). https://eody.gov.gr/en/epidemiological-statistical-data/annual-epidemiological-data/, (accessed on October 2024).
9. Parselia, E. et al. Satellite Earth observation data in epidemiological modeling of malaria, dengue and West nile virus: a scoping review. *Remote Sens.* **11**, 1862. https://doi.org/10.3390/rs11161862 (2019).
10. Barker, C. M. Models and surveillance systems to detect and predict West nile virus outbreaks. *J. Med. Entomol.* **56** (6), 1508–1515. https://doi.org/10.1093/jme/tjz150 (2019).
11. Ruiz, M. O. et al. Local impact of temperature and precipitation on West nile virus infection in Culex species mosquitoes in Northeast Illinois. *USA Parasites Vectors*. **3**, 19. https://doi.org/10.1186/1756-3305-3-19 (2010).
12. Chuang, T. W. & Wimberly, M. C. Remote sensing of Climatic anomalies and West nile virus incidence in the Northern great plains of the united States. *PLoS One*. **7**, e46882. https://doi.org/10.1371/journal.pone.0046882 (2012).
13. Manore, C. A. et al. Towards an early warning system for forecasting human West nile virus incidence. *PLoS Curr.* https://doi.org/10.1371/currents.outbreaks.f0b3978230599a56830ce30cb9ce0500 (2014).
14. Little, E., Campbell, S. R., Shaman, J. & County, S. Development and validation of a climate-based ensemble prediction model for West Nile Virus infection rates in Culex mosquitoes, New York. Parasit. Vectors. 9: 443. (2016). https://doi.org/10.1186/s13071-016-1720-1
15. Shand, L. et al. Predicting West nile virus infection risk from the synergistic effects of rainfall and temperature. *J. Med. Entomol.* **53**, 935–944. https://doi.org/10.1093/jme/tjw042 (2016).
16. Davis, J. K. et al. Integrating environmental monitoring and mosquito surveillance to predict vector-borne disease: prospective forecasts of a West nile virus outbreak. *PLoS Curr.* https://doi.org/10.1371/currents.outbreaks.90e80717c4e67e1a830f17feeaaf85de (2017).
17. Karki, S., Westcott, N. E., Muturi, E. J., Brown, W. M. & Ruiz, M. O. Assessing human risk of illness with West nile virus mosquito surveillance data to improve public health preparedness. *Zoonoses Public. Health*. **65**, 177–184. https://doi.org/10.1111/zph.12386 (2018).
18. Pepper, S. T., Dawson, D. E. & Dacko, N. Predictive modeling for West nile virus and mosquito surveillance in Lubbock. *Tex. J. Am. Mosq. Control Assoc.* **34**, 18–24. https://doi.org/10.2987/17-6714.1 (2018).
19. Farooq, Z. et al. Artificial intelligence to predict West nile virus outbreaks with eco-climatic drivers. *Lancet Reg. Health Eur.* **17**, 100370. https://doi.org/10.1016/j.lanepe.2022.100370 (2022). PMID: 35373173; PMCID: PMC8971633.
20. Ward, M. J. et al. A spatially resolved and environmentally informed forecast model of West Nile virus in Coachella Valley, California. GeoHealth,7, e2023GH000855. (2023). https://doi.org/10.1029/2023GH000855
21. McKenzie, V. J. & Goulet, N. E. Bird community composition linked to human West nile virus cases along the Colorado front range. *Ecohealth* **7**, 439–447. https://doi.org/10.1007/s10393-010-0360-8 (2010).
22. DeFelice, N. B., Little, E., Campbell, S. R. & Shaman, J. Ensemble forecast of human West nile virus cases and mosquito infection rates. *Nat. Commun.* **8**, 14592. https://doi.org/10.1038/ncomms14592 (2017).
23. Davis, J. K. et al. Improving the prediction of arbovirus outbreaks: a comparison of climate-driven models for West nile virus in an endemic region of the united States. *Acta Trop.* **185**, 242–250. https://doi.org/10.1016/j.actatropica.2018.04.028 (2018).
24. Bai, Z. & Zhang, Z. Dynamics of a periodic West nile virus model with mosquito demographics. *Commun. Pure Appl. Anal.* **21**, 3755–3775. https://doi.org/10.3934/cpaa.2022121 (2022).
25. Ferraguti, M., Dimas Martins, A. & Artzy-Randrup, Y. Quantifying the invasion risk of West nile virus: insights from a multi-vector and multi-host SEIR model. *One Health*. **17**, 100638. https://doi.org/10.1016/j.onehlt.2023.100638 (2023).
26. Marini, G. et al. West nile virus transmission and human infection risk in Veneto (Italy): a modelling analysis. *Sci. Rep.* **8**, 14005. https://doi.org/10.1038/s41598-018-32401-6 (2018).
27. Ewing, D. A., Purse, B. V., Cobbold, C. A. & White, S. M. A novel approach for predicting risk of vector-borne disease establishment in marginal temperate environments under climate change: West nile virus in the UK. *J. Royal Soc. Interface*. **18**, 20210049. https://doi.org/10.1098/rsif.2021.0049 (2021).

