

Executive Summary

With highly pathogenic avian influenza (HPAI) disrupting the poultry industry and wildlife, we developed a county-level machine learning framework for **classification** and **pseudo-forecasting** of HPAI outbreaks.

This framework supports proactive, data-driven disease management across the United States.

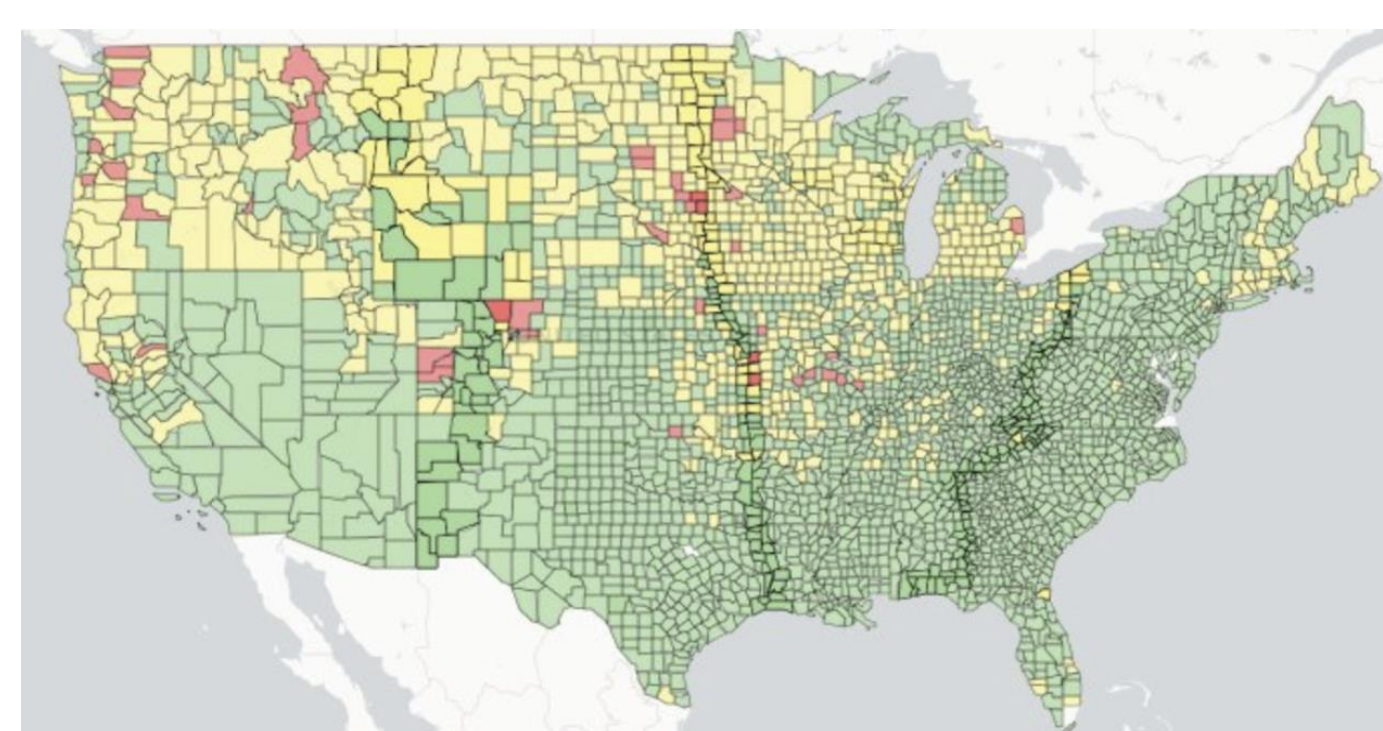


Figure 1. Pseudo-forecast Output Map

Background

Since 2022, HPAI outbreaks have caused over **\$1 Billion in damages** for egg and poultry-product consumers and has infected wild bird populations across the United States.

Current mitigation strategies are reactive, and no county-level national system exists to provide risk predictions.

This gap highlights the need for data-driven forecasting tools to support early intervention and disease suppression.

