

## Spatio-temporal Modelling of H5N1 Avian Influenza Outbreaks in Europe (2021–2024)

Mehak Jindal<sup>1,2</sup> Samsung Lim<sup>1,2</sup> C Raina MacIntyre<sup>2,3</sup>

<sup>1</sup> School of Civil and Environmental Engineering, University of New South Wales Sydney, New South Wales 2052, Australia

<sup>2</sup> Biosecurity Program, Kirby Institute, University of New South Wales Sydney, New South Wales 2052, Australia

<sup>3</sup> College of Health Solutions & College of Public Service & Community Solutions, Arizona State University, Tempe, United States.

m.jindal@unsw.edu.au, s.lim@unsw.edu.au, r.macintyre@unsw.edu.au

**Keywords:** Spatial Epidemiology, Spatio-temporal Modelling, Geospatial Health, Endemic–epidemic Model, Disease Forecasting, Avian Influenza

### Abstract

Highly Pathogenic Avian Influenza (HPAI), particularly the H5N1 strain, poses a significant ongoing threat to animal health, biodiversity and food security across Europe. Understanding where and when avian influenza risks intensify is essential for targeted surveillance and rapid response. This study develops a data-driven spatio-temporal framework that integrates geospatial, ecological and climatic datasets to explain and forecast the dynamics of H5N1 outbreaks between 2021 and 2024. Weekly country-level outbreak counts (208 weeks, 37 countries) were analysed using a hierarchical endemic-epidemic model with an assumption of Negative Binomial distribution. Environmental covariates, bird-species densities, and human population metrics were incorporated into endemic and autoregressive components. Model performance was evaluated using rolling one-step-ahead forecasts assessed by proper scoring rules (logarithmic score and ranked probability score) and calibration diagnostics. The proposed framework substantially outperformed a regression-only Negative Binomial baseline, reducing mean logS by approximately 29% and RPS by 49%, while exhibiting improved probabilistic calibration. Results indicate that H5N1 transmission is structured by ecological drivers and local persistence mechanisms rather than purely seasonal effects. Anseriformes, Charadriiformes and Pelecaniformes densities were identified as the key migratory bird families contributing to the viral spread. The endemic-epidemic model achieved high forecast accuracy, with majority of the observed weekly outbreak counts falling within central predictive intervals (RPS = 0.76, logS = 0.61). Overall, the proposed framework provides a scalable approach for integrating ecological and spatial information into early-warning systems for HPAI surveillance.

### 1. Introduction

Highly Pathogenic Avian Influenza (HPAI) continues to circulate across Europe, causing repeated epizootic waves that disrupt poultry production and wildlife in Europe (European Food Safety Authority (EFSA) et al., 2024). Since 2021, H5N1 subtype has spread rapidly among both domestic flocks and wildlife species, with increasing detection in mammals. The intensification and spread of outbreaks across Europe highlight the need for an effective early-warning system (Jung Kjær et al., 2023).

Most previous HPAI studies can be grouped into three categories. First, surveillance and genomic tracking to record the expansion of the H5 clade 2.3.4.4b and quantify the changing ecology in Europe. This included increasing circulation of the virus in wild birds with repeated spillover to poultry and mammalian infections (Fusaro et al., 2024; Xie et al., 2023). These studies provide valuable insight into viral evolution and transmission pathways but are primarily descriptive and do not directly address short-term outbreak prediction.

Second, a large body of work has employed spatial risk modelling approaches, including regression-based models, ecological niche modelling, and machine learning techniques, to identify environmental and anthropogenic drivers of HPAI occurrence. These studies consistently highlight the importance of wetland habitats, climatic conditions, wild bird community composition, and poultry production intensity as key predictors of outbreak risk (Schreuder et al., 2022; Opata et al., 2025; Hayes et al., 2025). While these approaches are effective in identifying spatial patterns and risk factors, they typically provide static risk maps and do not explicitly capture temporal dynamics or epidemic feedback processes.

Third, an emerging strand of research focuses on early-warning and forecasting systems, which aim to translate surveillance data into actionable predictions for outbreak preparedness. These approaches utilise time-series models and probabilistic forecasting frameworks to predict short-term outbreak dynamics and assess uncertainty, often within a One Health perspective that integrates human, animal, and environmental systems (Gneiting & Raftery, 2007). However, many existing models either rely on limited environmental covariates or do not explicitly incorporate ecological host dynamics and spatial transmission mechanisms.