28. Holcomb, K. M. et al. Predicted reduction in transmission from deployment of ivermectin-treated birdfeeders for local control of West nile virus. *Epidemics* **44**, 100697. https://doi.org/10.1016/j.epidem.2023.100697 (2023).
29. Angelou, A., Kioutsioukis, I. & Stilianakis, N. A climate-dependent spatial epidemiological model for the transmission risk of West Nile virus at local scale, One Health, 13: 100330, 2021, (2021). https://doi.org/10.1016/j.onehlt.2021.100330
30. Kioutsioukis, I. & Stilianakis, N. I. Assessment of West nile virus transmission risk from a weather-dependent epidemiological model and a global sensitivity analysis framework. *Acta Trop.* **193**, 129–141. https://doi.org/10.1016/j. actatropica.2019.03.003 (2019).
31. Fasano, A. et al. Stilianakis N. I. An epidemiological model for mosquito host selection and temperature-dependent transmission of West nile virus. *Sci. Rep.* **12**, 19946. https://doi.org/10.1038/s41598-022-24527-5 (2022).
32. World Health Organization. Daily COVID-19 cases and deaths by date reported to WHO, Available online: https://data.who.int/dashboards/covid19/data?n=c&m49=300 (accessed on March 2024).
33. DeFelice, N. B. et al. Use of temperature to improve West nile virus forecasts. *Plos Comput. Biol.* **14**, e1006047. https://doi.org/10.1371/journal.pcbi.1006047 (2018).
34. Bhowmick, S., Gethmann, J., Conraths, F. J., Sokolov, I. M. & Lentz, H. H. K. Locally temperature-driven mathematical model of West nile virus spread in Germany. *J. Theor. Biol.* **488**, 110117. https://doi.org/10.1016/j.jtbi.2019.110117 (2020).
35. Fesce, E. et al. Understanding West nile virus transmission: mathematical modelling to quantify the most critical parameters to predict infection dynamics. *PLoS Negl. Trop. Dis.* **17**, e0010252. https://doi.org/10.1371/journal.pntd.0010252 (2023).
36. Bernstein, A. S. et al. The costs and benefits of primary prevention of zoonotic pandemics. *Sci. Adv.* **8** (5), eabl4183. https://doi.org/10.1126/sciadv.abl4183 (2022).
37. Keyel, A. C. et al. Seasonal temperatures and hydrological conditions improve the prediction of West nile virus infection rates in Culex mosquitoes and human case counts in new York and Connecticut. *PloS One*. **14** (6), e0217854. https://doi.org/10.1371/journal.pone.0217854 (2019).
38. Wimberly, M. C., Davis, J. K., Hildreth, M. B. & Clayton, J. L. Integrated Forecasts Based on Public Health Surveillance and Meteorological Data Predict West Nile Virus in a High-Risk Region of North America. *Environ. Health Perspect.*, **130**(8), 87006. https://doi.org/10.1289/EHP10287 (2022).
39. Hadjichristodoulou, C. et al. West nile virus Seroprevalence in the Greek population in 2013: a nationwide cross-sectional survey. *PLoS One*. **10** (11), e0143803. https://doi.org/10.1371/journal.pone.0143803 (2015).
40. Tsioka, K. et al. Detection and molecular characterization of West nile virus in Culex pipiens mosquitoes in central Macedonia, Greece, 2019–2021. *Acta Trop.* **230**, 106391. https://doi.org/10.1016/j.actatropica.2022.106391 (2022).