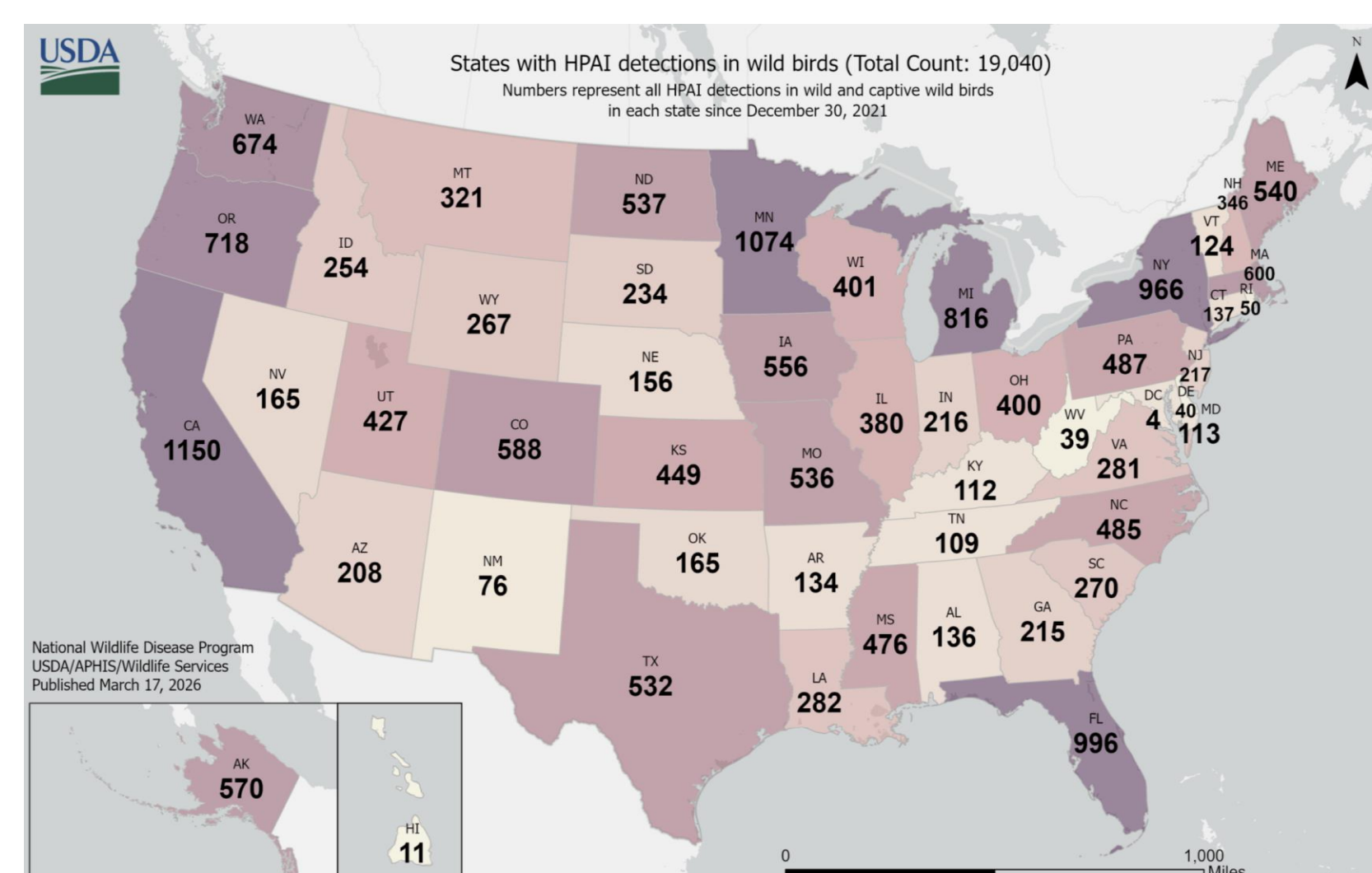


Figure 2. Nationwide distribution of HPAI outbreaks

Experimental Setup

The framework utilizes **ensemble-based machine learning modeling** that operates at the county-monthly level across the conterminous United States, incorporating environmental, climatological, and geospatial features. The two models developed include:

- a classification model that detects HPAI presence across the CONUS and
- a pseudo-forecasting model that predicts risk levels one month in advance.