Despite these advances, there remains a gap in integrating ecological host information, environmental drivers, and spatio-temporal transmission dynamics within a unified probabilistic forecasting framework. In particular, few studies combine large-scale surveillance data with ecological covariates and evaluate predictive performance using rigorous out-of-sample validation and proper scoring rules.

To address this gap, this study develops a spatio-temporal endemic–epidemic modelling framework to analyse weekly H5N1 outbreak counts across Europe between 2021 and 2024. By integrating environmental, ecological, and anthropogenic predictors within a probabilistic modelling framework, the study aims to (i) identify key drivers of outbreak dynamics, (ii) quantify the relative contributions of local persistence and spatial transmission, and (iii) evaluate predictive performance using rolling one-step-ahead forecasts and proper scoring rules. This approach provides a data-driven and interpretable framework for understanding and forecasting H5N1 outbreaks in support of early-warning systems and cross-border disease preparedness.

## 2. Methodology

### 2.1 Study area and data

The study region comprises 37 European countries represented as areal units in a spatial adjacency matrix. Outbreak data were obtained from the World Organisation for Animal Health – World Animal Health Information System (WOAH-WAHIS) and aggregated to weekly counts per country from January 2021 to December 2024 (208 weeks). Table 1 summarises the datasets used in this study.

To characterise environmental suitability and anthropogenic exposure, we compiled open covariates describing climate, human presence and accessibility, livestock intensity, and ecological bird context. Climate variables were derived from NASA Moderate Resolution Imaging Spectroradiometer Global Land Data Assimilation System (MODIS GLDAS) (0.25° resolution, monthly) and included specific humidity, snow melt, snow depth, total precipitation rate, soil moisture content, air temperature, transpiration and wind speed (Rodell et al., 2004). Anthropogenic proxies included population density from Gridded Population of the World (GPW, Center For International Earth Science Information Network-CIESIN-Columbia University, 2018) and poultry density from Gridded Livestock of the World (GLW; Gilbert et al., 2018). Ecological predictors were based on weekly bird distribution products (eBird Status & Trends), aggregated to 11 bird orders relevant for H5N1 transmission (Fink et al., 2023).

All raster’s were spatially aggregated to national scale and harmonised to weekly time steps to match the outbreak dataset. We log-transformed skewed covariates using  $\log(1+x)$ .

| Category                           | Dataset / Source                     | Spatial / Temporal resolution       | Use in model                      |
|------------------------------------|--------------------------------------|-------------------------------------|-----------------------------------|
| Outbreak surveillance <sup>1</sup> | WOAH–WAHIS H5N1                      | Country-level; weekly (2021–2024)   | Response variable (weekly counts) |
| Climate <sup>2</sup>               | NASA GLDAS (Noah)                    | 0.25°; monthly → weekly             | Temperature / moisture covariates |
| Population <sup>3</sup>            | WorldPop / GPW                       | 100 m–1 km; annual                  | Anthropogenic exposure proxy      |
| Livestock <sup>4</sup>             | Gridded Livestock of the World (GLW) | ~10 km (5 arc-min) or 1 km products | Poultry density proxy             |
| Ecology <sup>5</sup>               | eBird Status & Trends                | Weekly abundance maps               | Waterbird order densities         |

Table 1: Key datasets used in the H5N1 modelling workflow

### 2.2 Methods

The overall view of the steps followed is given in Figure 1.

**2.2.1 Feature selection:** A two-stage screening pipeline was used to reduce the initial set of thirty candidate predictors. These predictors represented environmental, ecological, and anthropogenic conditions hypothesised to influence H5N1 transmission dynamics. First, we employed the Boruta algorithm (Kursa et al., 2010) to evaluate the relative importance of each candidate variable. Boruta is a wrapper method built around the Random Forest classifier that iteratively compares the importance of real predictors against randomly permuted “shadow” variables. Variables that consistently outperform their shadow counterparts are classified as confirmed important, while those performing worse are rejected. This approach allows robust identification of predictors that contribute meaningful explanatory power while accounting for nonlinear relationships and interactions among variables. A second filtering step was applied to address potential multicollinearity among the retained predictors. Pairwise Pearson correlation coefficients were calculated for all remaining variables, and predictors exhibiting strong correlations ( $|r| > 0.8$ ) were examined to avoid redundancy in the model. In cases where highly correlated variables were identified, the variable with lower Boruta importance or weaker ecological interpretability was removed. This procedure ensured that the final predictor set remained statistically stable and ecologically meaningful while reducing the risk of inflated parameter estimates or unstable model coefficients.