41. Tsioka, K. et al. West nile virus in Culex mosquitoes in central Macedonia, Greece, 2022. *Viruses* **15**, 224. https://doi.org/10.3390/v15010224 (2023).
42. Becker, N. et al. Mosquitoes: Identification, Ecology and Control. 3rd edition.Cham, Switzerland: Springer Nature, p. 407. (2020). https://doi.org/10.1007/978-3-030-11623-1
43. Chaintoutis, S. C. et al. Surveillance and early warning of West nile virus lineage 2 using backyard chickens and correlation to human neuroinvasive cases. *Zoonoses Public. Health.* **62** (5), 344–355. https://doi.org/10.1111/zph.12152 (2015).
44. Delle Monache, L., Eckel, F. A., Rife, D. L., Nagarajan, B. & Searight, K. Probabilistic weather prediction with an analog ensemble. *Mon. Weather Rev.* **141**, 3498–3516. https://doi.org/10.1175/MWR-D-12-00281.1 (2013).
45. Sperati, S., Alessandrini, S. & Delle Monache, L. Gridded probabilistic weather forecasts with an analog ensemble. *Q. J. R Meteorol. Soc.* **143**, 2874–2885. https://doi.org/10.1002/qj.3137 (2017).
46. Pappa, A., Theodoropoulos, I., Galmarini, S. & Kioutsioukis, I. Analog versus multi-model ensemble forecasting: A comparison for renewable energy resources. *Renew. Energy*. **205**, 563–573. https://doi.org/10.1016/j.renene.2023.01.030 (2023).
47. Pappa, A. & Kioutsioukis, I. Forecasting particulate pollution in an urban area: from copernicus to Sub-Km scale. *Atmosphere* **12** (7), 881. https://doi.org/10.3390/atmos12070881 (2021).
48. Solomou, E. S. et al. Analog ensemble technique to post-process WRF-CAMx Ozone and particulate matter forecasts. *Atmos. Environ.* **256**, 118439. https://doi.org/10.1016/j.atmosenv.2021.118439 (2021).
49. Kouroupis, D., Charisi, K. & Pyrpasopoulou, A. The ongoing epidemic of West nile virus in Greece: the contribution of biological vectors and reservoirs and the importance of climate and socioeconomic factors revisited. *Trop. Med. Infect. Disease*. **8** (9), 453. https://doi.org/10.3390/tropicalmed8090453 (2023).

Acknowledgements

We thank EcoDevelopmet S.A. a private company implementing mosquito control projects in Greece for providing the Culex data.

Disclaimer

The views expressed in this article are purely those of the authors and may not, under any circumstances, be regarded as an official position of the European Commission.

Author contributions

Conceptualization: NIS, IKData curation: AA, AP, SGFormal analysis: AAFunding Acquisition: IKInvestigation: AA, APMethodology: AA, PVM, NIS, IKProject administration: IKSoftware: AA, AP, IKSupervision: IKValidation: AA, AP, IKVisualization: AAWriting – original draft: AAWriting – review & editing: PVM, NIS, SG, IK. All authors have approved the submitted version and have agreed to be personally accountable for authors' own contributions and to ensure that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

This work has been financed from the EIC 1st Horizon Prize "Early Warning for Epidemics".

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/1

0.1038/s41598-025-91996-9.

Correspondence and requests for materials should be addressed to I.K.