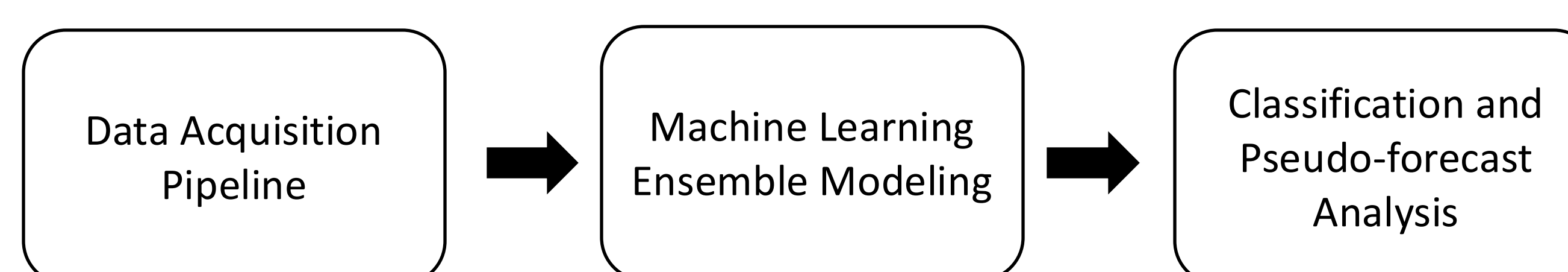


Figure 3. HPAI Machine Learning Overview

Feature importance, principal component analysis (PCA), and post-hoc analysis were used to interpret model behavior and identify key drivers of outbreak risk.

Experimental Test Plan

Both classification and forecasting models were evaluated on held-out county-month observations across the conterminous United States (CONUS). The classification test set consisted of mixed temporal samples, while the forecasting test set used fully unseen future months to simulate real-world deployment conditions.

Due to **extreme class imbalance** (~0.01% positive cases), **balanced accuracy** and **recall** were emphasized as primary evaluation metrics. Recall was prioritized to minimize missed HPAI detections, which are critical in disease surveillance contexts, while balanced accuracy ensured performance was not biased toward the majority negative class.

The classification model was assessed by comparing predicted presence/absence against observed HPAI detections at the county-month level. The pseudo-forecasting model was evaluated by comparing one-month-ahead risk classifications to detections observed in the subsequent month, enabling assessment of practical forecasting utility.

Confusion-matrix-based analysis was used to quantify false positives and missed detections, providing insight into tradeoffs between sensitivity and precision. Additional post-hoc interpretability analysis was conducted to confirm that dominant predictors aligned with known environmental and climatological drivers of HPAI spread.

These evaluation strategies ensure the framework is suitable for future research endeavors and real-world, forward-looking decision support.

