

A methodological framework for modeling plant virus occurrence using biometeorological data: insights from multi-crop case studies in Argentina

Franco Suarez^{1,2}, Brenda Gómez Montenegro², Carolina Dottori², Vanina Alemandri², Soledad de Breuil², Cecilia Bruno^{1,2}

¹Unidad de Fitopatología y Modelización Agrícola (UFYMA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Córdoba, Argentina

²Cátedra de Estadística y Biometría, Facultad de Ciencias Agropecuarias (FCA), Universidad Nacional de Córdoba, Argentina

INTRODUCTION

Viral diseases represent a major threat to the productive stability and resilience of modern agricultural systems. Their spatial and temporal occurrence is strongly influenced by complex interactions among viruses, vectors, hosts, and environmental conditions, making disease distribution difficult to predict using traditional statistical approaches. This complexity highlights the critical need to understand and characterize the macro-environmental drivers affecting crop health and to develop analytical tools that facilitate proactive decision-making. Within the context of precision agriculture, the integration of high-resolution, climate-based predictive models has emerged as a key strategy for building robust decision-support systems aimed at digital plant health surveillance and site-specific disease management. The objective of this study was to present a scalable methodological framework for modeling and mapping virus presence/absence in agricultural crops. The utility of this framework is demonstrated through two contrasting case studies conducted in major production regions of Argentina: wheat streak mosaic virus (WSMV) in wheat and groundnut ringspot orthotospovirus (GRSV) in peanut.

MATERIALS AND METHODS

The proposed framework integrates four sequential stages optimized for plant pathology modeling: (i) **georeferenced database compilation**: georeferenced virus presence/absence databases were compiled through systematic, spatial field sampling coupled with laboratory diagnostic confirmation. The presence of WSMV was confirmed via DAS-ELISA using commercial antisera (Agdia Inc., USA). Conversely, GRSV presence was confirmed by RT-LAMP (loop-mediated isothermal amplification), using the WarmStart Colorimetric LAMP 2X Master Mix with UDG (New England Biolabs) and, primer set specific for GRSV (Breuil et al. 2026). Amplification results were determined visually based on colorimetric change, a shift from pink to yellow indicated a positive reaction, while samples that remained pink were classified as negative; (ii) **biometeorological variable integration**: high-resolution environmental predictors were derived from the ERA5-Land reanalysis datasets at a 0.1° spatial resolution and hourly temporal frequency (Muñoz-Sabater et al., 2021). Target climatic variables, including temperature, precipitation, relative humidity, and wind speed, were extracted for both the pre-sowing window and the active crop cycle. These hourly data points were temporally aggregated into growing season and monthly predictors, (iii) **predictor selection and dimensionality reduction**: feature selection was performed using stepwise logistic regression, information criteria (AIC and BIC), statistical significance, and multicollinearity diagnostics, ensuring Variance Inflation Factor (VIF) below 5 (Suarez et al., 2023); and (iv) **predictive model fitting**: individual models were developed for each pathosystem according to their epidemiological characteristics. Specifically, a Random Forest Model was employed to capture non-linear relationships between climatic variables and virus presence for the WSMV-wheat pathosystem. Conversely, a Logistic Regression Model was fitted for the GRSV-peanut pathosystem to explicitly interpret the directional effect of environmental predictors. Model performance was evaluated using cross-validation and accuracy, sensitivity, specificity, and area under the ROC curve (AUC). Final model outputs were projected spatially into diagnostic risk maps using GIS environments.

RESULTS

For the WSMV-wheat pathosystem, the predictive Random Forest Model trained on historical data spanning 2006 to 2023 achieved an overall accuracy of 78% and an AUC of 0.85. The most influential spatial feature driving virus risk were regional wind speed patterns and relative humidity during the months preceding crop establishment. Projecting this model enable the generation of high-resolution, regional-scale risk maps detailing the probability of virus occurrence (Figure 1, left) (Montenegro et al., 2026). For the GRSV-peanut

pathosystem, the Logistic Regression Model, relying exclusively on winter biometeorological conditions prior to sowing, demonstrated satisfactory predictive performance. Independent validation yielded, an accuracy of 79%, a specificity of 87%, and a sensitivity of 61%. Temperature, wind speed, relative humidity, and localized precipitation were identified as key spatial determinants of disease development. This model was utilized to generate diagnostic phytopathological risk maps that effectively delineate high and low risk zones and evaluate their interannual spatial stability across the primary peanut-producing region of Argentina (Figure 1, right) (Dottori et al., 2025).