**2.2.2 Endemic-Epidemic Model:** The spatio-temporal dynamics of H5N1 detections were modelled using the endemic–epidemic (HHH4) model implemented in the surveillance R package (Meyer et al., 2017). The endemic–epidemic model decomposes disease incidence into three interpretable components that represent different transmission processes: an endemic component which captures the baseline risk, epidemic component which represents the within country persistence of the disease and the neighbourhood component which defined the cross-border transmission from neighbouring regions. Weekly H5N1 detections were modelled using a Negative Binomial distribution to account for overdispersion commonly observed in infectious disease count data. Equation 1 defines the conditional mean of the outbreak process.

$$\mu_{it} = \lambda_{it} y_{i,t-1} + \phi_{it} \sum_{j \neq i} w_{ji} y_{j,t-1} + e_{it} v_{it} \quad (1)$$

where  $i$  and  $j$  = country

$t$  = week

$\mu_{it}$  = expected number of H5N1 cases

$y_{i,t-1}$  = observed number of H5N1 detections

$w_{ji}$  = spatial weight i.e. the adjacency relationship

$e_{it}$  = exposure offset

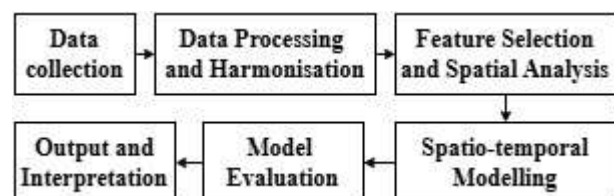


Figure 1: Methodological workflow for the spatio-temporal modelling of H5N1 outbreaks in Europe (2021–2024)

<sup>1</sup> <https://wahis.woah.org/#/event-management>

<sup>2</sup> <https://disc.gsfc.nasa.gov/>

<sup>3</sup> <https://www.earthdata.nasa.gov/data/projects/gpw>

<sup>4</sup> <https://www.fao.org/livestock-systems/global-distributions/en/>

<sup>5</sup> <https://ebird.org/home>

**2.2.3 Baseline benchmark:** To provide a benchmark for evaluating the predictive performance of the endemic–epidemic model, a Negative Binomial Generalised Linear Model (GLM) was implemented as a regression-based baseline. This baseline model incorporates the same covariates and exposure structure as the endemic component of the model but does not include autoregressive or neighbourhood transmission terms. It models outbreak counts as a function of explanatory variables. Equation 2 models the weekly H5N1 detections  $Y_{it}$  for country  $i$  at week  $t$  as

$$Y_{it} \sim \text{NegBin}(\mu_{it}, \theta) \quad (2)$$

where

$i$  = country  
 $t$  = week  
 $Y_{it}$  = observed number of H5N1 cases  
 NegBin = Negative Binomial distribution  
 $\mu_{it}$  = expected number of H5N1 cases  
 $\theta$  = dispersion parameter

**2.2.4 Model evaluation:** we compared a basic intercept-only model, a seasonal model and a covariate model. Model structures with alternative seasonal harmonics (one to three waves) and covariate combinations were compared using rolling one-step ahead predictions on a held-out test for 2024. Proper scoring rules for count forecasts were assessed using logarithmic score (logS) and ranked probability score (RPS). In addition, we visualised the predictive uncertainty using simulation-based fan chart (500 simulations).

Predictive performance of the models was evaluated using a rolling one-step-ahead forecasting framework, which assesses the ability of each model to predict the next observation based only on information available up to the previous time step. This approach avoids in-sample bias and mimics the conditions of an early-warning system. One-step-ahead residuals were visualised to assess for systematic bias in the modelling framework.