Experimental Results

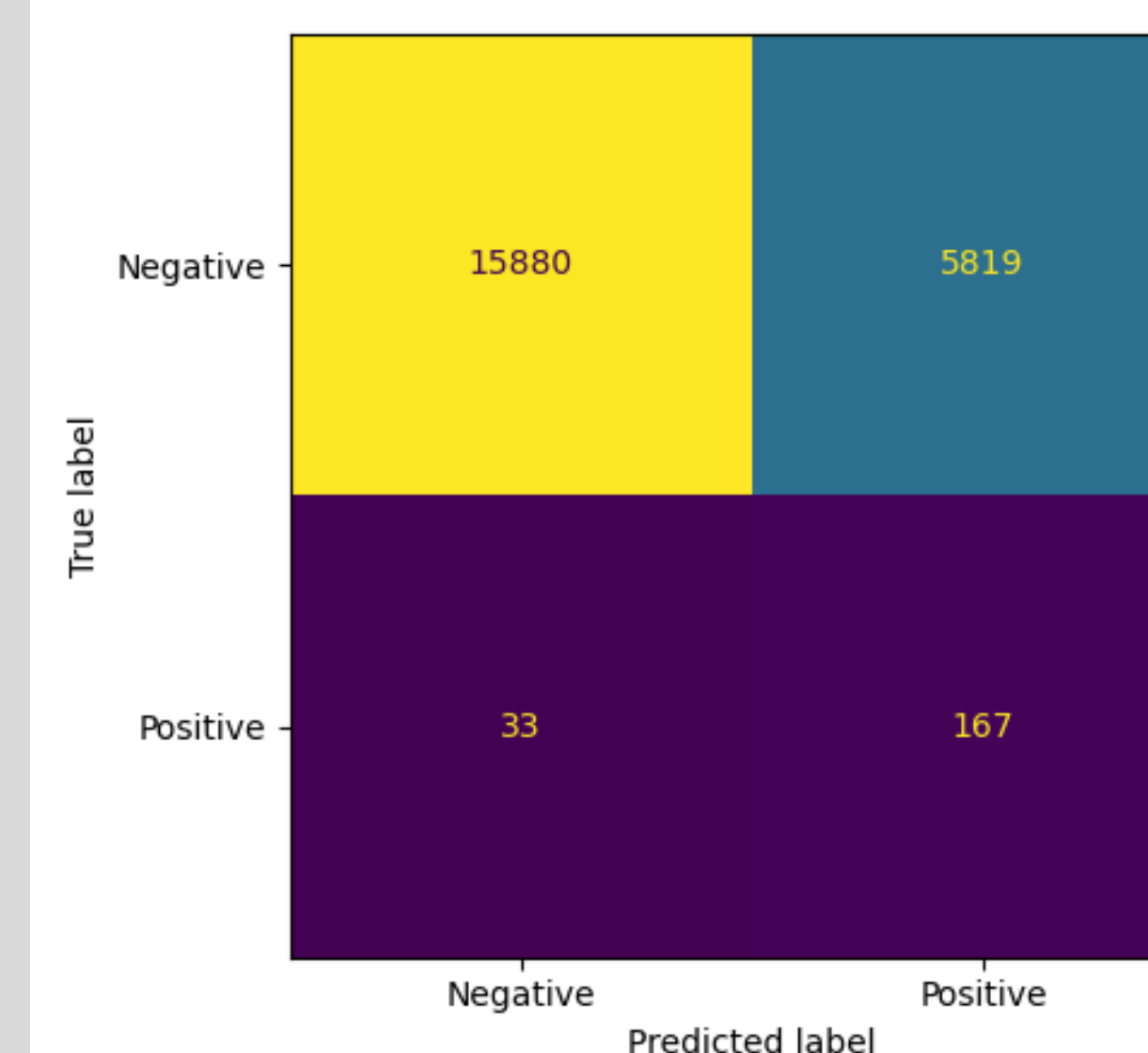


Figure 4. Classification Confusion Matrix

Modeling analysis revealed that the classification task achieved a balanced accuracy of **0.78** and **0.83**, respectively. Meanwhile, the pseudo-forecast model achieved **0.71** balanced accuracy and **0.66** recall. Feature importance confirmed ecological and climatological variables contributed the most to HPAI outbreaks.

