SÃO PAULO STATE UNIVERSITY – UNESP

JABOTICABAL CAMPUS

AGROMETEOROLOGICAL MODELING USING MACHINE LEARNING TO PREDICT MAIZE WHITE SPOT

Andrea Onelia Rodriguez Roa

Bachelor of Science in Physics

2026

SÃO PAULO STATE UNIVERSITY – UNESP

JABOTICABAL CAMPUS

AGROMETEOROLOGICAL MODELING USING MACHINE LEARNING TO PREDICT MAIZE WHITE SPOT

Andrea Onelia Rodriguez Roa

Advisor: Prof. Dr. Glauco de Souza Rolim

Co-advisor: Dr. Gisèle Maria Fantin

Thesis submitted to the School of Agricultural and Veterinary Sciences - UNESP, Jaboticabal Campus, as part of the requirements for the degree of Doctor in Agronomy (Plant Production).

2026

R696a

Rodriguez-Roa, Andrea Onelia

Agrometeorological modeling using machine learning to predict maize white spot / Andrea Onelia Rodriguez-Roa. -- Jaboticabal, 2026

84 p. : il., tabs., mapas

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Glauco de Souza Rolim

Coorientadora: Gisèle Maria Fantin

1. Meteorologia agrícola. 2. Modelagem de dados. 3. Machine learning. 4. Fitopatologia. 5. Milho. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Dados fornecidos pelo autor(a).

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: AGROMETEOROLOGICAL MODELING USING MACHINE LEARNING TO PREDICT MAIZE WHITE SPOT

AUTORA: ANDREA ONELIA RODRIGUEZ ROA
ORIENTADOR: GLAUCO DE SOUZA ROLIM
COORIENTADOR: COORIENTADORA: GISÈLE MARIA FANTIN

Aprovada como parte das exigências para obtenção do Título de Doutora em Agronomia (Produção Vegetal), pela Comissão Examinadora:

gov.br Documento assinado digitalmente
GLAUCO DE SOUZA ROLIM
Data: 13/04/2026 23:50:54-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. GLAUCO DE SOUZA ROLIM (Participaçao Virtual)
Departamento de Ciências Exatas / UNESP / Câmpus de Jaboticabal - FCAV

gov.br Documento assinado digitalmente
CLAUDIA VIEIRA GODOY
Data: 25/04/2026 15:38:22-0300
Verifique em https://validar.iti.gov.br

Pesquisadora Dra. CLAUDIA VIEIRA GODOY (Participaçao Virtual)
Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) / Londrina/PR

Assinado por: **Hélder José Chaves Fraga**
Num. de Identificação: 11973041
Data: 2026.04.23 20:02:08+01'00'

Pesquisador Dr. HÉLDER JOSÉ CHAVES FRAGA (Participaçao Virtual)
Departamento de Agronomia / Universidade de Trás-os-Montes e Alto Douro (UTAD) - Vila Real/Portugal

gov.br Documento assinado digitalmente
ODAIR APARECIDO FERNANDES
Data: 22/04/2026 16:01:50-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. ODAIR APARECIDO FERNANDES (Participaçao Virtual)
Departamento de Fitossanidade / UNESP / Câmpus de Jaboticabal - FCAV

gov.br Documento assinado digitalmente
ALAN RODRIGO PANOSSO
Data: 14/04/2026 08:20:04-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. ALAN RODRIGO PANOSSO (Participaçao Virtual)
Departamento de Ciências Exatas / UNESP / Câmpus de Jaboticabal - FCAV

Jaboticabal, 07 de abril de 2026.

AUTHOR ACADEMIC BACKGROUND

Andrea Onelia Rodriguez Roa is a Colombian national born in Bogotá, Colombia. She holds a B.S. degree in Physics from Universidad Distrital Francisco José de Caldas (2006). She earned an M.Sc. degree in Meteorology from the Universidad Nacional de Colombia (2011), supported by a fellowship from Conservation International, and completed a Specialization in Geomatics at the Universidad Militar Nueva Granada (2016).

She has more than ten years of professional experience in climatology and agroclimatology, with an emphasis on climate variability and climate change, climate modeling, the development of climate change scenarios, agroclimatic risk analysis, and the design of expert systems and decision-support tools.

She worked as a researcher and consultant at the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM), Colombia, between 2010 and 2014, contributing to projects focused on the evaluation of global climate models and the development of climate change scenarios for Colombia. She worked at the Colombian Corporation for Agricultural Research (Agrosavia) from 2011 to 2015, where she developed expert systems, agroclimatic early warning systems, and tools for agroclimatic risk management. She also participated in projects related to protected agriculture and the spatial modeling of pests and diseases.

She has served as a researcher at Agrosavia since 2015, leading and developing projects focused on agrometeorological modeling, risk analysis, agroclimatic zoning, and the application of data science techniques to optimize agricultural production systems.

In 2023, she began her Ph.D. in Agronomy at São Paulo State University (UNESP), Jaboticabal campus, Brazil, supported by a CAPES fellowship. She is a member of the Group of Agrometeorological Studies (GAS) at the same institution. Her research is conducted within the Digital Agriculture research line, with an emphasis on agrometeorology, predictive modeling, and machine learning applied to the sustainable management of agricultural systems.

“Chaos: When the present determines the future, but the approximate present does not approximately determine the future”

Edward Lorenz

To my parents, Blanca Roa and Gilberto Rodriguez, for their unconditional love and constant support, for always being by my side, sustaining me through every fall and celebrating each achievement.

ACKNOWLEDGMENTS

To São Paulo State University (UNESP) – School of Agricultural and Veterinary Sciences, Jaboticabal Campus, for the opportunity to pursue the Ph.D. in Agronomy (Crop Production).

To the Coordination for the Improvement of Higher Education Personnel (CAPES), Brazil (Funding Code 001), for the financial support provided through the fellowship that made this doctoral training possible.

To Professor Dr. Glauco de Souza Rolim, for his guidance throughout the development of this research and for the opportunity to join his research group. I am deeply grateful for the knowledge shared, the academic rigor, and the dedication to my training, as well as for his continued support over these years.

To the Biological Institute, for the institutional partnership established for the development of this research, and to Dr. Gisele Maria Fantin, for her trust in sharing information resulting from many years of work and for her valuable contributions to this thesis.

To the Department of Exact Sciences, its faculty members, administrative staff, and my colleagues from the Faculty and the GAS group, for their support, collaboration, and for fostering an academic environment that contributed to the development of this work.

To the Colombian Corporation for Agricultural Research (Agrosavia), for the institutional and professional support that enabled the completion of this doctoral training.

To my parents, siblings, nephews, and all my family, for their unconditional support, for the trust they placed in me, and for accompanying me with love and strength throughout every stage of this journey.

To my friends, Denise Kubilius, Mariana Dias, and Rafael de Lima, for sharing this academic stage, for the learning experiences built together, and for the moments lived both inside and outside the university. To Adolfo Leão, for his support, affection, and companionship during these years away from home.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

TABLE OF CONTENTS

CHAPTER 1 - General considerations 1
1.1 Introduction 1
1.2. Literature review 1
1.2.1 Maize cultivation and major foliar diseases 1
1.2.2 Agrometeorological conditions and plant disease development 3
1.2.3 Artificial intelligence in plant disease prediction 5
References 6
CHAPTER 2 - Influence of agrometeorological variables on the severity of Maize White Spot and its effect on yield in tropical and subtropical regions 10
2.1 Introduction 11
2.2 Materials and methods 13
2.2.1 Study area 13
2.2.2 Analysis of MWS severity and crop yield 16
2.2.3 Analysis of MWS severity and weather variables 17
2.3 Results and discussion 18
2.3.1 Yield response to MWS severity in maize crops 18
2.3.2 Correlation between agrometeorological variables and MWS severity 24
2.3.3 Ranges of influence of agrometeorological variables on MWS severity 26
2.4 Conclusions 31
References 32
CHAPTER 3 - Early prediction of maize white spot severity through machine learning and agrometeorological time windows 37
3.1 Introduction 37
3.2 Materials and methods 40
3.2.1 Study area 40
3.2.2 Field and weather data 40
3.2.3 Agrometeorological variables and derived indices 42
3.2.4 Hybrid resistance classification 43
3.2.5 Correlation-based definition of agroclimatic windows 44
3.2.6 Machine learning modeling approach for MWS severity prediction 44
3.2.7 Model interpretability 47

3.3 Results....47
3.3.1 Maize white spot severity and agrometeorological conditions....47
3.3.2 Daily correlations between agrometeorological variables and MWS severity....48
3.3.3 Identification of agroclimatic time windows....48
3.3.4 Model performance for MWS severity prediction....51
3.3.5 Model interpretability: SHAP analysis....54
3.4 Discussion....57
3.4.1 Predictive performance and modeling strategy....57
3.4.2 Role of host resistance in MWS severity prediction....58
3.4.3 Influence of agroclimatic conditions across phenological windows....58
4.4.4 Implications for disease forecasting and crop management....60
3.5 Conclusions....61
References....62
CHAPTER 4 - Final considerations....66
APPENDICES....68
Appendix A - Supplementary material of chapter 2....69
Appendix B - Supplementary material of chapter 3....71

AGROMETEOROLOGICAL MODELING USING MACHINE LEARNING TO PREDICT MAIZE WHITE SPOT

ABSTRACT - Maize (*Zea mays L.*) is one of the most important crops worldwide due to its contribution to food security and its economic relevance, particularly in tropical and subtropical regions. However, maize production is influenced by diverse biotic and abiotic factors, among which foliar diseases represent a major constraint on grain yield and quality. *Phaeosphaeria* leaf spot or Maize White Spot (MWS) is one of the most important foliar diseases, and its severity is strongly modulated by agrometeorological conditions. In this context, understanding the environmental factors that influence disease development is essential for improving prediction and management strategies. The objective of this study was to evaluate the relationships between agrometeorological variables and MWS severity and developed a machine learning-based predictive framework through the integration of environmental and crop-related variables. Disease severity data collected between 1998 and 2022 from major maize-producing municipalities in São Paulo State, Brazil, were combined with daily meteorological data obtained from the NASA POWER dataset. Correlation analyses were conducted to identify associations between disease severity and environmental variables, and a linear regression model was fitted to estimate the relationship between severity and yield. Then, four machine learning algorithms (Random Forest, Support Vector Regression, Multilayer Perceptron, and XGBoost) were evaluated alongside Multiple Linear Regression as a reference model, using agrometeorological variables aggregated over temporal windows throughout the crop cycle, together with crop-specific variables such as hybrid type, growing degree days to flowering, and cultivar resistance level. The results indicated that maize yield decreased by approximately 45.7 kg ha^{-1} (≈1%) for each 1% increase in disease severity. Higher severity levels were associated with relative humidity above 55%, soil moisture greater than 0.3, mean temperatures between 15-22 °C, maximum temperatures between 20-30 °C, diurnal temperature ranges of 8-16 °C, and wind speeds below 3 m s^{-1}. Machine learning models consistently outperformed the linear approach, with XGBoost achieving the best predictive performance (RMSE = 0.519; adjusted R^2 = 0.805 in the test dataset). Cultivar resistance was identified as the most influential predictor, followed by variables related to water availability during advanced vegetative and early reproductive stages. Overall, these findings provide quantitative support for plant protection risk and establish a foundation for the development of decision-support tools aimed at improving sustainable maize disease management under conditions of climate variability.

Keywords: Agrometeorology, Decision Support Systems, Machine Learning, Maize White Spot, Predictive modeling.

CHAPTER 1 - General considerations

1.1 Introduction

Maize (*Zea mays L.*) is one of the most important crops worldwide due to its contribution to food security and its economic relevance, particularly in tropical and subtropical regions (Tanumihardjo et al., 2020; FAO, 2022; Wang and Zhang, 2024). The United States, China, Brazil, Argentina, and India account for most of the global maize production (FAO, 2022; USDA, 2025b), with Brazil standing out as the leading producer in South America, the third largest producer worldwide, and the largest exporter of the crop (USDA, 2025b). This leadership is largely explained by the possibility of growing maize in two cropping seasons per year, the summer crop and the second season known as “safrinha”, which maximizes land use efficiency and annual yield (Cruz et al., 2021; Duarte et al., 2021).

Maize production is affected by a wide range of biotic and abiotic factors, among which foliar diseases are recognized as one of the main causes of yield reduction and grain quality loss (Codognoto et al., 2017; R. S. Silva et al., 2021). Among these diseases, *Phaeosphaeria* leaf spot, also known as maize white spot (MWS), has become one of the most relevant diseases in maize production systems, as it is associated with substantial yield losses that, under favorable environmental conditions and in susceptible cultivars, may exceed 50% (De Rossi et al., 2017; Embrapa, 2021; Xing et al., 2023).

However, the occurrence and development of MWS exhibit marked spatial and temporal variability, as they do not depend solely on cultivar susceptibility but rather on the complex interaction between environmental factors, particularly agrometeorological variables, and the phenological stage of the crop (Agrios, 2005a; Amorim et al., 2016; Xing et al., 2023). This interaction leads to contrasting responses across years, regions, and management systems, which hinders the epidemiological characterization of the disease, the identification of severity trends, and the estimation of its impact on yield, especially in tropical and subtropical environments characterized by high intra- and interannual climatic variability.

Despite advances in the understanding of MWS, a large proportion of existing studies has relied on traditional epidemiological and statistical approaches, which typically analyze the relationship between disease severity at a local scale and/or consider a limited number of environmental variables. Although these approaches have contributed to identifying factors that favor disease occurrence, they present limitations in capturing the nonlinear nature of the interactions among environment, host, and pathogen. Consequently, the ability of such methods to predict disease severity and anticipate risk scenarios remains restricted, particularly under highly variable climatic conditions.

In response to these limitations, recent advances in artificial intelligence (AI), and particularly in machine learning (ML) algorithms, have opened new opportunities for the analysis and prediction of crop diseases by enabling the simultaneous integration of multiple environmental, crop, and pathogen-related variables, as well as the modeling of complex nonlinear relationships. Recent studies have shown that these approaches can outperform traditional statistical methods in estimating and predicting the occurrence and intensity of plant diseases (Fenu and Malloci, 2021; González-Domínguez et al., 2023; Jensen et al., 2025). Nevertheless, in maize production systems, and particularly for MWS, the application of machine learning models remains poorly explored, despite their strong potential to represent dynamic interactions among environment, host, and pathogen and to anticipate disease risk scenarios.

Considering the importance of maize production, the impact of MWS on crop productivity, and the central role of agrometeorological variables in the epidemiological dynamics of the disease, the objective of this study is to evaluate the influence of these variables on MWS severity and to develop an agrometeorological model for its prediction using machine learning algorithms, based on field-observed disease data collected across multiple years and locations in the main maize-producing regions of São Paulo State, Brazil.

1.2. Literature review

1.2.1 Maize cultivation and major foliar diseases

Together with wheat and rice, maize is one of the most important agricultural crops worldwide (Fig. 1), both due to its high production volume and its wide range of uses, including contributions to food security, animal feed, and agro-industrial supply chains, as well as its broad adaptability to diverse agroecological conditions (Erenstein et al., 2022; Tanumihardjo et al., 2020). According to USDA (2025b), global maize production currently ranges between 1.22 and 1.30 billion metric tons per year, with production concentrated mainly in the United States (31%), China (24%), and Brazil (11%), followed by the European Union (5%), Argentina (4%), and India (3%). In this context, Brazil ranks as the world's third-largest maize producer, with an estimated production between 110 and 130 million metric tons, driven primarily by the expansion of second-season maize ("safrinha"), which has further strengthened its relevance in tropical agricultural systems and international grain markets (Duarte et al., 2021; USDA, 2025b).

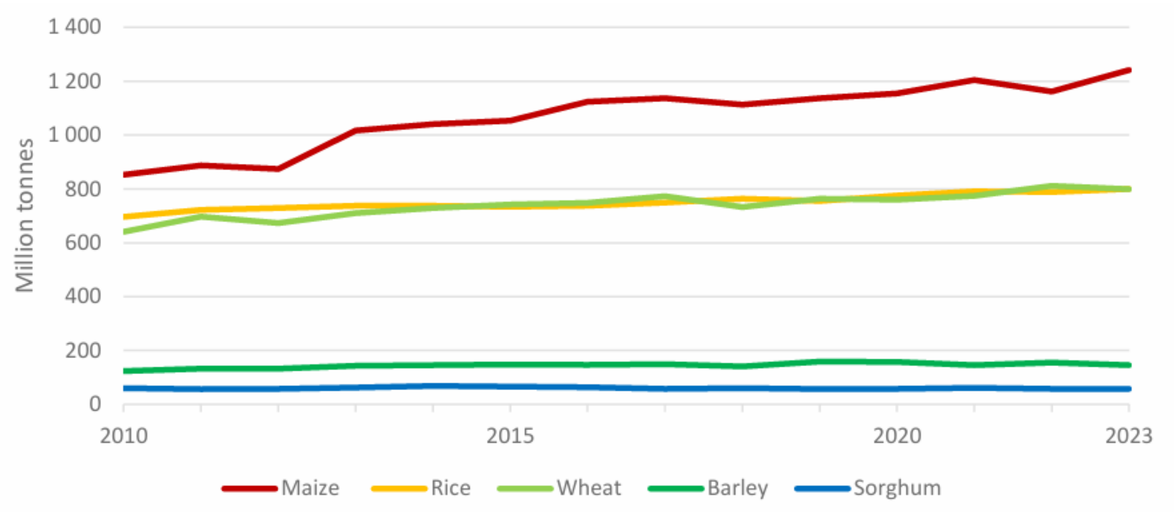


Fig. 1 World production of top cereals. **Source:** *FAO (2024),* Agricultural production statistics 2010–2023, FAOSTAT Analytical Brief No. 96.