For each week  $t$ , the model was fitted using observations up to week  $t - 1$ , and the predictive distribution for week  $t$  was generated through stochastic simulation. Forecast accuracy and calibration were then assessed using proper probabilistic scoring rules. Two primary evaluation metrics were used:

Logarithmic Score (logS) measures the probability assigned by the model to the observed outcome in Equation 3. Lower logS values indicate that the model assigns higher probability to observed events, reflecting better probabilistic accuracy.

$$\log S = -\log p(Y_t) \quad (3)$$

where  $t$  = time  
 $p(Y_t)$  = predictive probability of the observed count

Ranked Probability Score (RPS) evaluates the distance between the predicted cumulative distribution function and the empirical outcome in Equation 4. Lower RPS values indicate improved calibration and sharper predictive distributions.

$$\text{RPS} = \sum_k (F(k) - 1(Y_t \leq k))^2 \quad (4)$$

where  $k$  = count  
 $F(k)$  = predictive cumulative distribution function  
 $1$  = indicator function  
 $Y_t$  = observed count

Finally, predictive uncertainty was visualised through fan charts, which display simulated predictive trajectories and quantile intervals over time. These plots illustrate how forecast uncertainty evolves during both epidemic peaks and low-incidence periods.

### 3. Results

Figure 2 shows the aggregated weekly H5N1 detections across Europe from 2021 to 2024. We observe strong seasonal waves, with major peaks occurring during winter and early spring. The largest epidemic wave occurred in late 2021 and early 2022, followed by subsequent seasonal peaks in late 2022 and early 2023, late 2023 and early 2024 and finally late 2024 leading into early 2025. These seasonal patterns are consistent with known migratory bird movements and environmental conditions that facilitate viral persistence in aquatic environments.

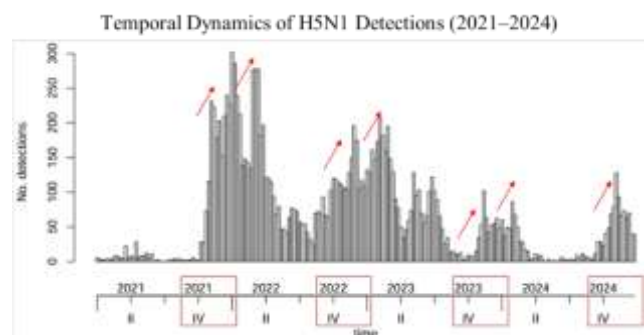


Figure 2: Temporal dynamics of H5N1 detections in Europe (2021–2024)

The Boruta algorithm evaluated the relative importance of each predictor by comparing the importance of real variables against randomly permuted shadow variables within a random forest framework. Temperature, humidity, snow depth, snow melt, goat and sheep density were the parameters discarded during the process. Furthermore, correlation analysis to reduce multicollinearity resulted in the removal of cattle density, transpiration and Accipitriformes bird family. The final set of predictors was reduced to a smaller subset of environmentally and ecologically meaningful variables. It included environmental variables (precipitation, soil moisture, wind speed), anthropogenic indicators (population density), landscape variables (wetland extent, elevation, slope, and distances to water bodies and roads), livestock densities, and bird community indicators representing multiple avian orders (Anseriformes, Charadriiformes, Ciconiiformes, Columbiformes, Galliformes, Gruiformes, Passeriformes, Pelecaniformes, Podicipediformes, Strigiformes).

Although several predictors were retained after feature selection, not all were incorporated simultaneously into the final endemic–epidemic model. Subsets of the retained predictors were evaluated through alternative endemic–epidemic model specifications, and their predictive performance was compared using proper scoring rules (RPS and logS). The final model retained a smaller set of covariates that consistently improved predictive performance and provided clear ecological interpretability. In particular, wind speed, wetland proximity, population density, and bird community densities (Anseriformes, Charadriiformes, and Pelecaniformes) emerged as the most informative predictors which were incorporated into the spatio-temporal modelling framework to evaluate their influence on outbreak persistence and spatial transmission.

Figure 3 presents the distribution of one-step-ahead residuals and the corresponding quantile–quantile (QQ) plot. The residual histogram is centred around zero, indicating that the model does not systematically overpredict or underpredict outbreak counts.