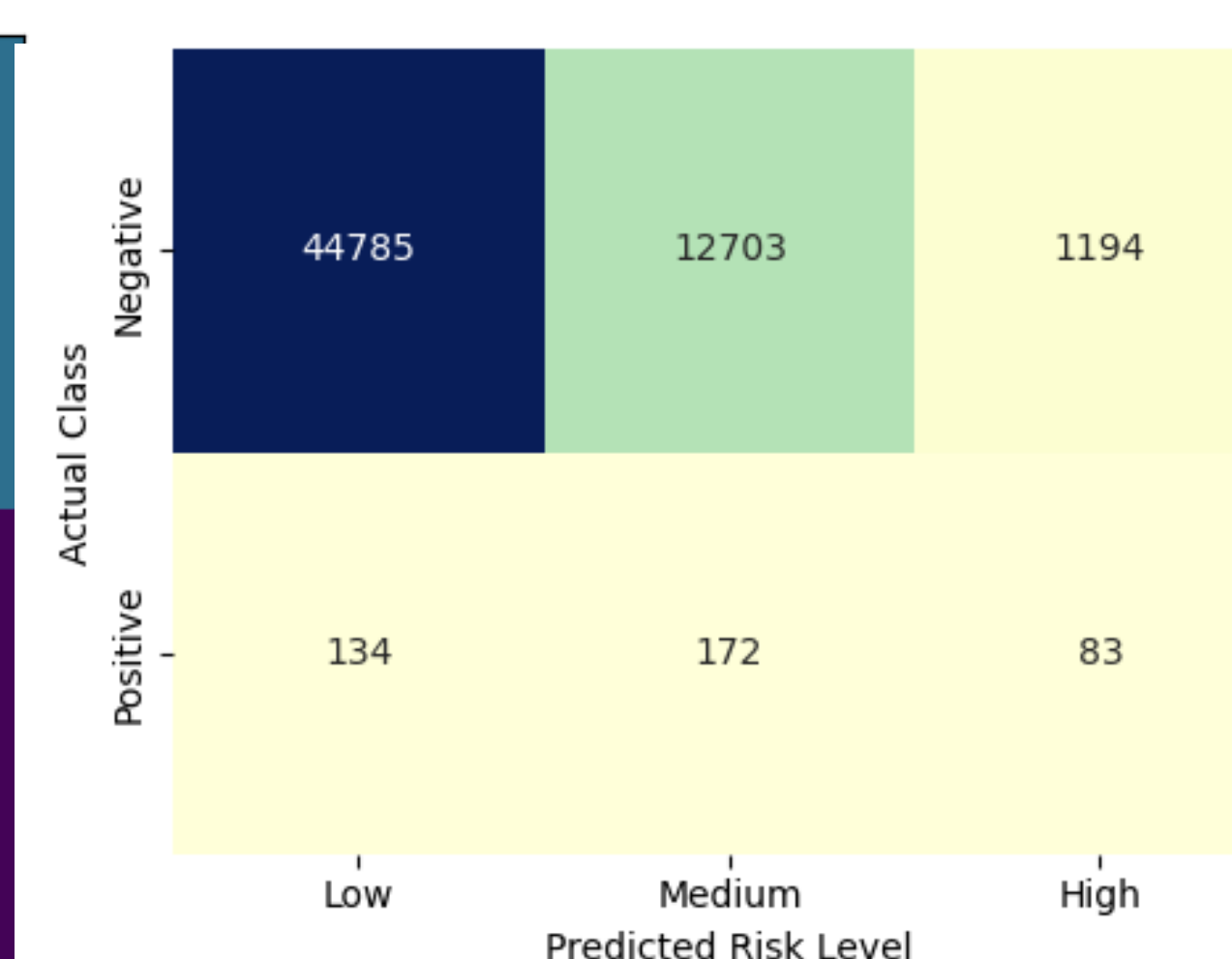


Figure 5. Pseudo-forecast Confusion Matrix

Conclusions

This work demonstrates a national-scale, forward-looking HPAI risk prediction framework capable of detecting outbreaks and pseudo-forecasting risk one month in advance across the United States. Beyond strong classification and pseudo-forecasting metrics, this framework supports a shift from reactive outbreak response to **proactive risk management**.

Future work will focus on:

- Improving spatiotemporal forecasting performance
- Integrating additional ecological and biological signals
- Deploying the framework as an operational decision-support tool

This research was conducted in collaboration with USDA-ARS and the Smart Agriculture Research Lab.

References

1. Mostafa, A., Nogales, A. & Martinez-Sobrido, L. Highly pathogenic avian influenza H5N1 in the United States: recent incursions and spillover to cattle. *npj Viruses* 3, 54 (2025).
2. Padilla, S. L., Baker, Q. J. & MacLachlan, M. J. The impact of HPAI trade restrictions on U.S. poultry exports in 2022–23. *Applied Eco Perspectives Pol* 47, 933–948 (2025).
3. Kilpatrick, A. M. et al. Predicting the global spread of H5N1 avian influenza. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19368–19373 (2006).
4. Breban, R., Drake, J. M., Stallknecht, D. E. & Rohani, P. The Role of Environmental Transmission in Recurrent Avian Influenza Epidemics. *PLoS Comput Biol* 5, e1000346 (2009).
5. United States Department of Agriculture Animal and Plant Health Inspection Service. Confirmations of Highly Pathogenic Avian Influenza in Commercial and Backyard Flocks.