West Nile virus forecast model submission form
Email completed form to vbd-predict@cdc.gov

Team name
Team Jornada

Team leader

Name	Institution	Email
Dr. Debra P. C. Peters	Agricultural Research Service, USDA	deb.peters@usda.gov

Other team members

Name	Institution	Email
John M. Humphreys	Agricultural Research Service, USDA	John.humphreys@usda.gov
Katie Young	Agricultural Research Service, USDA	kiy761@nmsu.edu
Lee Cohnstaedt		lee.cohnstaedt@usda.gov
Kathryn A. Hanley	New Mexico State University	khanley@nmsu.edu
Luis Rodriguez		

Model description
Provide a brief summary of the model methods with sufficient detail for another modeler to understand the approach being applied. If multiple models are used, describe each model and how they were combined.

We applied Bayesian hierarchical modeling to develop a two-part, joint spatiotemporal model (conditional autoregressive model) that concurrently estimated both neuroinvasive and non-neuroinvasive disease counts for each year and county over the period 2000-2018. As a joint model, two response/dependent variables were included; the first-level response variable consisted of non-neuroinvasive disease counts (data previously acquired) and the second-level response included the neuroinvasive counts provided in conjunction with the WNV Challenge competition. Because the model has two levels, estimated non-neuroinvasive counts were passed to the model's second level, such that they were used as predictors/covariates during neuroinvasive count estimation. In addition to the independent variables listed in the "Variables" section of this form, the model was designed to include two, separate spatial effects to control for spatial bias/error within each model level, as well as a third "shared" spatial effect to account for the spatial relationship (spatial correlation) between neuroinvasive and non-neuroinvasive counts. The reported case year was used to model a non-linear time trend for the period of record and a space-time interaction term (Year*County) was added to quantify the relationship between the spatial and temporal trends. Once the model's spatial and temporal effects were parametrized, expert judgement and iterative model fitting were used to identify potential independent variables. Following validation, forecasting was performed to estimate 2019 non-neuroinvasive and neuroinvasive case counts. The forecasts

for 2019 were then added to the original response variable (2000-2018) to undertake prediction for the target year 2020.

Variables
List each variable used and its temporal relationship to the forecast. If multiple models are used, specify which enter into each model.

1. Median Household Income (county and year specific)
2. Force of Infection (county and year specific estimate of past incidence)
3. Proportion of population over 54 years of age (county and year specific)
4. Geographic area of county (km^2, county specific)
5. Competent avian host richness (competent bird species reported in literature, county specific)
6. Average maximum temperature (summer season, May-August, spatial and year variation)
7. Average total precipitation (summer season, May-August, spatial and year variation)
8. Count of Mosquito detections (reported to ArboNet, county and year variation)
9. Count of Avian detections (reported to ArboNet, county and year variation)

Computational resources
Describe the programming languages and software tools that were used to write and execute the forecasts.

Data preprocessing and model coding were performed using the R-language for statistical programming. Due to high model dimensionally, the USDA Agricultural Research Service's Ceres High Performance Computer was used for model execution.

Publications
Note whether the model was derived from previously published work and, if so, provide references.

Associated manuscript currently being drafted.

Participation agreement
By submitting these forecasts, the team agrees to abide by the project rules and data use agreements.

Team lead name	Date