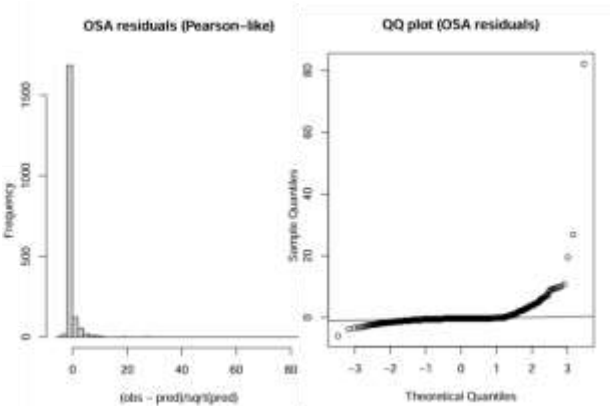


Figure 3: One-step-ahead residual diagnostics for the endemic–epidemic model

The model comparison across alternative seasonal structures and covariate combinations indicated that the configuration incorporating ecological covariates in both endemic and autoregressive components provided the best predictive performance ( $RPS = 0.61$ ,  $\log S = 0.75$ ). The endemic–epidemic model achieved substantially lower scores compared to the baseline GLM model, indicating improved probabilistic forecast accuracy (Table 2). The endemic–epidemic model reduced mean  $\log S$  by approximately 29% and  $RPS$  by nearly 50% relative to the GLM baseline.

Figure 4 presents the probabilistic forecast fan chart for the evaluation period (2023–2024). The coloured bands represent the 1–99 % predictive quantiles generated from 500 stochastic simulations, while the observed outbreak counts are shown as black points. The darker tones indicate higher probability density. The observed values fall predominantly within the central predictive intervals, indicating that the predictive distribution captures the variability of the outbreak process.

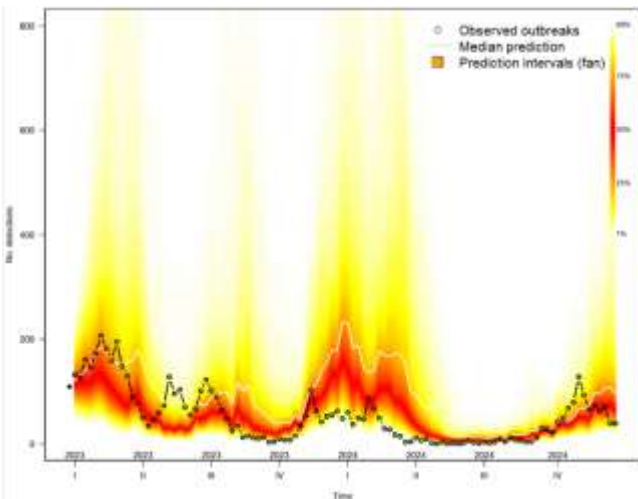


Figure 4: Probabilistic forecast of weekly H5N1 outbreak detections in Europe (2023–2024)

This study has a few limitations. First, the use of country-level aggregation may mask sub-national heterogeneity in outbreak dynamics. Second, surveillance data may be subject to reporting

bias and differences in detection capacity across countries. Third, while ecological covariates improve interpretability, they may not fully capture complex host–environment interactions. Future work could incorporate finer spatial resolution data and mobility-based transmission models.

| Model            | $\log S$ | RPS  |
|------------------|----------|------|
| Endemic–epidemic | 0.75     | 0.61 |
| GLM              | 1.06     | 1.19 |

Table 2: Comparison of predictive performance between the endemic–epidemic and GLM models.

#### 4. Discussion

The observed inter-annual variation in H5N1 outbreak magnitude across Europe suggests that, while seasonality plays an important role, additional ecological and environmental drivers substantially modulate outbreak intensity. Similar variability has been reported in European HPAI surveillance, where outbreak severity differs between years depending on migratory bird movements and prevailing environmental conditions (Fusaro et al., 2024). In particular, the high winter–spring peaks observed in the time series coincide with periods of increased waterbird aggregation, suggesting that host availability plays a primary role in shaping outbreak dynamics, with environmental conditions likely acting as a supporting factor for viral persistence. Conversely, the reduced outbreak activity during summer months corresponds to a decline in migratory bird densities, indicating that seasonal changes in host distribution rather than climatic variability alone may be the dominant driver of transmission patterns.

This temporal clustering indicates that H5N1 transmission dynamics are governed by a combination of seasonal ecological processes and short-term persistence mechanisms, reinforcing the need for spatio-temporal models capable of capturing both baseline risk and epidemic feedback. Identifying the key drivers underlying these patterns was therefore essential for understanding the mechanisms of transmission and improving predictive performance.