Maize productivity can be significantly affected by a wide range of diseases of fungal, bacterial, viral, and nematode origin, which represent some of the major limiting factors for yield worldwide (Aegro, 2023; Agrios, 2005a). In Brazil, numerous diseases affect maize; however, foliar diseases are among the most relevant from an agronomic and economic perspective. Among these, gray leaf spot (*Cercospora zeae-maydis*), maize white spot (*Phaeosphaeria maydis*), and northern leaf blight or turcicum leaf blight (*Exserohilum turcicum*) stand out as the most important foliar diseases affecting the crop (Aegro, 2023; Carvalho et al, 2016; Fantin and Duarte, 2009).

Maize white spot is considered one of the most important foliar diseases in tropical and subtropical maize-growing regions, with a wide distribution across major maize-producing areas in Brazil and worldwide (Moreira et al., 2009). The disease is prevalent in countries of Central and South America, Asia, and Africa and has also been reported more recently in (Carson, 2005; Costa et al., 2017; Xing et al., 2023). Under favorable environmental conditions and in susceptible cultivars, yield losses associated with MWS may exceed 50% (De Rossi et al., 2017; Embrapa, 2021; Xing et al., 2023).

Regarding the etiology of maize white spot, research has reported divergent findings. Initially, in Brazil, the disease was referred to as *Phaeosphaeria* leaf spot (PLS) and attributed to the fungal pathogen *Phaeosphaeria maydis* (Rane et al., 1966; Carson, 2005; Fantin, 2009; Fantin and Balmer, 1997). However, subsequent studies identified the bacterium *Pantoea ananatis* as a potential causal agent (Paccola-Meirelles et al., 2001; Bomfeti et al., 2007; Gonçalves et al., 2013). Due to these contrasting findings, the disease became more broadly referred to as Maize White Spot (MWS). To date, there is no definitive consensus regarding the primary etiological agent, which has limited the development of standardized management strategies (Xing et al., 2023).

Symptoms of MWS initially appear as small pale-green or chlorotic areas that progressively expand and develop a whitish, dry appearance, typically delimited by brown margins. Lesions may be round, oblong, elongated, or slightly irregular in shape, with well-defined margins, and range from 0.3 to 2.0 cm in length, being distributed across the leaf surface. In general, symptoms first appear on the lower leaves and rapidly progress toward the upper canopy (Carvalho et al., 2016; Fantin and Duarte, 2009).

MWS exhibits a polycyclic disease development, characterized by successive infection cycles and rapid disease spread once established in the crop. The inoculum can persist in crop residues, volunteer plants, or infected plant material, constituting an important source of primary infection in subsequent cropping cycles, and no alternative hosts have been clearly identified to date (Silva et al., 2001; Xing et al., 2023). Infection may occur both before and after flowering; however, the highest incidence and severity are typically observed during the grain-filling or milk-ripening stage (Xing et al., 2023).

In general, the management of MWS is based on an integrated approach that includes the use of cultivars with higher levels of resistance, the adoption of cultural practices aimed at reducing initial inoculum, such as residue management and crop rotation, and, in some cases, the application of fungicides. Nevertheless, the effectiveness of these measures is variable and strongly dependent on the production system and prevailing environmental conditions. Consequently, disease management strategies are primarily focused on prevention and on mitigating yield losses rather than on curative control (Fantin and Duarte, 2009; Xing et al., 2023).

1.2.2 Agrometeorological conditions and plant disease development

The development of plant diseases is closely determined by the interaction between the host, the pathogen, and the environment, forming the well-known disease triangle. In this context, environmental variables, particularly agrometeorological conditions, play a decisive role in the occurrence, intensity, and temporal dynamics of epidemic processes (Agrios, 2005; Bedendo et al., 2019).

Among the most relevant agrometeorological elements are air temperature, relative humidity, soil moisture, precipitation, solar radiation, wind, and leaf wetness duration, all of which can influence both pathogens and host plants. These factors affect in the different phases of the disease cycle, including pathogen survival, dissemination, infection, colonization, and reproduction, as well as in host plant growth, physiological status, and susceptibility (Bedendo et al., 2019; Hossain et al., 2024).

Air temperature is one of the most critical agrometeorological elements in the development of plant diseases, as it regulates pathogen growth, reproduction, virulence, and geographic distribution, while also influencing the expression of host defense mechanisms. Similarly, precipitation and atmospheric humidity play a fundamental role in spore germination, infection, and pathogen dispersal, particularly in the case of fungi and oomycetes, thereby increasing the risk of epidemics when prolonged periods of favorable moist conditions occur. Wind, in turn, acts as an important dispersal agent, facilitating the movement of pathogens over long distances and contributing to the emergence and spread of plant diseases (Bedendo et al., 2019; Hossain et al., 2024).

Soil moisture significantly influences the development of plant diseases, as high water availability favors the survival, germination, sporulation, and dispersal of numerous pathogens, particularly soilborne fungi and oomycetes, in addition to facilitating root infections. In contrast, water deficit conditions may reduce pathogen activity, although many organisms are able to survive through resistant structures. In the host plant, excessive soil moisture can cause physiological stress, reduced root oxygenation, and nutritional imbalances, thereby increasing susceptibility to diseases, whereas drought limits plant growth and weakens defense mechanisms. Overall, soil water status plays a key role in regulating disease development and severity throughout the disease cycle (Bedendo et al., 2019).

Leaf wetness duration is a key variable in the development of numerous plant diseases, as, in interaction with temperature, it determines the likelihood of infection. Prolonged leaf wetness favors fundamental pathogen processes such as spore germination, penetration into host tissues, and, in some cases, inoculum release, thereby acting as a triggering factor for epidemic outbreaks (Gillespie and Sentelhas, 2008).

Solar radiation or light acts primarily as an indirect factor in the development of plant diseases. In pathogens, radiation may influence the survival, viability, and activity of infective structures, generally in interaction with other environmental factors. In the host plant, light directly affects photosynthesis, growth, nutrition, and defense mechanisms; therefore, variations in radiation intensity or duration can modify the degree of susceptibility to diseases (Bedendo et al., 2019).

In the case of maize white spot (MWS), several studies have examined the influence of agrometeorological conditions on disease progression and severity, with particular emphasis on variables such as precipitation, temperature, and relative humidity. These studies indicate that MWS occurrence is closely associated with the accumulation of favorable thermal conditions combined with regular precipitation throughout the crop cycle, as well as with prolonged periods of leaf surface wetness and high relative humidity levels. Overall, the persistence of favorable temperature and moisture conditions explains the recurrent occurrence and increased severity of the disease in regions or periods characterized by consistent rainfall and moderate temperatures (Codognoto et al., 2017; Rolim et al., 2007; Xing et al., 2023).

1.2.3 Artificial intelligence in plant disease prediction

In agricultural production systems, plant disease management is a key determinant of crop productivity and, at the same time, one of the major challenges for reducing negative impacts on natural resources. The historical reliance on control strategies based on the application of plant protection products has raised increasing environmental, economic, and social concerns, which has driven the search for more sustainable approaches to crop health management. In this context, digital agriculture tools and artificial intelligence have emerged as fundamental support for decision-making (Jensen et al., 2025).

Artificial intelligence refers to the development of computational systems that imitate behaviors associated with human intelligence through interaction with their environment, the processing of input information, and the generation of outputs with performance comparable to that of humans in specific tasks (Zha, 2020). Machine learning constitutes one of its main branches and focuses on the development of models that learn relationships from data without being explicitly programmed (Candel, 2022). Within the field of phytopathology, machine learning has gained particular relevance due to its ability to model complex and nonlinear relationships among multiple environmental, biological, and management-related variables. Recent reviews highlight that these methods are especially well suited for plant disease prediction, as they enable the integration of large volumes of data associated with the host, the pathogen, and the environment, particularly in contexts characterized by high climatic variability and spatial heterogeneity (Fenu and Malloci, 2021; Jensen et al., 2025).

The application of machine learning models for plant disease prediction has increased markedly over the past two decades and has been widely used to estimate infection risk, disease occurrence, and disease severity across different spatial and temporal scales. These models allow the simultaneous integration of environmental and phenological information, as well as, in some cases, data derived from remote sensing or in-field monitoring systems. In particular, the incorporation of agrometeorological variables is key to disease prediction, as these variables directly regulate pathogen infection, development, and dispersal processes, and determine the spatiotemporal dynamics of epidemic outbreaks (Fenu and Malloci, 2021; Jensen et al., 2025).

Among the most commonly used approaches in plant disease prediction are support vector machines (Support Vector Machine, SVM), artificial neural networks (Artificial Neural Networks, ANN), Random Forest (RF) models, and, more recently, deep learning techniques such as convolutional neural networks (Convolutional Neural Networks, CNN) and long short-term memory models (Long Short-Term Memory, LSTM). According to recent reviews, SVM- and RF-based models have generally shown competitive performance across different plant disease prediction contexts (Fenu and Malloci, 2021; Jensen et al., 2025).

Overall, the literature indicates a sustained growth in the application of machine learning models for plant disease prediction, highlighting their potential to integrate environmental, phenological, and production-related information across different scales. Nevertheless, existing studies show a strong concentration on a limited number of crops and pathogens, as well as a still limited incorporation of data from multiple information sources in an integrated manner (Fenu and Malloci, 2021). In the case of foliar diseases of maize, and particularly white spot disease, studies combining machine learning approaches with agrometeorological variables remain scarce, reflecting an emerging line of research within the field of digital phytopathology.

References

Aegro (2023) Doenças do milho: sintomas e controle. Available at: https://blog.aegro.com.br/doencas-do-milho/

Ageitec (2021) Doenças Foliares. Embrapa. Available at: https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/milho/producao/pragas-e-doencas/doencas/doencas-foliares

Agrios, GN (2005) Plant Pathology. 5th ed. Elsevier, Amsterdam.

Amorim L, Rezende JAM, Bergamin Filho A, Camargo LEA (2016) Manual de Fitopatologia volume 2: Doenças das Plantas Cultivadas. 5a Edição, 772.

Bedendo IP, Amorim L, Mattos-Jr D (2019) Ambiente e Doença. In.: Amorim L, Marques JA, Bergamin A (Eds.) **Manual de Fitopatologia Princípios e Conceitos**. Ouro Fino: Agronômica Ceres, p. 93-103.

Bomfeti CA, Meirelles WF, Souza-Paccola EA, Casela CR, Ferreira AS, Marriel IE, Paccola-Meirelles LD (2007) Avaliação de produtos químicos comerciais, in vitro e in vivo, no controle da doença foliar, mancha branca do milho, causada por *Pantoea ananatis*. **Summa Phytopathologica** 33: 63-67. https://doi.org/10.1590/S0100-54052007000100009

Carson ML (2005) Yield loss potential of Phaeosphaeria leaf spot of maize caused by *Phaeosphaeria maydis* in the United States. **Plant Disease** 89(9): 986-988. https://doi.org/10.1094/PD-89-0986

Codognoto LC, Conde TT, Cipriano LP, Cipriano TL, Azevedo BA, Ferreira LR (2017) Mudanças climáticas e doenças na cultura do milho. **Revista Edutec** 2(1): 1-12.

Cruz JC, Pereira Filho IA, Pereira Duarte A (2021) Milho Safrinha. Agência Embrapa de Informação Tecnológica. Available at: https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/milho/producao/sistemas-diferenciais-de-cultivo/milho-safrinha

Costa RV, da Silva DD, Cota LV (2017) Evolução dos sintomas da mancha-branca no campo. Embrapa.

Silva HP, Fantin GM, Resende IC, Pinto N, Carvalho RV (2001) Manejo integrado de doenças na cultura do milho de safrinha. In Embrapa Milho e Sorgo-Artigo in **anais** de congresso (ALICE) Seminário Nacional de Milho Safrinha 6:113-144.

Carvalho RV, Pereira OAP, Camargo LEA (2016) Doenças do Milho. In.: Amorim L, Rezende JAM, Bergamin Filho A, Camargo LEA (Eds.) **Manual de fitopatologia: Doenças das plantas cultivadas**. Ouro Fino: Agronômica Ceres, p. 549-560.

Rossi RL, Gimenez Pecci MP, Guerra FA, Plazas MC, Brücher E, Guerra GD, Torrico Ramallo AK, Camiletti BX, Maurino MF, Barontini JM (2017) Enfermedades del maíz de siembra tardía causadas por hongos. In: El mismo maíz, un nuevo desafío: Compendio primer congreso de maíz tardío. Dow Agrosciences, p. 109–127.

Duarte AP, Tsunechiro A, Freitas R (2021) 3 décadas de inovações na cultura do milho safrinha: avanços e desafios. Campinas: Instituto Agronômico.

Erenstein O, Jaleta M, Sonder K, Mottaleb K, Prasanna BM (2022) Global maize production, consumption and trade: trends and R&D implications. **Food Security** 14(5): 1295-1319.

Fantin GM (2009) Mancha de *Phaeosphaeria* do milho [Comunicado técnico]. Secretaría de Agricultura e Abastecimiento, Estado de São Paulo. Available at: https://biologico.agricultura.sp.gov.br/publicacoes/comunicados-documentos-tecnicos/comunicados-tecnicos/mancha-de-phaeosphaeria-do-milho

Fantin GM, Balmer E (1997) Método de inoculação e evolução de sintomas da mancha foliar de *Phaeosphaeria maydis* em milho. **Summa Phytopathologica** 23(1): 64–65.

Fantin GM, Duarte A (2009) Manejo de doenças na cultura do milho safrinha. Campinas: Instituto Agrônomico.

FAO (2022) World Food and Agriculture – Statistical Yearbook 2022. Rome. https://doi.org/10.4060/cc2211en

FAO (2024) Agricultural production statistics 2010–2023. FAOSTAT Analytical Briefs, No. 96.

Fenu G, Malloci FM (2021) Forecasting plant and crop disease: an explorative study on current algorithms. **Big Data and Cognitive Computing** 5(1): 2. https://doi.org/10.3390/bdcc5010002

Gillespie TJ, Sentelhas PC (2008) Agrometeorology and plant disease management: a happy marriage. **Scientia Agricola** 65: 71–75. https://doi.org/10.1590/S0103-90162008000700012

Gonçalves RM, Figueiredo JEF, Pedro ES, Meirelles WF, Leite RP, Sauer AV, Paccola-Meirelles LD (2013) Etiology of Phaeosphaeria leaf spot disease of maize. **Journal of Plant Pathology** 95: 559–569.

González-Domínguez E, Caffi T, Rossi V, Salotti I, Fedele G (2023) Plant disease models and forecasting: changes in principles and applications over the last 50 years. **Phytopathology®** 113(4): 678–693. https://doi.org/10.1094/PHYTO-10-22-0362-KD.

Hossain MM, Sultana F, Mostafa M, et al. (2024). Plant disease dynamics in a changing climate: impacts, molecular mechanisms, and climate-informed strategies for sustainable management. **Discover Agriculture** 2(1): 132.

Jensen A, Brown P, Groves K, Morshed A (2025) Next generation crop protection: A systematic review of trends in modelling approaches for disease prediction. **Computers and Electronics in Agriculture** 234: 110245. https://doi.org/10.1016/j.compag.2025.110245.

Moreira JUV, Bento DAV, de Souza AP, de Souza CL (2009) QTL mapping for reaction to *Phaeosphaeria* leaf spot in a tropical maize population. **Theoretical and Applied Genetics** 119: 1361-1369. https://doi.org/10.1007/s00122-009-1140-0

Ortega-Candel JMO (2022) Big data, machine learning y data science en python. Ra-Ma Editorial. 406 p.

Silva RS, Campos HD, Ribeiro LM, Braz GBP, Magalhães WB, Bueno JN (2021) Danos na cultura do milho em função da redução de área foliar por desfolha artificial e por doenças. **Summa Phytopathologica** 46: 313–319.

Tanumihardjo SA, McCulley L, Roh R, Lopez-Ridaura S, Palacios-Rojas N, Gunaratna NS (2020). Maize agro-food systems to ensure food and nutrition security in reference to the Sustainable Development Goals. **Global Food Security** 25: 100327. https://doi.org/10.1016/j.gfs.2019.100327

USDA (2025a). Corn: World production [Production, Supply and Distribution database]. Foreign Agricultural Service. Available at: https://www.fas.usda.gov/data/production/commodity/0440000

USDA (2025b) World Agricultural Supply and Demand Estimates. Office of the Chief Economist, World Agricultural Outlook Board.

Wang G, Zhang X (2024) Maize in Global Food Security: Role and Challenges. **Field Crop** 7.

Xing E, Fan X, Jiang F, Zhang Y (2023) Advancements in Research on Prevention and Control Strategies for Maize White Spot Disease. **Genes** 14(11): 2061. https://doi.org/10.3390/genes14112061

Zha J (2020) Artificial intelligence in agriculture. **Journal of Physics: Conference Series** 1693(1): 012058. Doi:10.1088/1742-6596/1693/1/012058

CHAPTER 2 - Influence of agrometeorological variables on the severity of Maize White Spot and its effect on yield in tropical and subtropical regions[1]

ABSTRACT - Maize production is affected by several biotic and abiotic factors, with diseases being a major cause of yield and quality loss. The maize white spot (MWS) is one of the most important diseases and is widespread in tropical and subtropical production areas. Its development depends on the interactions among plants, pathogens, and environmental conditions. Understanding the environmental conditions that favor disease development is essential for predicting disease occurrence in different regions and for implementing effective crop management strategies. This study evaluated the relationship between agrometeorological variables and the severity of MWS, as well as the impact of the disease on maize yield under field conditions. The analysis was based on 25 years of MWS severity and yield data collected from 29 production sites in São Paulo, Brazil. Statistical approaches include correlation analysis, linear regression models, and pattern identification in the data distribution. The observed trend suggests that the maize yield decreases by approximately 45.7 kg ha^{-1} (≈ 1%) for each 1% increase in MWS severity. The variables most strongly associated with increased MWS severity are relative humidity (>55%), soil moisture (>0.3), mean temperature (15-22 °C), maximum temperature (20-30 °C), diurnal temperature range (8-16 °C), and wind speed (<3 m s^{-1}). These findings improve the understanding of the effects of the environment on disease progression, expand the knowledge of key factors and reveal new associations. The results provide valuable input for predictive models and management strategies to mitigate the impact of MWS on maize production.