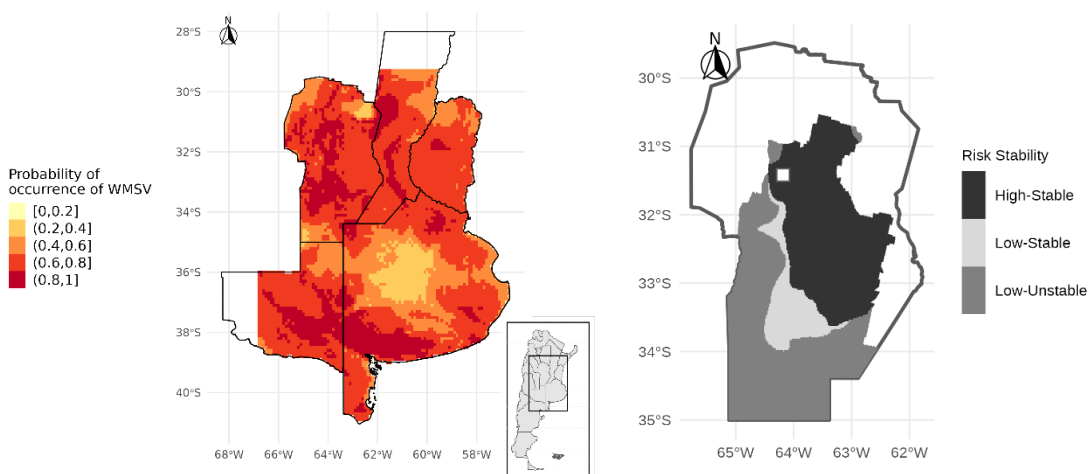


Figure 1. Predictive risk mapping outputs. Spatial probability of occurrence for WSMV in wheat (left). Interannual spatial risk stability zones for GRSV in peanut crops (right).

CONCLUSIONS

The findings demonstrate that the proposed methodological framework is highly robust, flexible, and adaptable for modeling plant virus risk across diverse macro-climatic environments. By integrating standard statistical methods with machine learning algorithms according to the distinct epidemiology of each pathogen, we successfully transformed raw, gridded climate data from the ERA5-Land reanalysis into operational, spatially explicit smart farming tool. Although the evaluated pathosystem differs substantially in vector biology, crop cycle and data available, the framework proved readily generalizable. This analytical workflow establishes a solid foundation for the deployment of early warning systems, targeted pest monitoring, and variable-rate crop protection applications, ultimately enhancing plant health resilience and site-specific management efficiency under increasing climate variability.

REFERENCES

- de Breuil, S., López, S., Dottori, C., Nome, C., & Adkins, S. (2026). Visual loop-mediated isothermal amplification (LAMP) assay for specific detection of groundnut ringspot and tomato spotted wilt orthotospoviruses. *Journal of Plant Pathology*, 108, 763–768.
- Dottori, C., Suarez, F., Córdoba, M., Giannini-Kurina, F., & de Breuil, S. (2025). Predicting and mapping orthotospovirus risk in peanut crops using biometeorological data. *European Journal of Agronomy*, 170, 127742. <https://doi.org/10.1016/J.EJA.2025.127742>
- Montenegro, B. E. G., Suárez, F. M., Balzarini, M., Giolitti, F., Martino, J., Bruno, C., & Alemandri, V. (2026). Climate-Associated Risk Patterns of Wheat Streak Mosaic Virus (WSMV) in Wheat-Producing Regions of Argentina. *Journal of Phytopathology*, 174(3), e70334. <https://doi.org/10.1111/JPH.70334>
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J. N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383.
- Suarez, F., Bruno, C., Kurina Giannini, F., Giménez Pecci, M. P., Rodríguez Pardina, P., & Balzarini, M. (2023). Marriage between variable selection and prediction methods to model plant disease risk. *European Journal of Agronomy*, 151, 126995. <https://doi.org/10.1016/J.EJA.2023.126995>