The feature selection results further support this interpretation. The dominance of bird community predictors over climatic variables suggests that host availability and ecological structure are stronger determinants of outbreak dynamics than short-term environmental variability alone. The retained predictors collectively represent three fundamental processes driving H5N1 transmission: environmental suitability for viral persistence, host community availability, and anthropogenic amplification pathways. This aligns with previous studies highlighting the central role of water-associated bird species and wetland habitats in maintaining and spreading avian influenza viruses.

However, not all selected predictors were included in the final endemic–epidemic model. Incorporating a large number of covariates in the endemic component can lead to overparameterisation and reduced predictive stability, particularly when modelling weekly surveillance data across multiple countries. As a result, alternative model specifications were evaluated, and only those predictors that consistently improved predictive performance and maintained ecological interpretability were retained. This approach ensured a balance between model complexity and robustness.

Residual diagnostics provide further insight into model performance. The one-step-ahead residuals are centred around zero, indicating that the model adequately captures the overall

temporal structure of the outbreak dynamics without systematic bias. Nevertheless, the residual distribution exhibits a right-skewed tail, indicating that the model occasionally underestimates large outbreak events. The QQ plot confirms that deviations occur primarily in the upper tail, corresponding to extreme epidemic peaks. Such behaviour is typical of infectious disease processes, where rare but significant amplification events can occur due to local ecological or anthropogenic conditions that are difficult to anticipate fully.

The model estimates indicate that H5N1 outbreaks across Europe from 2021 to 2024 were primarily driven by ecological and anthropogenic factors rather than purely seasonal cycles. Higher human population density and elevated densities of Anseriformes, Charadriiformes, and Pelecaniformes were associated with increased baseline risk and enhanced short-term persistence of outbreaks. These bird orders are strongly linked to aquatic habitats and migratory flyways, which are recognised pathways for long-distance virus dispersal. The association with human population density likely reflects indirect effects, such as poultry production intensity, farm density, and surveillance capacity, rather than direct human involvement in transmission. Importantly, comparison with the baseline regression model highlights the added value of explicitly modelling epidemic processes. The substantial reduction in both logS and RPS demonstrates that incorporating autoregressive persistence and spatial interaction mechanisms significantly improves predictive performance. This indicates that H5N1 transmission cannot be adequately represented using regression-based approaches alone, as they fail to capture the temporal dependence and spatial connectivity inherent in epidemic systems.

The probabilistic forecasts further illustrate the strengths of the modelling framework. The widening of predictive intervals during periods of high incidence reflects increased uncertainty associated with epidemic amplification, while narrower intervals during low-incidence periods indicate more stable endemic conditions. The model successfully reproduces both the rapid increase in outbreak counts during epidemic onset and the subsequent decline, demonstrating its ability to capture the full temporal evolution of outbreak dynamics.

These findings highlight the value of probabilistic forecasting for surveillance and decision-making. The ability to anticipate periods of elevated outbreak risk enables more targeted monitoring and timely intervention strategies, particularly in regions with high ecological suitability and host density. This is especially relevant for cross-border disease management, where coordinated responses are required to mitigate the spread of H5N1 across Europe.

Overall, the results demonstrate that combining ecological information with spatio-temporal epidemic modelling provides a robust framework for understanding and predicting H5N1 transmission dynamics.

## 5. Conclusion

This study developed a spatio-temporal modelling framework to analyse and forecast H5N1 outbreak dynamics across Europe using an endemic–epidemic approach. By integrating environmental, ecological, and anthropogenic covariates within a probabilistic framework, the model successfully captured the temporal evolution and seasonal variability of outbreaks between 2021 and 2024. Feature selection results highlighted the dominant role of waterbird community structure, wetland habitats, and population density in shaping outbreak risk,

reinforcing the importance of ecological drivers in HPAI transmission. Compared to a regression-based baseline model, the proposed framework demonstrated substantially improved predictive performance and calibration, indicating that explicitly modelling outbreak persistence and spatial interactions provides significant advantages for forecasting infectious disease dynamics. The ability of the model to generate realistic probabilistic forecasts highlights its potential utility for early-warning systems and surveillance planning. Overall, this work demonstrates that combining spatio-temporal modelling with ecological data offers a robust and scalable approach for understanding and predicting H5N1 outbreaks.