Keywords: disease, *Phaeosphaeria* leaf spot, climate, relative humidity, temperature, *Zea mays*.

---

[1] This chapter corresponds to the scientific article published in the *International Journal of Biometeorology* 69:3613–3626, 202. https://doi.org/10.1007/s00484-025-03046-8

2.1 Introduction

Maize is one of the most produced grains in the world, with the United States, China, Brazil, Argentina and India being the major producers (FAO 2022; USDA 2024). Brazil, in particular, is the largest producer of maize in South America and the third largest producer in the world. It is also the world's leading exporter of this grain (USDA 2025) and is the second most cultivated crop in the country after soybeans (FAO 2022). Brazil's large production is due to its ability to harvest twice a year: in the summer season and in the second season, called "safrinha". The second season is sown between January and March after the first harvest, generally soybean (Cruz et al. 2021). This double production capacity allows Brazil to maximize land use (Duarte et al. 2021) and annual yield, and to satisfy both domestic demand and exports.

Maize productivity and quality are affected by biotic and abiotic factors, with diseases being a major cause of losses (Codognoto et al. 2017). *Phaeosphaeria* leaf spot (PLS), or maize white spot (MWS), is an important foliar disease in tropical and subtropical maize regions (Moreira et al. 2009) and occurs in Central and South America, Asia, and Africa (Carson 1999). While reported in the U.S., its impact is minimal due to the high resistance of commercial hybrids and the fact that warm and dry conditions in major production regions, such as the Midwestern maize belt, are not favorable for disease development (Carson 2005). Its damage varies with environmental conditions and plant developmental stage (Amorim et al. 2016). The disease was initially attributed to the fungus *Phaeosphaeria maydis* (Rane et al. 1966; Carson 2005; Fantin 2009; Fantin and Balmer, 1997). However, other studies have identified the bacterium *Pantoea ananatis* as a possible causal agent (Paccola-Meirelles et al. 2001; Bomfeti et al. 2007; Gonçalves et al. 2013), and some authors have even suggested that MWS may result from the combined action of multiple pathogens (Amorim et al. 2016). To date, no definitive consensus has been reached regarding the primary etiological agent of the disease, which has hindered the development of standardized management strategies (Xing et al. 2023).

In Brazil, MWS is a prevalent maize disease (Bosqui et al. 2018; Codognoto et al. 2017). Before the late 1980s and early 1990s, it appeared at the end of the growing cycle without causing significant damage. However, since then, MWS has become increasingly significant due to spatial and temporal migration of maize cultivation towards higher-altitude regions, characterized by milder temperatures. Although these conditions are favorable for maize productivity, they simultaneously create an environment conducive to proliferation of foliar diseases such as MWS (Fantin and Duarte 2009). Today, MWS is widespread in Brazil's maize-producing areas, potentially causing substantial yield loss (Fantin and Duarte 2009; Duarte et al. 2021).

According to Embrapa (2021), MWS can cause greater than 60% yield loss in Brazil under favorable conditions with susceptible cultivars. Xing et al. (2023) reported 10-50% losses in Southwest China, whereas Carson (2005) reported an 11-13% reduction in susceptible U.S. hybrids. Although specific data for other countries are lacking, increased disease frequency and severity have been reported in Argentina (De Rossi et al. 2017), and susceptible maize genotypes have been observed in South Africa (Adam et al. 2017).

Plant diseases result from the interaction of a susceptible host, a pathogen, and favorable environmental conditions, and are influenced mainly by meteorological factors (Agrios 20(FA. Sawazaki et al. (1997) reported that rainfall uniformity was the most critical factor in MWS development. Similarly, Fantin et al. (2006) reported that MWS severity tends to be greater during extended rainy periods in maize cultivation, particularly after the vegetative phase.

Several studies have examined both the meteorological drivers of MWS and its impact on maize yield. Meteorological conditions such as precipitation, temperature, and relative humidity play a key role in disease development: Rolim et al. (2007) found that MWS is favored by approximately 2,900 °C of accumulated degree days and around 350 mm of rainfall during the first five months of the crop cycle in São Paulo, while Codognoto et al. (2017) reported that the disease develops under relative humidity above 60% and moderate temperatures (15-20 °C), and Xing et al. (2023) highlighted the importance of seasonal fluctuations in temperature and humidity. Together, these findings indicate that sustained surface moisture and persistent favorable thermal and humidity conditions explain the recurrent occurrence of MWS in regions with regular rainfall and moderate temperatures. In parallel, multiple studies have demonstrated a negative effect of MWS on maize yield: Sawazaki et al. (1997)

reported a strong negative correlation between disease severity and grain weight across 30 cultivars, Pegoraro et al. (2001) found a significant correlation (r = -0.45) between severity and yield in six hybrids across five growing seasons, and Fantin et al. (2006) quantified yield losses ranging from 250 kg ha^{-1} at 5% severity to 1,930 kg ha^{-1} at 30%, confirming that increasing foliar severity leads to substantial yield reductions.

These losses were confirmed by Fantin et al. (2010), who found that each 1% increase in MWS severity (within 1-10%) leads to an approximate 2% yield reduction, while at higher severity levels losses follow a quadratic pattern and become proportionally smaller. From a physiological perspective, Godoy et al. (2001) showed that *Phaeosphaeria* leaf spot can reduce net photosynthesis by up to 40% at 10-20% severity, with transpiration decreasing in proportion to the loss of healthy tissue; however, quantitative evidence of these effects at high infection levels remains limited.

Although a few studies have addressed the impact of MWS on maize production, research directly linking its severity to yield remains limited. Even fewer investigations have correlated meteorological variables with severity levels, and most of the available studies are either outdated, geographically restricted, or based on short-term observations. These limitations highlight the need for new research that integrates meteorological and disease occurrence data to better understand the effects of MWS severity on crop performance and the role of agrometeorological conditions in disease dynamics.

Therefore, the present work has two main objectives: (1) to investigate the trend of the relationship between the MWS severity and crop yield under field conditions and (2) to evaluate the influence of agrometeorological variables on MWS severity.

2.2 Materials and methods

2.2.1 Study area

This study was conducted in the state of São Paulo, southeastern Brazil, between 20°- 25°S and 44°-53°W (Fig. 1). The region features a wide range of altitudes, from sea level on the coast to over 2,000 m in the Serra da Mantiqueira. However, elevation variation is generally moderate, with about 85% of the territory between 300 and 900 m. The predominant landform is a relatively uniform plateau extending east (Nunes et al. 2023). According to Köppen's climate classification, São Paulo encompasses both tropical savanna (Aw) climates in the northern and western regions and humid subtropical (Cfa) climates in the central, southern, and higher-elevation areas (Alvares et al., 2013; Rolim, Camargo, et al., 2007). This climatic heterogeneity, within a single political unit, allows for the analysis of disease dynamics across distinct tropical and subtropical conditions.

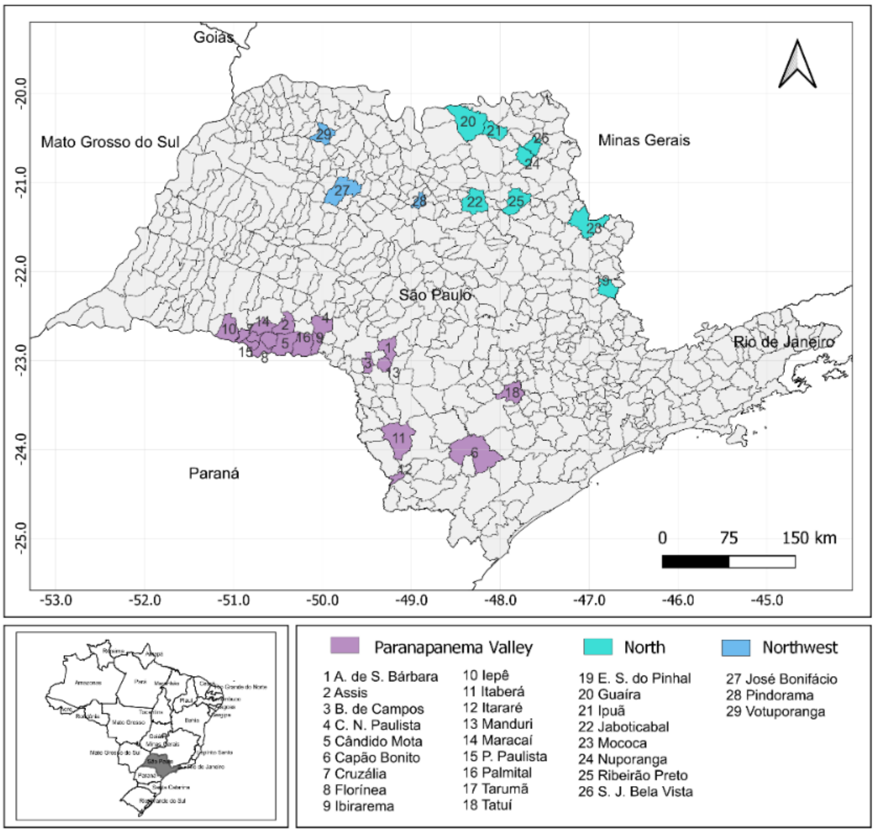


Fig. 1 Map of the state of São Paulo showing the geographical location of the municipalities where MWS severity assessments were conducted.

2.2.1.1 Disease data

The Maize Project, coordinated by IAC/APTA (Institute of Agronomy, São Paulo Agency for Agricultural Technology) with support from the Biological Institute (IB), has monitored maize pathogens and disease management since 1998. A key objective is to track severity fluctuations of major maize diseases in second season ("safrinha") cultivar trials conducted across various regions of São Paulo (IB 2024). This long-term effort has generated over two decades of valuable data, including disease severity (e.g., MWS), crop performance, and yield.

For this study, MWS severity and crop yield data were obtained from IAC/APTA Maize Project reports covering 1998–2022. The data originate from trials in 29 municipalities at altitudes of 400-850 m in key production regions of the Paranapanema Valley and North/Northwest (N/NW) São Paulo, spanning 1,5849 km^2 (Fig. 1). Fields were sown in February–March and harvested in July–August. Recent results are available in annual reports at https://zeamays.com.br/.

Between 15 and 40 trials were conducted annually, one or two per municipality, each involving 20 to 64 cultivars across production regions. All trials followed a randomized block design with four five-meter rows per plot and three to four replications. Disease was assessed at the final dough kernel stage using the Agroceres diagrammatic scale (Agroceres 1993), which ranges from 1 to 9 and corresponds to 0, 1, 2.5, 5, 10, 25, 50, 75, or more than 75% of affected leaf area (Fantin and Duarte 2021). Severity and yield for each hybrid were calculated as the means of all replications, resulting in a single severity score and average yield per trial. A fifth-order polynomial equation was used to convert scores to percentages. Analyses included plots without fungicide treatment and some with a single, generally preventive, application.

To perform this analysis, severity data were first grouped into three categories based on percentage values. The data were fitted to a gamma distribution, which best represents their behavior, high frequency at low values and tapering at higher ones (Wilks, 2011). Defined by shape and scale parameters, this distribution is ideal for modeling continuous, positive, and skewed variables such as MWS severity. The adjusted data were then divided into terciles: the lowest tercile was labeled "Resistant" (R), the middle "Moderately Resistant" (MR), and the highest "Susceptible" (S).

2.2.1.2 Weather data

Daily meteorological data were retrieved for each municipality for the June–August (JJA) quarters, corresponding to the disease assessment period, over the years 1998-2022. These data were sourced from the NASA POWER (Prediction of

Worldwide Energy Resources) database dataset (https://power.larc.nasa.gov/data-access-viewer/), which provides climate information derived from solar and meteorological inputs based on NASA research. The dataset supports applications in renewable energy, agriculture, and energy efficiency, and incorporates satellite observations and reanalysis models such as MERRA-2 (NASA 2021).

The following variables were included: mean, maximum, and minimum air temperature (°C), dew point temperature (°C), global solar radiation (MJ m^{-2} d^{-1}), relative humidity (%), precipitation (mm), wind speed (m s^{-1}), and soil moisture at the surface and root zone. All variables were retrieved at a spatial resolution of 0.5° latitude/longitude (approximately 55 km). Soil moisture is expressed as an index from 0 (completely dry) to 1 (fully saturated). Surface moisture refers to the top 0-5 cm of soil, while root zone moisture includes depths from the surface down to 100 cm.

Additionally, the following variables were calculated: reference evapotranspiration (mm) using the FAO Penman-Monteith method (Allen et al. 1998); vapor pressure deficit (kPa); diurnal temperature range (°C), defined as the difference between daily maximum and minimum temperatures; and a water availability index, calculated as the ratio between precipitation and evapotranspiration (dimensionless).

Daily data for all variables were aggregated on a quarterly scale (JJA) to perform correlation analysis and evaluate the intervals of influence on MWS severity.

2.2.2 Analysis of MWS severity and crop yield

To analyze the MWS severity and maize yield in the study region, we employed boxplot plots to visualize the distribution and variability of the data. In addition, to establish the relationship between these two variables, we performed simple linear regression analysis via ordinary least squares (OLS). The analysis focused on the slope coefficient (equation (*1*), and its standard error and statistical significance (< 0.05) to evaluate this relationship. Nonsignificant coefficients were left at zero since there is insufficient statistical evidence to claim that they differ from zero or that they significantly influence the mean value of the dependent variable.

The primary objective of this analysis was to examine the trend in the relationship between mean MWS severity (independent variable, X) and mean maize yield (dependent variable, Y) with the aim of comprehending how severity affects the production system, thereby reflecting the dynamics observed under field conditions. Importantly, the linear regression did not aim to formulate a predictive model but rather aimed to explore and characterize the interaction between the two variables, thereby providing a more nuanced perspective on the processes that influence the system.

$$\beta = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \quad (1)$$

where β is the slope coefficient, X is the independent variable, Y is the dependent variable, and $\bar{X}$ and $\bar{Y}$ represent the means of X and Y, respectively.

Regression analyses were performed for the entire state as well as for the Paranapanema Valley and North/Northwest regions, in addition to each severity group (R, MR and S). However, to compare the slope coefficients between severity classes, it was necessary to normalize the independent variable (X, severity %) within each group. This was due to the asymmetric distribution of severity (gamma distribution), which resulted in a difference in the amplitude or range of X among the three classes. This difference affects the calculation of β, which is directly associated with the variance of X. The greater the range of X is, the greater the variance of X, thereby influencing the denominator of the β calculation (equation (*1*).

2.2.3 Analysis of MWS severity and weather variables

To ascertain the relationship between MWS severity and meteorological variables, a Spearman correlation analysis was conducted, and its significance was evaluated at the 0.01 level, as this method captures nonlinear relationships and handles nonparametric data (MacFarland and Yates 2016).

The generation of boxplots was undertaken to facilitate the visualization of MWS severity across a range of meteorological variables, as well as by severity groups (R, MR, and S). This approach enabled the analysis of the influence of each variable on severity and the establishment of relevant climatic thresholds at varying severity levels. Additionally, the boxplots contributed to the identification of disease distribution according to climatic ranges and severity levels.

Simple linear regressions were also performed between MWS severity (now dependent variable, Y) and meteorological variables (independent variable, X), with the objective of analyzing their relationships by means of the slope coefficient, standard deviation and statistical significance. For these regressions, the medians of the ranges of the meteorological variables represented in the boxplot graphs were used, along with the mean severity corresponding to each median. Similarly to the severity and yield analysis, nonsignificant coefficients were assigned a value of zero.

2.3 Results and discussion

A database containing information on MWS severity was consolidated for the main maize-producing regions in the state of São Paulo, covering the period from 1998-2022. This database includes the municipalities of the Paranapanema Valley and the North and Northwest zones of the state and integrates yield data. Severity scores were transformed into percentages of affected leaf area using the adjusted polynomial equation (Equation (*2*). Data were grouped into terciles based on the gamma distribution to balance group sizes: resistant (R; 0-0.4%), moderately resistant (MR; 0.41-2.0%), and susceptible (S; >2.0%). Although the scale extends to 100%, the threshold for group S was set at 2.0% because most data (66%) fell below this value.

$$\text{MWS severity (\%)} = -0.0158\,nota^5 + 0.3031\,nota^4 - 1.7098\,nota^3 + 3.8741\,nota^2 - 2.4016\,nota - 0.0378 \qquad (2)$$
$$R^2 = 0.9994$$

2.3.1 Yield response to MWS severity in maize crops

During the evaluation period, the average maize yield in São Paulo State was estimated at 5,135 kg ha^{-1}, lower than typical summer crop yields due to unfavorable weather conditions, particularly during reproductive stages. Reduced rainfall and heat availability from April-May significantly contribute to this reduction. Average yields are similar across regions, though variability differs notably. In the Paranapanema Valley, yield averaged 5,117 kg ha^{-1} with higher variability (coefficient of variation, CV=32%; standard deviation, SD=1,656 kg ha^{-1}). In contrast, the North/Northwest region had slightly higher stability, with an average yield of 5,218 kg ha^{-1} (CV=25%; SD=1,280 kg ha^{-1}) (Fig. 2A). While climatic conditions likely influence yield differences, average

MWS severity and its variability were notably higher in the Paranapanema Valley, with more outliers than in the North/Northwest region (Fig. 2B). This increased disease intensity potentially contributes to greater yield variability, adding uncertainty to final yields.

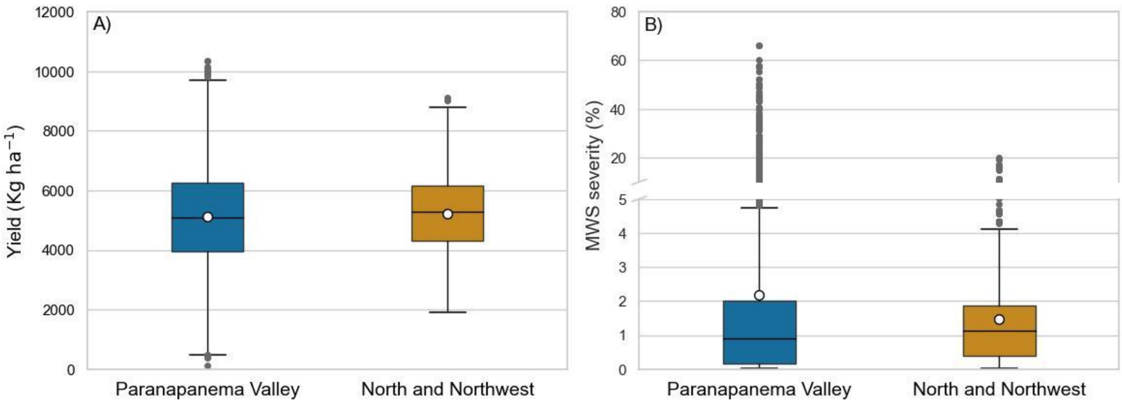


Fig. 2 Maize yield and MWS severity in two regions of São Paulo State. A) Maize yield and B) MWS severity (%) of the maize white spot (MWS) for the two regions in the state of São Paulo. Symbols: ○mean, ─ median, □ 50% of data, ┬ ┴ 99% of data, ● outliers.

Severity analysis by group (Fig. 3) indicated higher values in the valley (S=6.70%, MR=1.15%, R=0.12%) compared to the North/Northwest region (S=3.71%, MR=1.20%, R=0.16%). In the Paranapanema valley, average maize yield decreased progressively from the R group to the MR and S groups, with the greatest reduction observed in the S group, and the yield variability was slightly higher in the R and MR groups than in the S group (Appendix A - Fig. S1A), possibly due to factors other than MWS severity, such as cultivar adaptation to environmental conditions. However, the clear trend of reduced yields at higher severity levels highlights the importance of MWS severity as a limiting factor in the production system of the region.

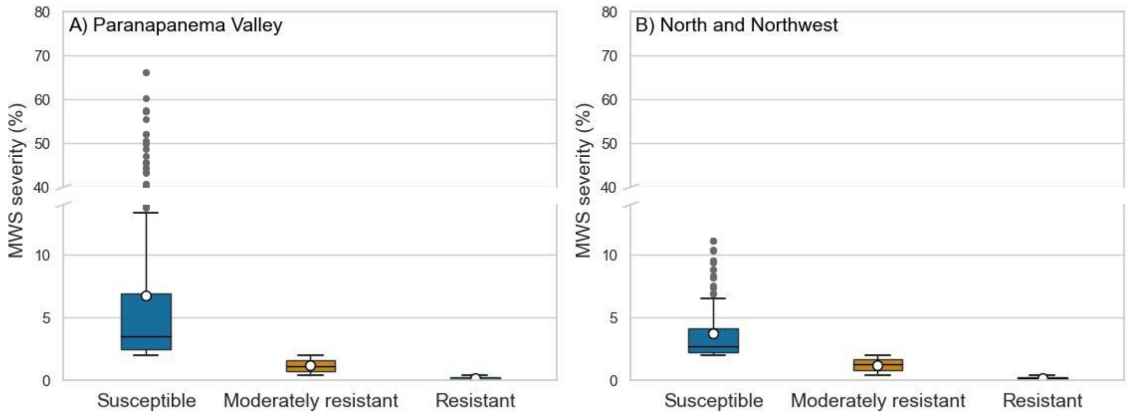


Fig. 3 MWS severity for two regions of the state of São Paulo. Symbols: ○mean, ─ median, □ 50% of data, ┬ ┴ 99% of data, ● outliers.

In the North/Northwest region, average yields and variability were similar across severity groups, with the R group slightly higher, but the S group yield marginally exceeded that of MR (Appendix A - Fig. S1B). This pattern contrasts with expectations that productivity would decrease with increasing MWS severity. Given the lower severity percentages in this region, other factors such as insect-transmitted diseases, including maize bushy stunt and viral infection, may have exerted a greater impact in recent years (Fantin and Duarte, 2021), and likely may have had a stronger influence on yield. Additionally, climatic events, such as droughts or rainfall shortages, which are common in this region, may reduce yield more significantly than MWS, as they not only limit disease development but also negatively affect plant growth and productivity (Bergamaschi et al. 2004).

In both regions, yield variations may be influenced by other factors, considering the wide range of variables involved in field conditions, as well as the productive adaptation of each cultivar to its environment and the differences in productive potential among hybrids. Nevertheless, the higher severity percentages in the valley are indicators of more favorable conditions for the development of MWS in this region. In the North/Northwest, all municipalities register values below 10% severity, except for Mococa and Votuporanga (Appendix A - Fig. S2B). In contrast, in the valley, the severity considerably exceeds this threshold in most of the municipalities (Appendix A - Fig. S2A).

For the analysis of the relationship between the percentage of MWS severity and crop yield, the slope coefficient (β) and confidence interval (standard error) of the linear regression were considered (Fig. 4). Considering the two regions, Paranapanema Valley and North/Northwest, the coefficients were -43.31 and -52.74 kg ha^{-1}, respectively, for each 1% severity (Fig. 4). This means that for each 1% increase in MWS severity, the crop yield tends to decrease by approximately 43.31 kg ha^{-1} and 52.74 kg ha^{-1}, respectively.

Statewide, each 1% increase in severity results in an average yield loss of 45.70 ± 2.38 kg ha^{-1} (Fig. 4A). Extrapolating this, a 10% severity corresponds to a yield reduction of approximately 457 kg ha^{-1} (8.9% of the average yield of 5,135 kg ha^{-1}), and a 20% severity to 914 kg ha^{-1} (17.8%). In the Paranapanema Valley, the average reduction is 43.31 ± 2.85 kg ha^{-1} per 1% severity (Fig. 4B), and in the North/Northeast region, 52.74 ± 3.82 kg ha^{-1} (Fig. 4C). A 20% severity would reduce yields by around 866 kg ha^{-1} (16.9%) in the valley and 1055 kg ha^{-1} (20.2%) in the North/Northeast.

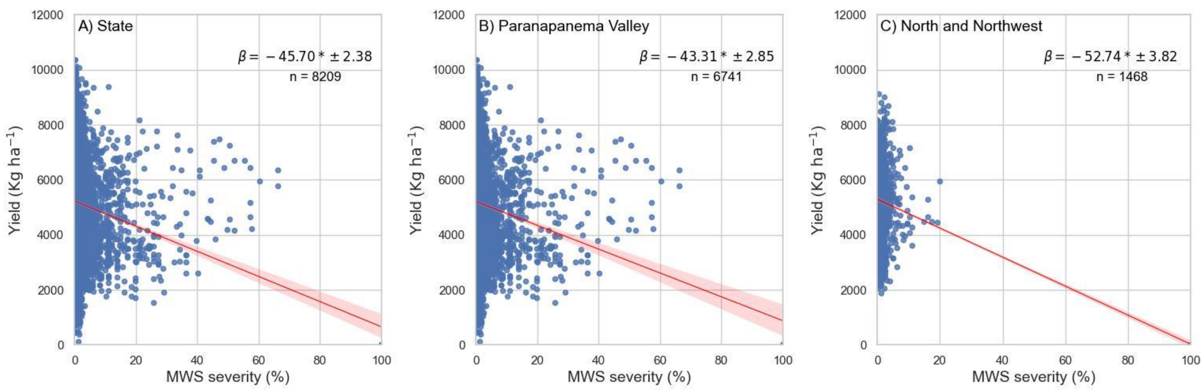


Fig. 4 Dispersion of yield and MWS severity data. Symbols: * indicates significance at the 0.05 level.

Our results showed that each 1% increase in MWS severity led to an average maize yield reduction of approximately 1%. Fantin et al. (2010), studying the relationship between severity and yield in São Paulo, reported a 2% reduction per 1% increase in severity. This difference can be explained by the severity ranges considered: we analyzed the full 0-100% range, while Fantin et al. (2010) focused on 1-10% and used only experiments with a significant correlation between severity and productivity, grouped into severity classes. This selective approach minimized the variability of both the independent (X, severity %) and dependent (Y, yield kg ha^{-1}) variables, resulting in a higher slope coefficient and coefficient of determination (R^2).

Some authors have reported in plants evaluated at the dough kernel stage that a 1% affected leaf area (in so-called resistant cultivars) causes damage of approximately 250 kg ha^{-1}, which is considered not significant and does not justify the use of fungicides. A severity of 2.5% (moderately resistant) typically results in average damage of 500-600 kg ha^{-1}, which is generally considered the threshold for the application of fungicides. At a severity level of 5% (moderately susceptible), the damage typically reaches approximately 800-900 kg ha^{-1}, thereby suggesting its utilization. However, when the severity reaches 10% or more (susceptible), the damage is generally substantial, rendering its application indispensable (Fantin et al. 2010; Fantin et al. 2015).

Several studies have correlated the severity of MWS with maize yield by comparing treatments with and without fungicide application. Although these studies did not perform regression analyses, the reported relationships between severity levels and productivity fall within the range of values observed in the present study.

Carvalho et al. (2023) conducted a study in Cafelândia, Paraná (Brazil), in 2022 to evaluate the impact of a fungicide mixture on maize white spot. The trial included 20 off-season maize cultivars with varying levels of resistance. The treatment reduced MWS severity from 20.76% to 4.79% and provided an average yield protection of 17.83%, with gains ranging from 0 to 111.31% among hybrids. Average yield increased from 8,412.5 kg ha^{-1} to 9,425.5 kg ha^{-1} with fungicide application.

Furthermore, in a study conducted at eight locations in 2020 including the states of Paraná, São Paulo, Goiás, and the Federal District Custódio et al. (2020) compared the control efficiency of ten fungicide mixtures using one susceptible hybrid per site. The average MWS severity was 21.2% in the control group and 6.3% with fungicide treatments, with control efficiency ranging from 62% to 81%. Yields increased from 6,621 kg ha^{-1} in the control to 7,524 kg ha^{-1} on average in treated plots, representing gains between 11% and 16%.

Fantin and Duarte (2009) reported that significant yield reduction typically begins at score three on the Agroceres scale (Agroceres 1993), corresponding to 2.5% of affected leaf area, or 2.2% according to our equation. Given that 2.5% is commonly considered the threshold for cost-effective fungicide use, the 2.0% lower limit adopted for the susceptible group (S) in our study offers a reasonable reference point for identifying the onset of potential yield reductions.

Because most severity data in this study are concentrated at lower levels, analyzing regression coefficients over a narrower severity range may be useful. Although susceptible cultivars can show high severity at the dough kernel stage, grain is still produced because severity was not high all the time. This is related to the concept of healthy leaf area duration described by Bergamin Filho and Amorim (1996). Additionally, the Agroceres scale shows limited differentiation at the upper end: score seven equals 75%, while score eight represents 75% or more. Due to the importance of this disease, breeding for resistance, focused on eliminating highly susceptible hybrids, has long been a priority, though somewhat reduced in the last decade with the development of more effective fungicides (Braga et al. 2021; Agrofit 2025). Nonetheless, as the severity scale spans from 0 to 100%, this study included all levels to provide an overall view of the disease's field impact.

At severity levels near zero, data show high dispersion (Fig. 4), limiting the ability to assess the specific impact of the disease on yield, particularly in low-yield cases. In this range, chance factors such as frost, drought, other diseases, and pests likely play a more prominent role than white spot itself. Conversely, in cases of high yield with low severity, favorable climatic conditions, cultivar adaptation, and hybrid potential may contribute. These variables also introduce random variability; thus, our focus was to identify the overall trend in the relationship between yield and MWS severity.

A comparison of slope coefficients among severity groups shows that the rates of yield reduction in the R and MR categories are substantially lower than in the S group, both statewide and in the Valley and North/Northwest regions (Table 1). These coefficients may vary depending on how the severity classes are defined. In our case, the S group includes values starting at 2% affected leaf area, and results indicate that severity above this threshold has a much stronger impact on yield, with a more rapid reduction. While this trend confirms the negative association between severity and yield, it is also possible that the rate of decline may stabilize or reach a plateau, as suggested by Fantin et al. (2010).

Table 2. Slope coefficients between crop yield and severity by resistance group (with the normalized ranges). Symbols: * indicates significance at the 0.05 level.

	Resistant	Moderately resistant	Susceptible
São Paulo State	$-2.17^{*} \pm 1.04$	$-3.95^{*} \pm 0.89$	$-36.40^{*} \pm 2.34$
Paranapanema Valley	$0.00^{NS} \pm 1.18$	$-4.43^{*} \pm 1.06$	$-30.25^{*} \pm 2.77$
North and Northwest	$-8.72^{*} \pm 1.94$	$-0.00^{NS} \pm 1.58$	$-52.11^{*} \pm 4.01$

2.3.2 Correlation between agrometeorological variables and MWS severity

Environmental conditions are the main disease-modulating factor when pathogens and nonresistant cultivars are present (Agrios 2005). When evaluating the relationships between agrometeorological variables and MWS severity, we found that the variables with the strongest correlations (greater than 0.3 in at least one region; Fig. 5) were relative humidity, surface soil moisture, root zone soil moisture, diurnal temperature range, maximum, mean, and minimum air temperatures, dew point, wind speed, vapor pressure deficit, and solar radiation. Relative humidity and soil moisture showed positive correlations, while all others were negative. Although these coefficients may be considered moderate in general terms, they are meaningful in the context of field-based plant disease studies, where multiple uncontrolled factors influence disease development (Kozak et al. 2012). All correlations were statistically significant at the 0.01 level.

The correlations of these variables were similar across both regions of the state, though slightly higher in the valley. An exception was observed for surface and root zone soil moisture, which were 0.05 and 0.07 greater, respectively, in the North/Northwest. Relative humidity and diurnal temperature range (DTR) showed the strongest associations: the former was positively correlated, and the latter negatively. The relative humidity correlations were 0.36 for the entire state, 0.46 for the valley, and 0.38 for the North/Northwest, while those for DTR were -0.42, -0.49, and -0.42, respectively (Fig. 5).

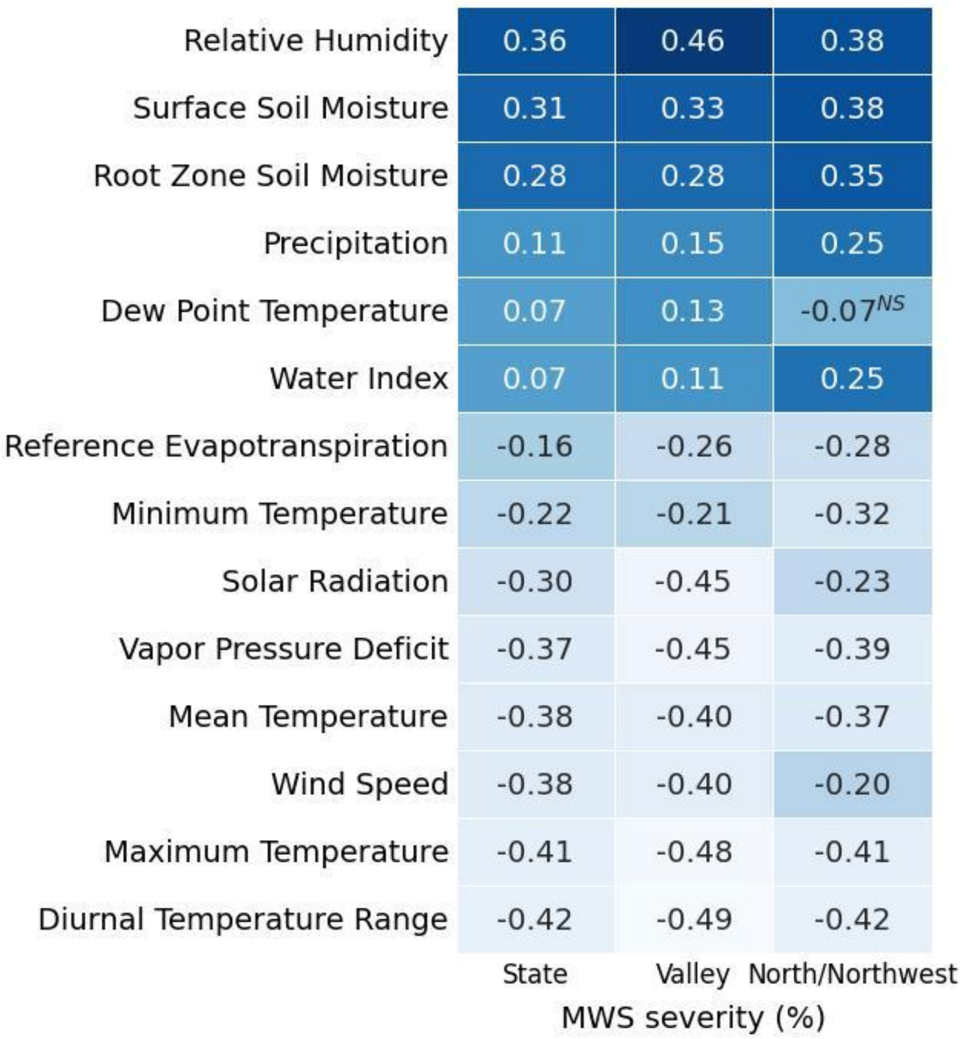


Fig. 5. Correlation analysis of climatic variables and MWS severity for the state of São Paulo, Paranapanema Valley and the north/northwest region. Symbols: ^{NS} correlation is not significant at the 0.01 level.