## References

- Center For International Earth Science Information Network-CIESIN-Columbia University. (2018). *Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11* [Dataset]. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H49C6VHW>
- European Food Safety Authority (EFSA), Abrahantes, J. C., Aznar, I., Boom, M., Catalin, I., Dórea, F., Grant, M., Mulligan, K. F., Zancanaro, G. (2024). Avian Influenza annual report 2024. *EFSA Journal*, 23(12), e9761. <https://doi.org/10.2903/j.efsa.2025.9761>
- Fink, D., Auer, T., Johnston, A., Strimas-Mackey, M., Ligocki, S., Robinson, O., Hochachka, W., Jaromczyk, L., Crowley, C., Dunham, K., Andrew, Davies, I., Rodewald, A., Ruiz-Gutierrez, V., Wood, C. (2023). *eBird Status and Trends* (Version 2022) [Dataset]. Cornell Lab of Ornithology, Ithaca, New York. <https://doi.org/10.2173/ebirdst.2022>
- Fusaro, A., Zecchin, B., Giussani, E., Palumbo, E., Agüero-García, M., Bachofen, C., Bálint, Á., Banihashem, F., Banyard, A. C., Beerens, N., Bourg, M., Briand, F.-X., Brøjer, C., Brown, I. H., Brugger, B., Byrne, A. M. P., Cana, A., Christodoulou, V., Dirbakova, Z., Monne, I. (2024). High pathogenic avian influenza A(H5) viruses of clade 2.3.4.4b in Europe—Why trends of virus evolution are more difficult to predict. *Virus Evolution*, 10(1), veae027. <https://doi.org/10.1093/ve/veae027>
- Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T. P., Vanwambeke, S. O., Wint, G. R. W., Robinson, T. P. (2018). Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Scientific Data*, 5(1), 180227. <https://doi.org/10.1038/sdata.2018.227>
- Gneiting, T., Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477), 359–378. <https://doi.org/10.1198/016214506000001437>
- Hayes, S., Hilton, J., Mould-Quevedo, J., Donnelly, C. A., Baylis, M., Brierley, L. (2025). Ecology and environment predict spatially stratified risk of H5 highly pathogenic avian influenza clade 2.3.4.4b in wild birds across Europe. *Scientific Reports*, 16(1), 997. <https://doi.org/10.1038/s41598-025-30651-9>
- Jung Kjær, L., Ward, M. P., Boklund, A. E., Larsen, L. E., Hjulsgager, C. K., Kirkeby, C. T. (2023). Using surveillance data for early warning modelling of highly pathogenic avian influenza in Europe reveals a seasonal shift in transmission 2016–2022. *Scientific Reports*, 13(1), 15396. <https://doi.org/10.1038/s41598-023-42660-7>
- Kursa, M. B., Jankowski, A., Rudnicki, W. R. (2010). Boruta – A System for Feature Selection. *Fundamenta Informaticae*, 101(4), 271–285. <https://doi.org/10.3233/FI-2010-288>
- Meyer, S., Held, L., Höhle, M. (2017). Spatio-Temporal Analysis of Epidemic Phenomena Using the R Package surveillance. *Journal of Statistical Software*, 77, 1–55. <https://doi.org/10.18637/jss.v077.i11>

- Opata, M. R., Lavarello-Schettini, A., Semenza, J. C., Rocklöv, J. (2025). Predictiveness and drivers of highly pathogenic avian influenza outbreaks in Europe. *Scientific Reports*, 15(1), 20286. <https://doi.org/10.1038/s41598-025-04624-x>
- Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, 85(3), 381–394. <https://doi.org/10.1175/BAMS-85-3-381>
- Schreuder, J., de Knecht, H. J., Velkers, F. C., Elbers, A. R. W., Stahl, J., Slaterus, R., Stegeman, J. A., de Boer, W. F. (2022). Wild Bird Densities and Landscape Variables Predict Spatial Patterns in HPAI Outbreak Risk across The Netherlands. *Pathogens*, 11(5), 549. <https://doi.org/10.3390/pathogens11050549>
- Xie, R., Edwards, K. M., Wille, M., Wei, X., Wong, S.-S., Zanin, M., El-Shesheny, R., Ducatez, M., Poon, L. L. M., Kayali, G., Webby, R. J., Dhanasekaran, V. (2023). The episodic resurgence of highly pathogenic avian influenza H5 virus. *Nature*, 622(7984), Article 7984. <https://doi.org/10.1038/s41586-023-06631-2>