The weakest associations were found for precipitation, dew point temperature, and the water index (positive correlations), and for reference evapotranspiration and minimum temperature (negative correlations). All were statistically significant at the 0.01 level, except dew point temperature in the North/Northwest region.

Unlike the findings of Sawazaki et al. (1997) and Rolim et al. (2007), who identified precipitation as a key factor in white spot occurrence, this study did not observe a strong association between precipitation and MWS severity, although the correlation was statistically significant. Sawazaki et al. (1997) emphasized rainfall uniformity as the most relevant factor, while Xing et al. (2023) noted that white spot tends to occur in environments with consistent rainfall and moderate temperatures.

These findings suggest that variables such as the number of consecutive rainy days or cumulative rainfall in clustered events may be more informative than total rainfall over extended periods. Although combined variables or the use of extreme event indices may yield stronger correlations, they were analyzed separately in this study, which may have prevented more complex interactions from emerging. Future analyses could benefit from using compound or extreme event indices to better capture conditions favoring disease development.

2.3.3 Ranges of influence of agrometeorological variables on MWS severity

The onset and progression of foliar diseases such as MWS are strongly influenced by a complex interplay between pathogen biology, plant physiological status, and environmental conditions. Among these environmental factors, temperature and humidity are particularly critical, as they regulate both the physiological susceptibility of the host and the infection dynamics of the pathogen. Temperature influences all stages of pathogen development, while humidity plays a key role in spore release, dispersal, and infection processes (Gillespie and Sentelhas 2008).

In the case of MWS, our results indicate that severity is favored under moderately warm and humid conditions, characterized by relative humidity levels above 55% (

Fig. ***6***A) and mean temperatures ranging from 15 to 22 °C (Fig. 7C). These conditions likely support prolonged leaf wetness duration and increased stomatal opening, which together facilitate pathogen penetration and colonization. The highest severity levels were recorded at relative humidity above 60% and temperatures around 20 °C (

Fig. ***6***A, Fig. 7C), corroborating previous findings (Codognoto et al. 2017;Moreira 2004; Getap 2022). The DTR also showed a strong inverse association with MWS severity. Lower values (between 8 and 16 °C) were associated with higher severity, particularly in the susceptible group (Fig. 7A). Narrower temperature ranges during the day may create more stable conditions for pathogen activity. In addition, limited thermal fluctuation may reduce the plant’s capacity to perceive and respond to environmental stress, as key defense mechanisms are often activated by daily thermal variation. Under such conditions, prolonged stomatal opening and slower cuticle recovery could facilitate infection by foliar pathogens, as observed in other maize diseases such as *Cercospora zeae-maydis* (Beckman and Payne 1982; Gillespie and Sentelhas 2008).

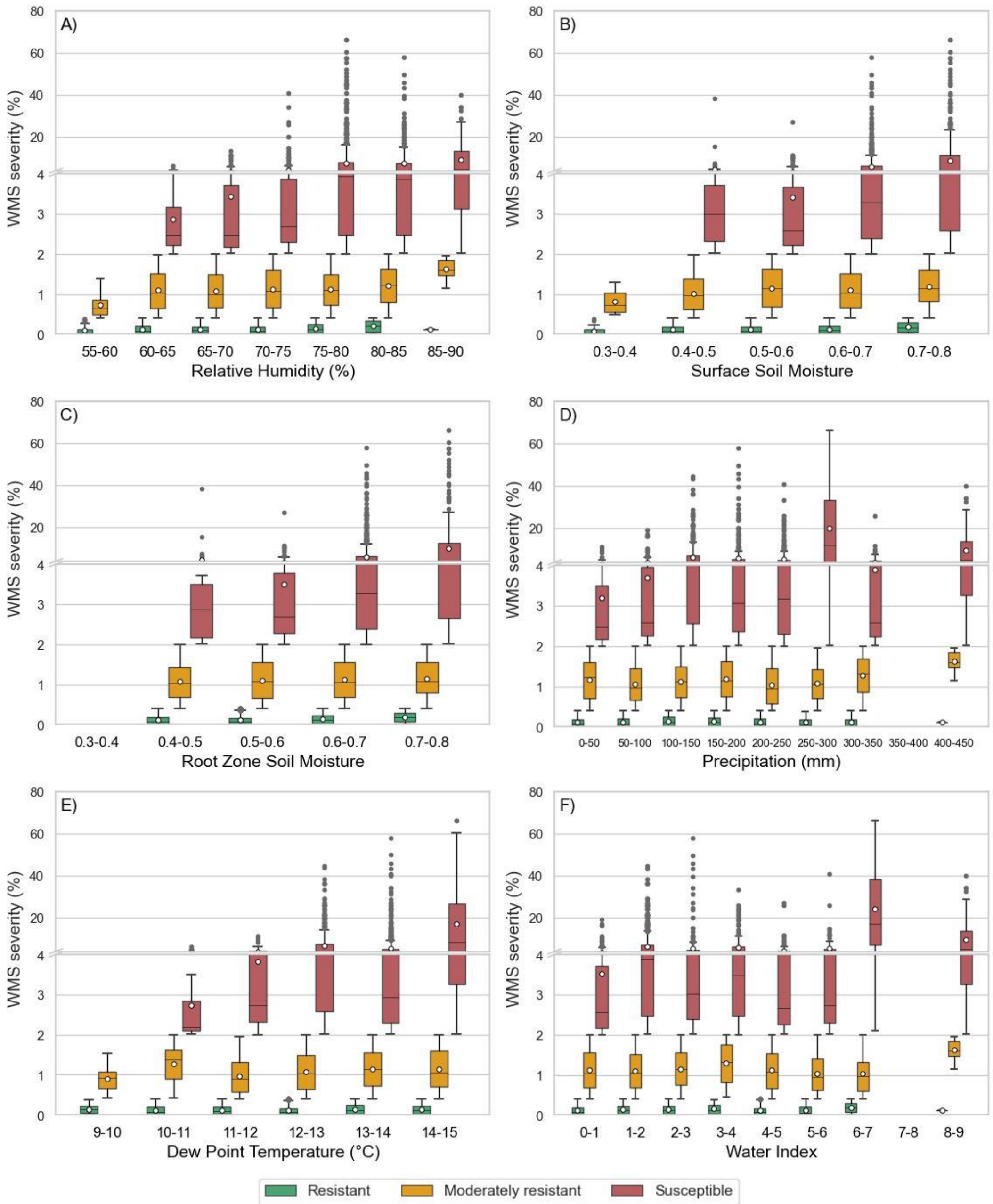


Fig. 6 The relationship between MWS severity and the weather variables with the positive correlations. The lower y-axis represents the severity percentage ranging from 0% to 4% and the upper y-axis covers the ranges from 4% to 100%. Symbols: ○mean, ─ median, □ 50% of data, ┬ ┴ 99% of data, ● outliers.

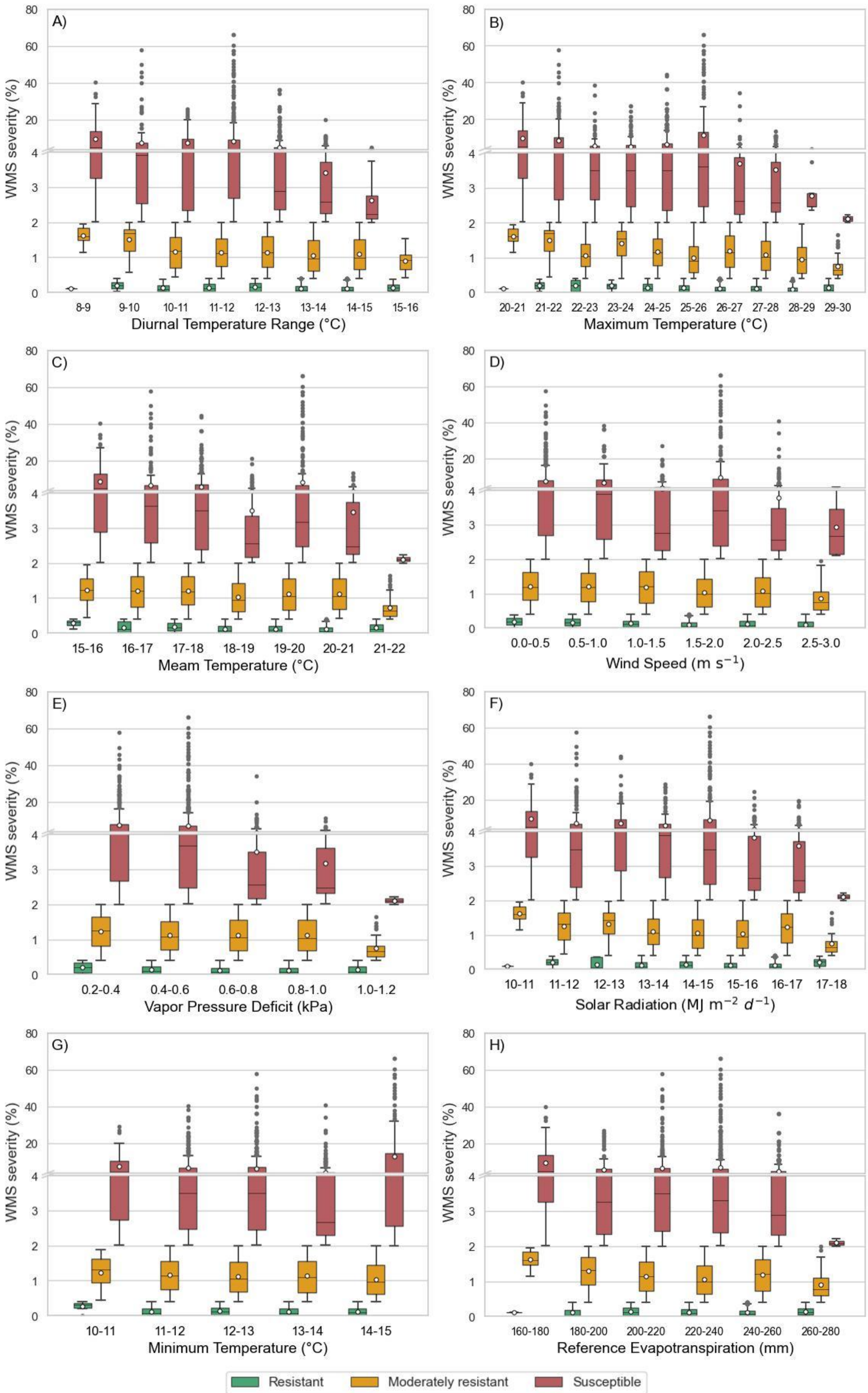


Fig. 7 Relationship between MWS severity and the weather variables with the negative correlations. The lower y-axis represents the severity percentage ranging from 0% to 4% and the upper y-axis covers the ranges from 4% to 100%. Symbols: ○mean, ─ median, □ 50% of data, ┬ ┴ 99% of data, ● outliers.

Likewise, soil moisture in both surface and root zones showed a positive association with MWS severity, particularly above 0.6 (

Fig. ***6***B and

Fig. ***6***C). This positive relationship may reflect the overall moisture status of the soil–plant–atmosphere system, where wetter soils typically indicate higher atmospheric humidity, which favors disease development. Moreover, increased root and surface moisture supports plant growth but also creates conditions conducive to pathogen development. Although no studies have specifically described the infection process of the white spot pathogen, it is known that higher soil moisture promotes stomatal opening in leaves, facilitating the entry of various pathogens, such as *Cercospora zeae-maydis* in maize (Beckman and Payne 1982).

Water availability in the crop environment plays a central role in the development of MWS, as it regulates both atmospheric humidity and the physiological state of the plant. Our results indicate that MWS severity increases under moderate rainfall, particularly between 100 and 300 mm over a three-month period (

Fig. ***6***D), suggesting that sustained canopy moisture, rather than extreme rainfall, creates favorable conditions for infection. This trend aligns with prior observations by Rolim et al. (2007), who found higher incidence with approximately 350 mm of accumulated rainfall, calculated from the sowing date. At the same time, severity tended to decline under higher reference evapotranspiration (ETo) values, likely due to increased atmospheric demand and reduced leaf surface moisture. The inverse relationship observed above 260 mm of ETo supports this interpretation (Fig. 7H).

The water index, which integrates both precipitation and ETo, further reinforces this pattern. Most severe cases occurred when the index exceeded one (

Fig. ***6***F), indicating a moisture surplus in the crop system. Taken together, these findings suggest that an intermediate hydric balance, sufficient to maintain relative humidity and leaf wetness without excess runoff or spore loss, is a key factor driving MWS severity under field conditions.

The combined effects of minimum temperature, dew point, and vapor pressure deficit (VPD) offer deeper insight into the role of atmospheric moisture in MWS development. MWS severity was most prominent when minimum temperatures ranged from 10 to 15 °C (Fig. 7G) and dew point temperatures exceeded 10 °C (

Fig. ***6***E), consistent with previous findings that associate MWS with moderate nighttime cooling and high relative humidity (Pinto 1997). These thermal and humidity conditions favor condensation on the leaf surface, sustaining leaf wetness and enhancing pathogen infection. Supporting this, we observed that MWS severity was highest under low VPD values (0.2 to 0.6 kPa) (Fig. 7E), indicating environments with low atmospheric demand and high moisture retention. As VPD increased, MWS severity decreased, suggesting that drier air limits the conditions necessary for pathogen establishment. Together, these variables illustrate how subtle shifts in atmospheric moisture dynamics can significantly influence disease progression.

Wind speed and solar radiation further modulate the microenvironmental conditions affecting MWS severity. MWS was observed under wind speeds below 3.0 m s^{-1}, with different severity levels across groups and a decreasing trend above 2.0 m s^{-1} (Fig. 7D). Such low wind conditions may reduce evaporation and help maintain humid air layers near the leaf surface, thereby enhancing the likelihood of infection. In contrast, higher wind speeds tend to accelerate drying of leaf surfaces, which may hinder pathogen establishment despite their known role in dispersal. Similarly, increased solar radiation, particularly above 17 MJ m^{-2} d^{-1} (Fig. 7F) was linked to lower severity, possibly due to elevated temperature and evapotranspiration leading to drier conditions. However, field observations suggest that sunlight may also promote lesion appearance, especially on more exposed leaf areas. This dual role highlights the complexity of radiation effects on disease expression and progression.

To complement the correlation analysis and the influence ranges of the agrometeorological variables, we calculated the slope coefficients from the linear regression with severity percentages (Appendix A - Table S1), which indicate the magnitude and direction of each variable's effect on the disease. Coefficients equal to zero reflect non-significant values, suggesting no relevant influence on mean severity. Among positively associated variables, only relative humidity (for MR and S groups) and surface soil moisture (for the R group) had significant coefficients. Variables with negative associations showed significant effects in at least one severity class.

Overall, the coefficients for the lower (R) and moderate (MR) severity groups were low, indicating limited influence of agrometeorological variables in these cases. In contrast, the higher severity group (S) was more sensitive to these variables. Notably, maximum temperature and wind speed had significant coefficients across all severity groups, suggesting both play important roles in disease dynamics.

The results also showed a clear pattern of reduced severity under higher temperature conditions, especially in the S group. Extreme values of maximum and mean temperature had a more pronounced effect, indicating reduced pathogen activity at elevated temperatures. Additionally, DTR ($\beta = -0.46$ in the S group) significantly contributed to severity reduction, highlighting the role of diurnal temperature variation as a key factor in disease management.

Finally, the coefficients for radiation and vapor pressure deficit, particularly in the S group, confirmed the inverse relationship with MWS severity, reinforcing how dry conditions and high radiation limit disease development. Although evapotranspiration showed significant coefficients in the MR and S groups, their low values suggest a limited effect on severity.

2.4 Conclusions

The analysis of yield and MWS severity indicated that white spot is a limiting factor for maize production in the Paranapanema Valley. Conversely, its impact is reduced in the North/Northeast region, likely due to the area's warmer and drier climate, which is less favorable for the pathogen. Quantitatively, the results show that maize yield decreases by approximately 45.7 kg ha^{-1} (≈ 1%) for every 1% increase in MWS severity, considering the full 0–100% severity range.

The variables most strongly associated with MWS severity were relative humidity, soil moisture, thermal amplitude, maximum and mean temperatures, wind speed, vapor pressure deficit, and solar radiation. Among them, relative humidity and diurnal temperature range showed the highest correlations, positive for the former and negative for the latter.

Humidity-related variables exhibited a direct relationship with severity, reinforcing that wetter conditions favor disease development. In contrast, thermal variables showed an inverse relationship, confirming that moderate temperatures are more conducive to white spot. Overall, the most favorable environmental conditions for disease occurrence include high relative humidity (>55%) and soil moisture (>0.3), moderate temperatures (15–22 °C), low wind speed (<3 m s^{-1}), and low vapor pressure deficit (0.2–1.2 kPa).

This study also contributes new insights by incorporating previously unexamined variables, particularly diurnal temperature range and soil moisture, which not only showed correlations with MWS severity but also enhanced our understanding of the environmental conditions that influence its development.

While this study provides valuable insights into the environmental factors associated with MWS severity and its effect on maize yield, some limitations should be noted. Factors such as soil properties, agronomic practices, extreme weather events, and/or co-infection with other diseases may also affect productivity but were not included in this analysis. In addition, the agrometeorological variables used represent mean conditions over a three-month period (June-August), aligned with the timing of disease assessment, rather than continuous monitoring at finer temporal scales. This temporal resolution may overlook short-term fluctuations critical to disease dynamics. Together with the lack of consensus on the primary causal agent of MWS, this limits the depth of biophysical interpretations involving environmental interactions. Nonetheless, the findings provide a strong foundation for future research focused on predictive modeling, disease management, and experimental design.

References

Adam E, Deng H, Odindi J, Abdel-Rahman EM, Mutanga O (2017) Detecting the early stage of phaeosphaeria leaf spot infestations in maize crop using in situ hyperspectral data and guided regularized random forest algorithm. **Journal of Spectroscopy** 2017:6961387. https://doi.org/10.1155/2017/6961387.

Agrios GN (2005) Plant pathology. 5th ed. Elsevier, Amsterdam.

AGROCERES (1993) Guia Agroceres de sanidade. Agroceres, São Paulo, 56 p.

AGROFIT (2025) Ministério da Agricultura, Pecuária e Abastecimento -Coordenação-Geral de Agrotóxicos e Afins. Available at: https://agrofit.agricultura.gov.br/agrofit_cons/principal_agrofit_cons.

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. FAO, Rome, 300 p.

Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G (2013) Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift** 22:711–728. https://doi.org/10.1127/0941-2948/2013/0507.

Amorim L, Rezende JAM, Bergamin Filho A, Camargo LEA (2016) Manual de fitopatologia: doenças das plantas cultivadas. 5ª ed. Agronômica Ceres, Viçosa.

Beckman PM, Payne GA (1982) External growth, penetration, and development of Cercospora zeae-maydis in corn leaves. **Phytopathology** 72:8 10–815. https://doi.org/10.1094/PHYTO-72-810.

Bergamaschi H, Dalmago GA, Bergonci JI, Bianchi CAM, Müller AG, Comiran F, Heckler BMM (2004) Distribuição hídrica no período crítico do milho e produção de grãos. **Pesquisa Agropecuária Brasileira** 39: 831-839.

Bergamin Filho A, Amorim L (1996) Doenças de plantas tropicais: epidemiologia e controle econômico. Agronômica Ceres, São Paulo.

Bomfeti CA, Meirelles WF, Souza-Paccola EA, Casela CR, Ferreira AS, Marriel IE, Paccola-Meirelles LD (2007) Avaliação de produtos químicos comerciais, in vitro e in vivo, no controle da doença foliar, mancha branca do milho, causada por *Pantoea ananatis*. **Summa Phytopathologica** 33: 63–67. https://doi.org/10.1590/S0100-54052007000100009

Bosqui PA, Ramari T, Silva C, Schmidt Filho E, Gasparotto F (2018) Efeito da época de aplicação de fungicida no controle de doenças foliares na cultura do milho safrinha. **Enciclopédia Biosfera** 15:168–178. https://doi.org/10.18677/EnciBio_2018A39.

Braga K, Fantin LH, Cateri MG et al (2021) Novos fungicidas foliares são eficientes no controle da mancha branca e promovem ganho de produtividade no milho: uma metanálise. **Anais** do XVI Seminário Nacional de Milho Safrinha. Instituto Agronômico, Assis, SP.

Carson ML (1999) Vulnerability of US maize germplasm to Phaeosphaeria leaf spot. **Plant Disease** 83:462–464. https://doi.org/10.1094/PDIS.1999.83.5.462.

Carson ML (2005) Yield loss potential of Phaeosphaeria leaf spot of maize caused by Phaeosphaeria maydis in the United States. **Plant Disease** 89:986–988. https://doi.org/10.1094/PD-89-0986.

Carvalho AG de Tonini V, Roy JMT, Pereira HV, Somavilla JC, Pereira VGC, Sulzbach A (2023) Severidade de mancha-branca em resposta da aplicação de fungicida e rendimento em híbridos de milho. **Anais** do 17º Seminário Nacional de Milho Safrinha, 28-30 de novembro de 2023, Campo Grande, MS. Maracaju: Fundação MS.

Codognoto LC, Conde TT, Cipriano LP, Cipriano TL, Silva BAA, Ferreira LR (2017) Mudanças climáticas e doenças na cultura do milho. **Revista Edutec** 2(1): 1-12.

Cruz JC, Pereira Filho IA, Duarte AP (2021) Milho safrinha. Agência Embrapa de Informação Tecnológica. Available at: https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/milho/producao/sistemas-diferenciais-de-cultivo/milho-safrinha.

Custódio AAP, Utiamada CM, Madalosso T et al (2020) **Eficiência de fungicidas no controle da mancha branca do milho - segunda safra 2020**. Boletim Técnico No. 96. IDR-Paraná, Londrina.

De Rossi RL, Gimenez Pecci M de la P, Guerra FA, Plazas MC, Brücher E, Guerra GD, Torrico Ramallo AK, Camiletti BX, Maurino MF, Barontini JM (2017) Enfermedades del maíz de siembra tardía causadas por hongos. **El mismo maíz, un nuevo desafío: Compendio primer congreso de maíz tardío**. Dow AgroSciences Argentina, Buenos Aires, pp 109-127.

Duarte AP, Tsunechiro A, Freitas R (2021) 3 décadas de inovações na cultura do milho safrinha: avanços e desafios. Instituto Agronômico, Campinas, SP.

Embrapa (2021) Doenças Foliares. Agência Embrapa de Informação Tecnológica). https://www.embrapa.br/Agencia-de-Informacao-Tecnologica/Cultivos/Milho/Producao/Pragas-e-Doencas/Doencas/Doencas-Foliares. Accessed 12 July 2024

Fantin GM, Balmer E (1997) Método de inoculação e evolução de sintomas da mancha foliar de Phaeosphaeria maydis em milho. **Summa Phytopathologica** 23(1): 64–65.

Fantin GM, Duarte A (2009) Manejo de doenças na cultura do milho safrinha. Instituto Agrônomico, Campinas, SP.

Fantin GM, Duarte A (2021) Evolução da ocorrência e intensidade de doenças fúngicas foliares e enfezamento/viroses no milho safrinha, em função da resistência das cultivares, durante 24 anos no estado de São Paulo. **Anais** do XVI Seminário Nacional de Milho Safrinha - 3 décadas de inovações: avanços e desafios. Instituto Agronômico, Assis, SP.

Fantin GM, Duarte AP, Cruz FA, Dudiendas C, Castro JL, Pereira JOF, Gallo PB, Bortoleto N, Ramos VJ, Ribeiro J L (2006) Efeito da Mancha de Phaeosphaeria sobre a Produtividade do Milho Safrinha no Estado de São Paulo. **Anais** Congresso Nacional de Milho e Sorgo (Vol. 26). ABMS - Embrapa Milho e Sorgo, Sete Lagoas, MG.

Fantin GM, Duarte AP, Dudienas C (2010) Quantificação de Danos causados por diferentes Níveis de Severidade de Doenças Foliares à Produtividade do Milho. **Anais** XXVIII Congresso Nacional de Milho e Sorgo (Ed.), XXVIII Congresso Nacional de Milho e Sorgo. Associação Brasileira de Milho e Sorgo, Goiânia, GO.

Fantin GM, Duarte AP, Sawazaki E (2015) Utilização de fungicidas e desempenho de híbridos em ambientes de produção do milho safrinha no Estado de São Paulo. **Anais** do XIII Seminário Nacional de Milho Safrinha - 30 anos de inovação em produtividade e qualidade. Universidade Estadual de Maringá (UEM), Maringá-PR.

FAO (2022) World food and agriculture - statistical yearbook 2022. FAO, Roma.

Getap (2022) Mancha-branca no milho. https://getap.agr.br/2022/08/11/mancha-branca-no-milho/. Accessed 11 September 2024

Gillespie TJ, Sentelhas PC (2008) Agrometeorology and plant disease management: a happy marriage. **Scientia Agricola** 65:71–75. https://doi.org/10.1590/S0103-90162008000700012.

Gonçalves RM, Figueiredo JEF, Pedro E dos S, Meirelles WF, Leite RP, Sauer AV, Paccola-Meirelles, LD (2013) Etiology of *Phaeosphaeria* leaf spot disease of maize. **Journal of Plant Pathology** 95: 559–569.

Godoy CV, Amorim L, Bergamin Filho A (2001) Alterações na fotossíntese e na transpiração de folhas de milho infetadas por Phaeosphaeria maydis. **Fitopatologia Brasileira** 26: 209-215. https://doi.org/10.1590/S0100-41582001000200017

IB (2024) Caracterização de patógenos do milho e manejo de doenças da cultura. APTA. http://www.sgp.apta.sp.gov.br/sgp.apta/projeto/caracterizao-de-patgenos-do-milho-e-manejo-de-doenas-da-cultura/2301. Accessed 14 July 2024

Kozak M, Krzanowski W, Tartanus M (2012) Use of the correlation coefficient in agricultural sciences: problems, pitfalls and how to deal with them. **Anais** Da Academia Brasileira de Ciências 84: 1147–1156.

MacFarland TW, Yates JM (2016) Introduction to nonparametric statistics for the biological sciences using R. Springer, Cham

Moreira JUV (2004) **Mapeamento de QTLs para reação à mancha de *Phaeosphaeria* em milho**. 100 f. Tese (Doutorado em Genética e Melhoramento de Plantas) -Universidade de São Paulo, Piracicaba.

Moreira JUV, Bento DAV, Souza AP, Souza CL (2009) QTL mapping for reaction to Phaeosphaeria leaf spot in tropical maize. **Theoretical and Applied Genetics** 119:1361–1369. https://doi.org/10.1007/s00122-009-1140-0.

NASA (2021) Prediction of Worldwide Energy Resources (POWER) Project. NASA Technical Reports Server. https://ntrs.nasa.gov/citations/20210023937. Accessed 6 September 2024

Nunes LH, Gabriel GH, Marengo JA (2023) More erratic and more extreme: trends in precipitation in São Paulo State, Brazil. **American Journal of Climate Change** 12:140–171. https://doi.org/10.4236/ajcc.2023.121008.

Paccola-Meirelles LD, Ferreira AS, Meirelles WF, Marriel IE, Casela CR (2001) Detection of a bacterium associated with a leaf spot disease of maize in Brazil. **Journal of Phytopathology** 149(5): 275–279. https://doi.org/10.1046/j.1439-0434.2001.00614.x

Pegoraro DG, Vacaro E, Nuss CN, dal Soglio FK, Sereno MJC de M, Barbosa Neto JF (2001) Effect of planting date and fertilization on *Phaeosphaeria* leaf blight in maize. **Pesquisa Agropecuária Brasileira** 36: 1037-1042. https://doi.org/10.1590/S0100-204X2001000800005.

Pinto N. (1997) Eficiência de fungicidas no controle de doenças foliares do milho. **Summa Phytopathologica** 23: 271-274.

Rane MS, Payak MM, Renfro BL (1966) A Phaeosphaeria leaf spot of maize. **Indian Phytopathol. Soc. Bull** 3: 8–10.

Rolim G de S, Pedro Júnior MJ, Fantin GM, Brunini O, Duarte AP, Dudienas C (2007) Modelo agrometeorológico regional para estimativa da severidade da mancha de *Phaeosphaeria* em milho safrinha no Estado de São Paulo, Brasil. **Bragantia** 66: 721-728. https://doi.org/10.1590/S0006-87052007000400023

Sawazaki E, Dudienas C, Paterniani M, Galvao JCC, Castro JL, Pereira J (1997) Reação de cultivares de milho à Mancha de Phaeosphaeria no estado de São Paulo. **Pesquisa Agropecuaria Brasileira** 32(6): 553–558.

USDA (2024) Foreign Agriculture Service. https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?startrow=11&cropid=0440000&sel_year=2024&rankby=Production. Accessed 12 August 2024.

USDA (2025) World Agricultural Supply and Demand Estimates. United States Department of Agriculture, Washington, D.C.

Wilks D S (2011) Statistical methods in the atmospheric sciences. Third Edition. Academic press, 675p.

Xing E, Fan X, Jiang F, Zhang Y (2023) Advancements in Research on Prevention and Control Strategies for Maize White Spot Disease. **Genes** 14(11): 2061. https://doi.org/10.3390/genes14112061.

CHAPTER 3 - Early prediction of maize white spot severity through machine learning and agrometeorological time windows

ABSTRACT - Maize white spot (MWS) is one of the most important foliar diseases of maize in tropical and subtropical regions, with severity strongly modulated by agrometeorological conditions. This study aimed to develop a machine learning based approach to predict MWS severity by integrating agrometeorological and crop variables. Disease severity data collected between 1998 and 2022 in 16 municipalities of São Paulo State, Brazil, were combined with daily meteorological data from the NASA POWER dataset. Thermal, radiative and water availability variables, along with derived agrometeorological indices, were aggregated into temporal windows along the crop cycle, defined using changepoint detection applied to the dynamics of correlations between agrometeorological variables and disease severity. Crop variables, including hybrid type, growing degree days to flowering and cultivar resistance estimated via mixed models, were also incorporated. Four machine learning algorithms (Random Forest, Support Vector Regression, Multilayer Perceptron and XGBoost) were evaluated against a multiple linear regression baseline. Machine learning models consistently outperformed the linear approach, with XGBoost achieving the best performance (RMSE = 0.519; adjusted R^2 = 0.805 on the test set). Model interpretability based on SHAP values identified cultivar resistance as the most influential predictor, followed by variables related to water availability during advanced vegetative and early reproductive stages. The proposed approach enables early anticipation of MWS severity using agrometeorological information, representing a promising decision support tool for maize disease management.

Keywords: disease prediction, Phaeosphaeria leaf spot, artificial intelligence, decision support systems, Python.

3.1 Introduction

Maize (*Zea mays L.*) is one of the most important crops worldwide, playing a key role in global food security, forage production, and agro-industrial systems, particularly in tropical and subtropical regions (Tanumihardjo et al., 2020; FAO, 2022; Wang and Zhang, 2024). Despite advances in genetic improvement and agronomic management, maize productivity remains highly sensitive to biotic factors, among which foliar diseases are a major cause of crop damage and yield losses (Codognoto et al., 2017; Silva et al., 2021). In this context, *Phaeosphaeria* leaf spot, also known as maize white spot (MWS), has become one of the most relevant foliar diseases in maize production systems, being associated with substantial yield reductions that, under favorable environmental conditions and in susceptible cultivars, may exceed 50% (Rossi et al., 2017; Embrapa, 2021; Xing et al., 2023).

Because plant diseases result from the interaction between host, pathogen, and environmental conditions, particularly meteorological factors (Agrios, 2005), previous studies have shown that the development and severity of Maize White Spot (MWS) are favored by the combined effects of several agrometeorological variables, especially high atmospheric and soil moisture levels, prolonged periods of leaf wetness, and moderate temperature ranges (Sawazaki et al., 1997; Codognoto et al., 2017; Xing et al., 2023; Rodriguez-Roa et al., 2025). In tropical and subtropical regions, where these conditions exhibit pronounced intra and interannual variability throughout the crop cycle, disease severity consequently varies substantially among seasons and locations, hindering its characterization and reliable anticipation under field conditions.

The ability to anticipate the occurrence and intensity of plant diseases plays a central role in agronomic decision-making, as it directly influences key strategies such as cultivar selection, crop protection planning, and the optimization of fungicide use. For diseases such as MWS, which are closely associated with agrometeorological variables and characterized by high spatial and temporal variability, the development of predictive tools is essential to support timely crop management. Furthermore, the increasing demand for more efficient and sustainable production systems, aimed at reducing unnecessary chemical pesticide applications and strengthening early warning systems, has driven interest in approaches capable of supporting disease risk monitoring and epidemic prediction in agricultural systems (Morkeliūnė et al., 2021).

In recent years, advances in sensor technologies, the Internet of Things (IoT), and remote sensing, which have expanded both the availability and the spatial and temporal coverage of environmental information, together with the development of artificial intelligence (AI), have driven the increasing adoption of machine learning (ML) techniques for building predictive models of plant diseases (Fenu and Malloci, 2021; González-Domínguez et al., 2023; Jensen et al., 2025). In this regard, the systematic review conducted by Jensen et al. (2025), primarily focused on crops within the cereal, grapevine, potato, and citrus groups, highlights a progressive shift from traditional statistical approaches to machine learning based models over the past two decades. These approaches have enabled the development of models with high predictive performance, clearly contributing to improved estimation of disease risk and epidemic dynamics.

According to recent review studies (Fenu and Malloci, 2021; Jensen et al., 2025) machine learning based models applied to crop disease prediction have mainly focused on a limited set of plant pathologies, including potato late blight (*Phytophthora infestans*), rice blast (*Magnaporthe oryzae*), wheat stripe rust (*Puccinia striiformis*), and downy mildews affecting various crops. Both reviews report a clear predominance of classical ML algorithms, particularly Support Vector Machine and Support Vector Regression (SVM and SVR), Random Forest (RF), artificial neural networks (ANN), and Bayesian networks (BN). In terms of performance, SVM, SVR, and RF consistently exhibit high predictive accuracy, especially in models driven by meteorological variables, when compared with traditional statistical approaches and other ML methods. Regarding input variables, most models rely primarily on environmental factors, mainly temperature, humidity, and precipitation, whereas crop related variables and pathogen related variables remain scarcely incorporated.

Despite advances in the use of machine learning based approaches for predicting foliar diseases in a wide range of crops, their application to MWS remains limited. To the best of our knowledge, the main prior contribution is the study by Rolim et al. (2007), who developed agrometeorological models aimed at estimating disease severity at a regional scale based on accumulated temperature and precipitation over the crop cycle. Although these models represented an important contribution to understanding the relationship between climate and disease in maize “safrinha” systems, they rely on traditional statistical approaches, do not incorporate ML techniques, and do not explicitly account for the differential influence of meteorological conditions across distinct crop development stages on MWS severity. In this context, there remains a need to develop more flexible predictive approaches that integrate agrometeorological information at multiple temporal scales throughout the crop cycle, to improve risk anticipation and support decision-making in MWS management.

Based on the above, the objective of this study was to develop and evaluate a machine learning based predictive model for MWS severity by integrating agrometeorological information through temporal windows along the crop cycle. The proposed approach combines agrometeorological and crop related variables to identify the maize development stages with the greatest influence on disease severity and to enhance the ability to anticipate disease risk. The results of this study aim to contribute to the development of decision support tools for more timely and efficient management of MWS in maize production systems.

3.2 Materials and methods

3.2.1 Study area

The study was carried out in the state of São Paulo, located in southeastern Brazil, between 20°- 25°S and 44°-53°W (Fig.8) totaling 25 million ha. Although São Paulo presents a broad elevational gradient, most of its territory exhibits moderate relief, with nearly 85% of the area positioned between 300 and 900 m. The landscape is largely characterized by a relatively homogeneous plateau that extends eastward (Nunes et al. 2023). According to the Köppen climate classification, São Paulo comprises tropical savanna (Aw) climates in its northern and western regions, as well as humid subtropical (Cfa) climates in central, southern, and high-altitude zones (Alvares et al., 2013; Rolim, Camargo, et al., 2007).

3.2.2 Field and weather data

3.2.2.1. Field disease assessments

The MWS severity data used in this study were obtained from field evaluations conducted as part of the Maize Project, coordinated by IAC/APTA (Institute of Agronomy / São Paulo Agency for Agricultural Technology) with support from the Biological Institute (IB). One of the objectives of this project is to monitor the main diseases affecting second-season maize (safrinha) in the major producing regions of the state of São Paulo, through experiments established each year with maize cultivars considered to be of greatest relevance for the period evaluated.

Between 1998 and 2022, 15 to 40 field experiments were conducted annually, one or two per location, evaluating 20 to 64 maize cultivars in each trial. All experiments followed a randomized complete block design, with plots consisting of four five-meter rows and three or four replications. Planting was carried out between February and March, and harvest occurred between July and August.

MWS severity was visually assessed using the Agroceres diagrammatic scale (Agroceres 1993), which employs scores from 1 to 9 corresponding to 0, 1, 2.5, 5, 10, 25, 50, 75, and more than 75% of leaf area affected. For each hybrid, severity was calculated as the mean of all replications at each location, resulting in a single severity score per trial.

For this study, severity data were obtained for maize single-cross (SC), double-cross (DC), and three-way cross (TC) hybrids, collected over a 15-year period across 16 municipalities in the Paranapanema Valley and northern regions of the state of São Paulo (Fig. 1). Evaluations conducted approximately 115 days after planting were considered, which on average correspond to the R4 phenological stage (dough stage), with assessment dates concentrated between June and July each year. The analyses included plots without pesticide treatment and plots with a single application, generally of a preventive nature.

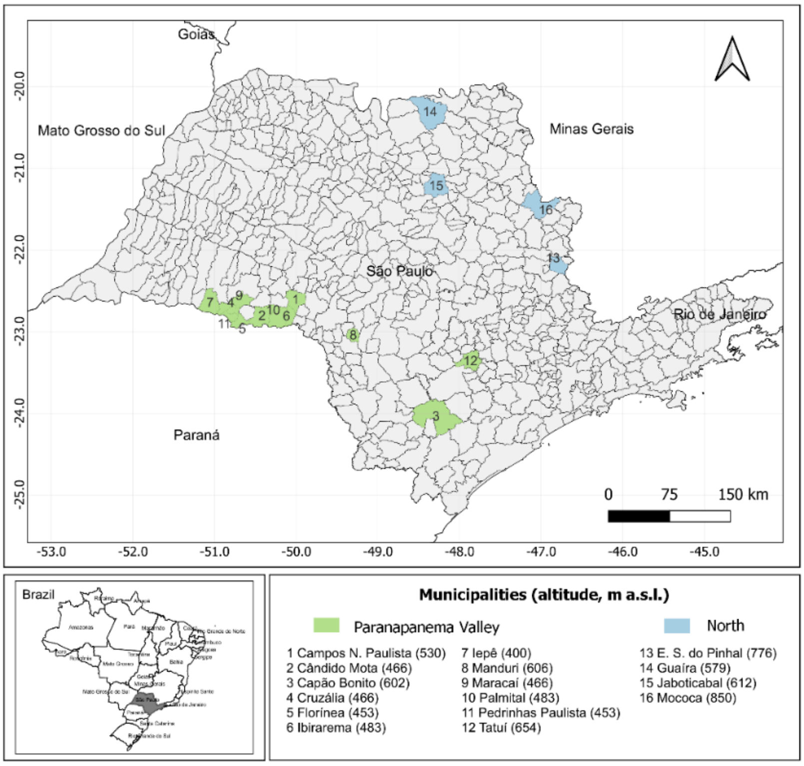


Fig.8. Study area showing the 16 municipalities where field data were collected in the Paranapanema Valley and northern regions of the state of São Paulo, Brazil.

3.2.2.2. Weather dataset and meteorological variables

Meteorological data were obtained from the NASA POWER dataset (https://power.larc.nasa.gov) at a daily temporal resolution for the period 1998-2022, using the geographic coordinates of each municipality included in the study. This dataset, which integrates satellite observations and reanalysis models, has a spatial resolution of 0.5° latitude/longitude (approximately 55 km), providing a consistent representation of regional atmospheric conditions. The variables considered were: mean (Tmean), maximum (Tmax), and minimum (Tmin) air temperature (°C); solar radiation at the top of the atmosphere (Qo, MJ m^{-2} d^{-1}); incoming solar radiation at the Earth's surface (GSR, MJ m^{-2} d^{-1}); relative humidity (RH, %); precipitation (PPT, mm); and wind speed (WS, m s^{-1}).

3.2.3 Agrometeorological variables and derived indices

Based on the meteorological variables obtained, several indices and derived agrometeorological variables were calculated. This included reference evapotranspiration (ETo, mm), estimated using the FAO Penman-Monteith method

(Allen et al., 1998); vapor pressure deficit (VPD, kPa); diurnal temperature range (DTR, °C); and an atmospheric water availability index (WI, dimensionless), calculated as the ratio between precipitation (PPT) and ETo.

Soil water storage (SWS, mm) was also estimated through a sequential water balance following the methodology of Thornthwaite and Mather (1955). This procedure used the daily values of PPT and ETo together with the available water capacity (AWC) specific to each municipality within the study area. The AWC information was obtained from the Irrigation Atlas (ANA, 2021) via the SNIRH system (https://metadados.snirh.gov.br). In addition, leaf wetness duration (LWD, h) was calculated as the number of hours with relative humidity greater than or equal to 90 %, following the approach described by de Lima et al. (2022).

All variables were computed at a daily time scale for the 1998–2022 period. However, for the modeling step, only the values corresponding to the 115 days preceding the disease assessment date were considered for each year and locality included in the study, corresponding to the average number of days between sowing and disease assessment. This interval represents the average number of days between planting and the end of the dough stage (R4). Furthermore, growing degree days from planting to flowering (GDDF), approximately 70 days after planting, were estimated using a base temperature of 10 °C and an upper threshold temperature of 30 °C for maize, as proposed by (Nleya et al., 2016).

3.2.4 Hybrid resistance classification

Hybrid resistance to MWS was classified using a mixed-model approach to adjust disease severity for environmental effects and obtain a robust measure of genetic resistance. Because not all hybrids were evaluated across all years and municipalities, and considering that environmental variability strongly influences disease severity, we adopted a methodology based on linear mixed models following Piepho et al., 2003).

First, a linear mixed model was fitted with environment (year and municipality combined) as a random effect to correct for environmental variation in disease severity. The residuals from this model were then used to fit a second mixed model including hybrid as a random effect. The Best Linear Unbiased Predictors (BLUPs) obtained for each hybrid were interpreted as environment-adjusted severity estimates and thus as indicators of genetic resistance.

Hybrids were subsequently classified into three resistance categories: Resistant (R), Moderately Resistant (MR), and Susceptible (S), according to the tertiles of their BLUP values, herein denoted as cultivar resistance (CR). This approach allowed us to identify genetic differences in resistance in a robust manner while minimizing environmental influence.

3.2.5 Correlation-based definition of agroclimatic windows

A daily correlation analysis was conducted between MWS severity and each agrometeorological variable (Tmean, Tmax, Tmin, GSR, PPT, RH, WS, ETo, VPD, DTR, WI, SWS, and LWD) using Spearman's correlation coefficient. Subsequently, the absolute values of the daily correlations were summed to integrate the multivariate agroclimatic influence into a single univariate time series, enabling the identification of seasonal patterns and the delineation of some relevant time windows. Spearman's correlation was selected because it does not require the assumption of normality in the data distribution.

To objectively define these windows and avoid arbitrary segmentation, we applied the PELT (Pruned Exact Linear Time) changepoint detection algorithm proposed by Killick et al. (2012) to the univariate agroclimatic time series. This algorithm partitions the series into homogeneous intervals by optimizing the number and position of breakpoints through a linear penalty function. Each resulting window was used to calculate aggregated agrometeorological variables and was associated with a specific phenological stage of maize growth. This allowed the assessment of the relationship between disease severity and the crop cycle. Additionally, these aggregated variables were subsequently used as predictor features in the modeling stage.

3.2.6 Machine learning modeling approach for MWS severity prediction

The overall workflow consisted of three stages preprocessing, processing and outcomes (Fig. 2). The preprocessing stage included the previously described steps:

calculation of agrometeorological indices and derived variables, aggregation of these variables into temporal windows, and classification of hybrid resistance. This step consolidated all the input information required to construct the dataset used for modeling.

For the processing (modeling) stage, the final dataset included 13 agrometeorological variables aggregated within the climatic windows, along with three crop-related variables (hybrid type, CR and GDDF), used as input features. MWS severity was employed as the target variable. In total, 2,446 observations were available. The dataset was divided into training (70%) and testing (30%) subsets. Four machine learning algorithms were evaluated: Random Forest Regression (RF), Support Vector Regression (SVR), Multilayer Perceptron (MLP), and EXtreme Gradient Boosting (XGBoost), along with a multiple linear regression (MLR) model. All models were implemented in Python using the Scikit-Learn library.

Each algorithm was calibrated by optimizing its hyperparameters on the training subset using a grid search procedure (GridSearchCV), which evaluated different parameter combinations through cross-validation using the coefficient of determination as the primary metric. Once the best configuration was identified, the optimal algorithm was refitted using the full training data and subsequently used to generate predictions for both the training and testing subsets. Finally, model performance was evaluated using the root mean squared error (RMSE) as a metric of accuracy, adjusted coefficient of determination (adjusted R^2) as precision, and systematic error (ES) as bias.

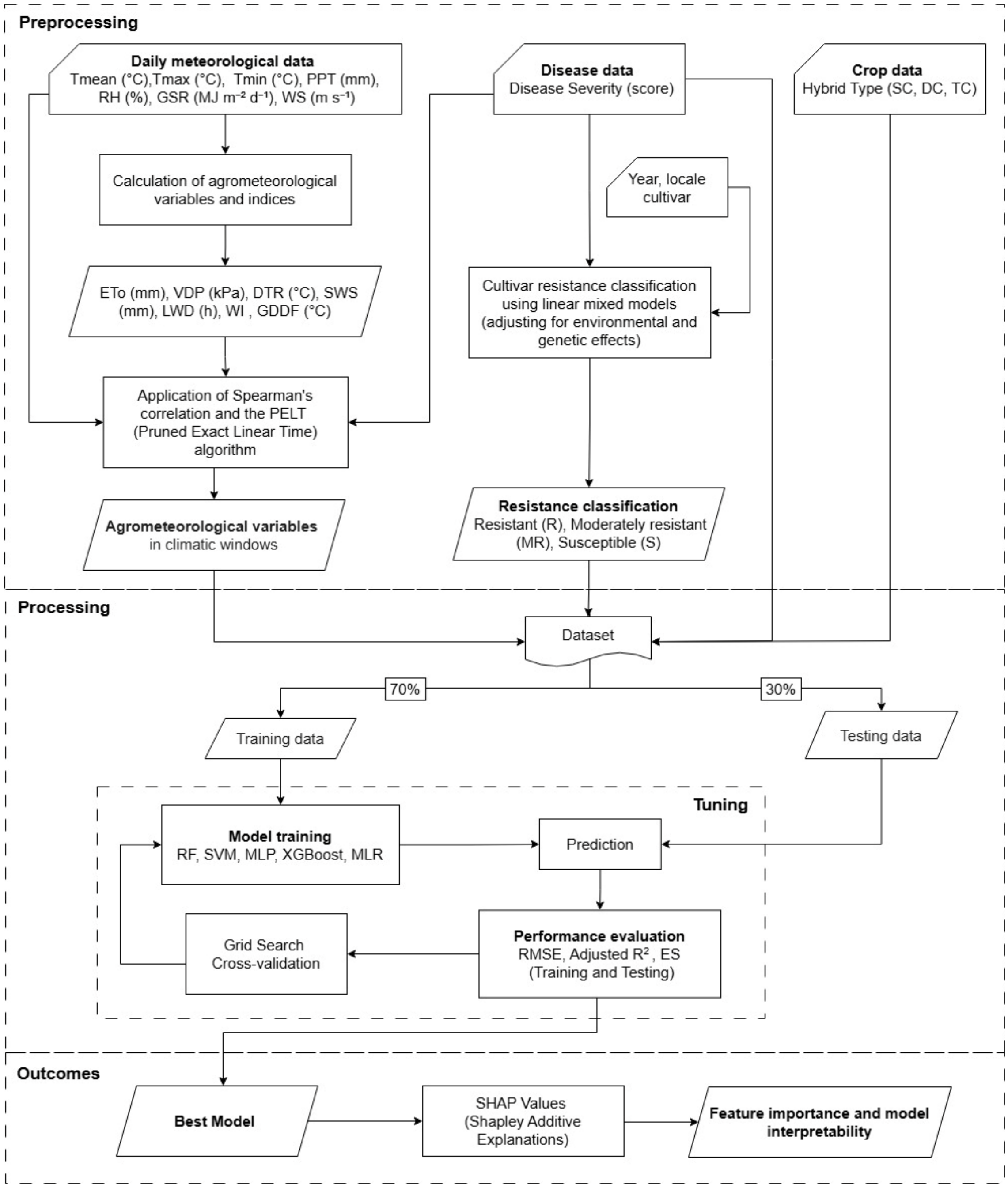


Fig. 9. Flow chart of the research process. Symbols: Tmean, mean temperature; Tmax, maximum temperature; Tmin, minimum temperature; PPT, precipitation; RH, relative humidity ; GSR, solar radiation at the Earth's surface; WS, wind speed; ETo, reference evapotranspiration; VDP, vapor pressure deficit; DTR, diurnal temperature range; SWS, soil water storage; LWD, leaf wetness duration; WI, atmospheric water availability index; GDDF, growing degree days from planting to flowering; SC, maize single-cross hybrids; DC, maize double-cross hybrids; TC, maize three-way hybrids; RF, Random Forest Regression; SVM, Support Vector Regression; MLP, Multilayer Perceptron; XGBoost, eXtreme Gradient Boosting; MLR, multiple linear regression; RMSE, root mean squared error; AdjustedR2, adjusted coefficient of determination; ES, systematic error.

3.2.7 Model interpretability

After identifying the best-performing algorithm based on the evaluation metrics from the testing subset, this model was analyzed (Fig. 2) to determine how each predictor contributed to the final MWS severity estimates. Model interpretability was assessed using SHAP (Shapley Additive Explanations) values, which quantify both the magnitude and direction of each predictor's contribution to individual predictions. Global SHAP importance plots were used to identify the variables with the greatest overall influence and to examine how different values of each predictor affected the model output. This provided an interpretable representation of the relationships captured by the model and clarified the role of each predictor in shaping the final severity estimates.

3.3 Results

3.3.1 Maize white spot severity and agrometeorological conditions

Based on the Agroceres diagrammatic scale, MWS severity exhibited a right-skewed distribution, with most observations below 10%. The mean value was close to 1%, and most observations fell within the 0-2.5% range. Despite this concentration of low values, several high-severity outliers were observed, with maximum levels approaching 75% (Fig. S3).

The agrometeorological conditions during the 115 days preceding the disease assessment were characterized by moderate temperatures and predominantly humid conditions. Mean air temperature was 20.3 °C (range: 16.9-23.9 °C), with mean maximum and minimum temperatures of 26.2 °C (21.7-29.7 °C) and 15.5 °C (13.2-19.0 °C), respectively. Mean global solar radiation was 15.3 MJ m^{-2} day^{-1} (11.2-18.7 MJ m^{-2} day^{-1}), while cumulative reference evapotranspiration totaled 313 mm (244-393 mm). Relative humidity ranged from 65.4% to 87.8% (mean = 78.9%) and was associated with a low vapor pressure deficit (mean = 0.52 kPa; range = 0.24-0.83 kPa). Total precipitation averaged 327 mm (93-564 mm), indicating substantial interannual and spatial variability. Leaf wetness duration averaged 6.9 h day^{-1} (3.5-9.0 h day^{-1}), and mean wind speed was 1.55 m s^{-1} (0.36-2.29 m s^{-1}) (Table S2).

3.3.2 Daily correlations between agrometeorological variables and MWS severity

The daily correlations between MWS severity and the agrometeorological variables exhibited clear temporal patterns over the 115 days preceding the disease assessment, a period corresponding to the average interval from sowing to the dough grain stage.

Regarding the thermal variables (Tmean, Tmax, Tmin, DTR, and GSR), the daily correlations with MWS severity were predominantly negative across most of the 115-day period (Fig. S2). These negative correlations were more consistent during the early part of the crop cycle, whereas closer to the disease assessment they became more variable in both magnitude and sign.

The water-related variables (SWS, PPT, RH, LWD, ETo, WI, VPD, and WS) exhibited more heterogeneous correlation patterns with MWS severity across the 115-day period (Fig. S3). Soil water storage (SWS), relative humidity (RH), and leaf wetness duration (LWD) showed predominantly positive correlations, particularly from the middle of the period onward. In contrast, vapor pressure deficit (VPD) and wind speed (WS) displayed predominantly negative correlations throughout most of the period.

The cumulative influence of all agrometeorological variables, summarized as the daily sum of the absolute Spearman correlation coefficients (Fig. S4), increased progressively over the 115-day period, with a marked rise from approximately day 40 onward. Higher cumulative values were consistently observed toward the end of the period, reflecting an overall increase in the combined magnitude of the correlations as the assessment date approached.

3.3.3 Identification of agroclimatic time windows

Five agroclimatic time windows (W1 to W5) were identified based on the changepoints detected in the cumulative daily correlations (Fig.10.). These intervals could be associated with some of the major phenological phases of second-season maize in the state of São Paulo. W1 (days 1-40) covered the period from emergence to mid-vegetative development (VE-V6/V8), while W2 (days 41-70) encompassed advanced vegetative growth up to flowering (V8-VT). W3 (days 71-85) spanned the interval from flowering to the onset of the kernel blister stage (R2). W4 (days 86-105) extended from the kernel blister stage to the onset of kernel dough development (R2-R4), and W5 (days 106-115) covered the kernel dough phase (R4) up to the time of disease assessment.

The correlations obtained revealed distinct patterns across the different time windows. Based on the mean absolute correlation values, the period between emergence and mid-vegetative growth (W1) exhibited the lowest overall association between agrometeorological variables and MWS severity, with a mean cumulative correlation of approximately 2.1. In contrast, higher mean values were observed in subsequent windows, particularly during the advanced vegetative growth to male flowering stage (W2, ≈3.4) and the kernel dough stage (W5, ≈4.2), indicating a stronger overall correlation structure during these periods.

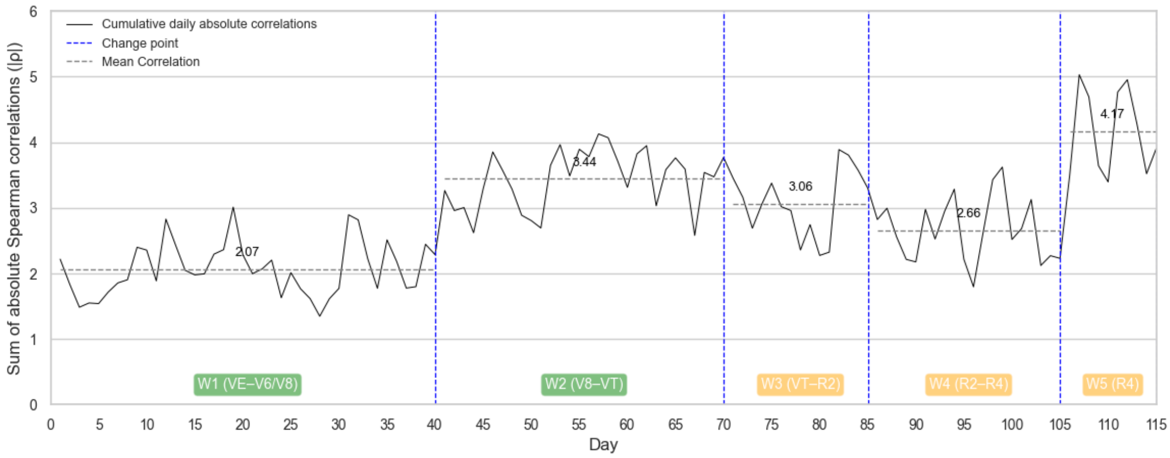


Fig.10. Agroclimatic time windows (W1–W5) identified from changepoints in the cumulative daily sum of absolute Spearman correlation coefficients between MWS severity and agrometeorological variables. Symbols: ─ daily absolute correlations sum, --- detected changepoints, --- mean correlation value within each window. Phenological stages: VE, emergence; VT, tasseling (male flowering); R2, kernel blister stage; R4, kernel dough stage, W1, emergence–mid vegetative growth; W2, advanced vegetative growth–male flowering; W3, flowering–kernel blister onset; W4, kernel blister–kernel dough onset; W5, kernel dough stage.

Building on the defined agroclimatic time windows, correlations between MWS severity and the agrometeorological variables aggregated within each interval were examined (Fig. 11. and Fig. S7). Clear contrasts in both the sign and magnitude of the correlations were observed among windows. From emergence through mid-vegetative growth (W1), most variables exhibited low coefficients, close to zero or non-significant, indicating weak association during this initial period. In contrast, during advanced vegetative growth up to male flowering (W2), several variables showed higher correlation magnitudes, with positive coefficients for SWS (ρ ≈ 0.48), PPT (ρ ≈ 0.48), RH (ρ ≈ 0.40), and WI (ρ ≈ 0.51), while Tmax and VPD displayed negative correlations of comparable magnitude.

In subsequent windows, similar patterns persisted, although with variations in intensity. From flowering to the onset of the kernel blister stage (W3) and from the kernel blister stage to the onset of kernel dough (W4), positive correlations associated with water-related variables remained, although with generally moderate values. During the kernel dough phase (W5), the highest correlation magnitudes were again observed, including strong positive coefficients for SWS (ρ ≈ 0.46), RH (ρ ≈ 0.49), and leaf wetness duration (LWD; ρ ≈ 0.54), along with pronounced negative correlations for VPD (ρ ≈ -0.54) and temperature-related variables. Overall, these results highlight marked temporal variability in the relationships between agrometeorological conditions and MWS severity throughout the crop cycle.

Fig. 11. Spearman correlation coefficients between agrometeorological variables and MWS severity aggregated within the defined agroclimatic time windows (W1–W5). Colors indicate the magnitude and direction of the correlations. Symbols: NS correlations not significant at the 0.01 probability level; Tmean, mean temperature; Tmax, maximum temperature; Tmin, minimum temperature; PPT, precipitation; RH, relative humidity ; GSR, solar radiation at the Earth's surface; WS, wind speed; ETo, reference evapotranspiration; VDP, vapor pressure deficit; DTR, diurnal temperature

range; SWS, soil water storage; LWD, leaf wetness duration; WI, atmospheric water availability index.

3.3.4 Model performance for MWS severity prediction

Based on the temporal window analysis, the models were trained using agrometeorological variables aggregated over windows W1, W2, and W3. Together, these windows encompass approximately the first 85 days of the crop cycle, leaving an average interval of about 30 days before the disease severity assessment, which occurs around the kernel dough stage (R4).

3.3.4.1 Training and testing performance of the evaluated models

The machine learning-based models (Table S3) outperformed multiple linear regression, exhibiting lower RMSE values and higher adjusted R^2 in both datasets (Table 3). Among them, XGBoost achieved the best performance in both training and testing, followed closely by Random Forest, SVR, and MLP, which showed similar performance levels. In contrast, the multiple linear regression model yielded the highest errors and the lowest adjusted R^2 values. The systematic error (ES) was comparable across the evaluated models and showed limited variation between the training and testing datasets.

Table 3. Performance metrics of the five evaluated algorithms in the training and testing datasets for MWS severity prediction. RMSE and ES are expressed in disease severity units (score), whereas adjusted R2 is dimensionless.

	Training			Testing		
Model	**RMSE**	**Adjusted R^2**	**ES**	**RMSE**	**Adjusted R^2**	**ES**
RF	0.423	0.886	1.265	0.526	0.800	1.210
SVR	0.444	0.874	1.265	0.528	0.798	1.211
MLP	0.453	0.868	1.265	0.529	0.797	1.210
XGBoost	0.420	0.887	1.265	0.519	0.805	1.210
MLR	0.611	0.761	1.265	0.642	0.702	1.210

3.3.4.2 Observed vs. predicted severity

A clear correspondence between predicted and observed MWS severity was observed in the training dataset, with most machine learning models showing a close alignment with the 1:1 reference line (Fig. 12). Predictions generated by the nonlinear algorithms, particularly XGBoost and Random Forest, were characterized by a tighter clustering of points around the diagonal. In contrast, the multiple linear regression model exhibited a wider dispersion of predicted values across the observed range.

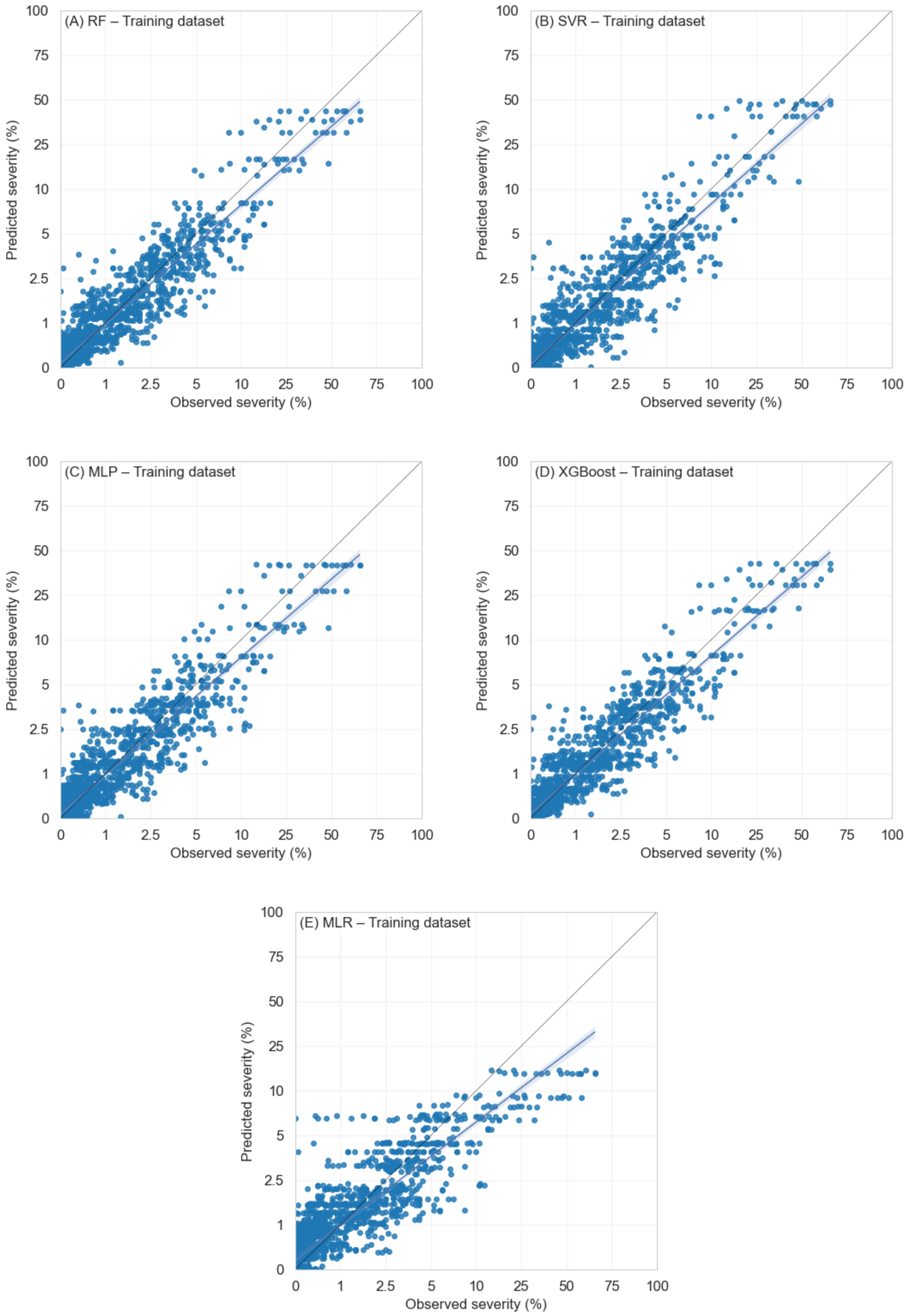
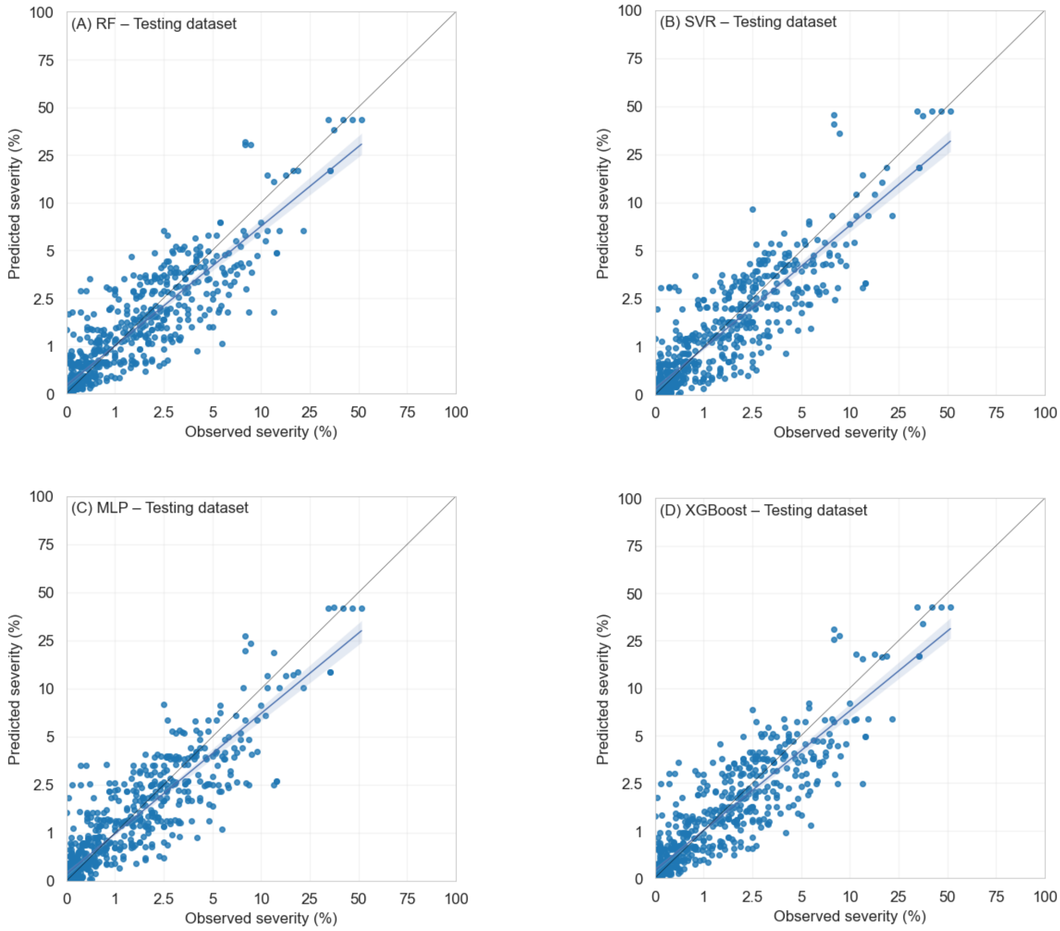


Fig. 12. Scatterplots of predicted versus observed MWS severity for the training dataset using Random Forest, SVR, MLP, XGBoost, and multiple linear regression (MLR). The solid 1:1 line indicates perfect agreement between observed and predicted values. Axes are expressed as percentages of affected leaf area.

A similar pattern was observed for the testing dataset, although with a slightly greater dispersion of points compared with the training results (Fig. 13). The nonlinear models generally maintained a close correspondence with the 1:1 reference line, with XGBoost showing the tightest clustering around the diagonal among the evaluated approaches. Random Forest and SVR also exhibited good agreement between predicted and observed values, albeit with increased variability. In contrast, the multiple linear regression model displayed the largest deviations from the diagonal across the observed range.



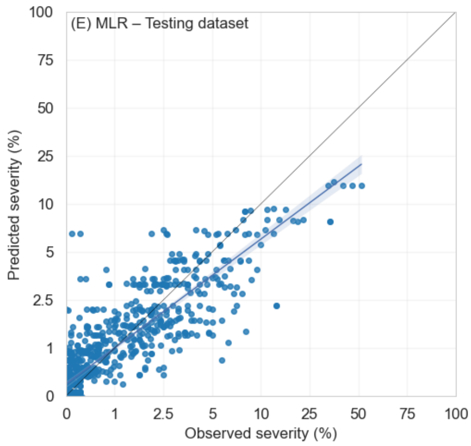


Fig. 13. Scatterplots of predicted versus observed MWS severity for the testing dataset using Random Forest, SVR, MLP, XGBoost, and multiple linear regression (MLR). The solid 1:1 line indicates perfect agreement between observed and predicted values. Axes are expressed as percentages of affected leaf area.

3.3.5 Model interpretability: SHAP analysis

Since the evaluated models produced very similar results across the performance metrics, the selection of the algorithm for the interpretability analysis was based on the model showing the strongest overall performance in both modeling phases. Under this criterion, XGBoost was identified as the most robust model, presenting the lowest error values and the highest adjusted R^2 in both the training and testing subsets. Therefore, XGBoost was selected as the final model for the interpretability analysis and for exploring the factors driving maize white spot severity.

The SHAP analysis (Fig. 14) identified cultivar resistance as the most influential predictor in the XGBoost model, with higher resistance levels associated with lower predicted disease severity. Among the agrometeorological variables, water-related factors showed prominent contributions, particularly soil water storage (SWS) and the water index (WI) during the V8-VT (W2) and VT-R2 (W3) periods. Wind speed (WS) and precipitation (PPT) during W2 and W3 also ranked among the most influential predictors, although with comparatively smaller contributions.

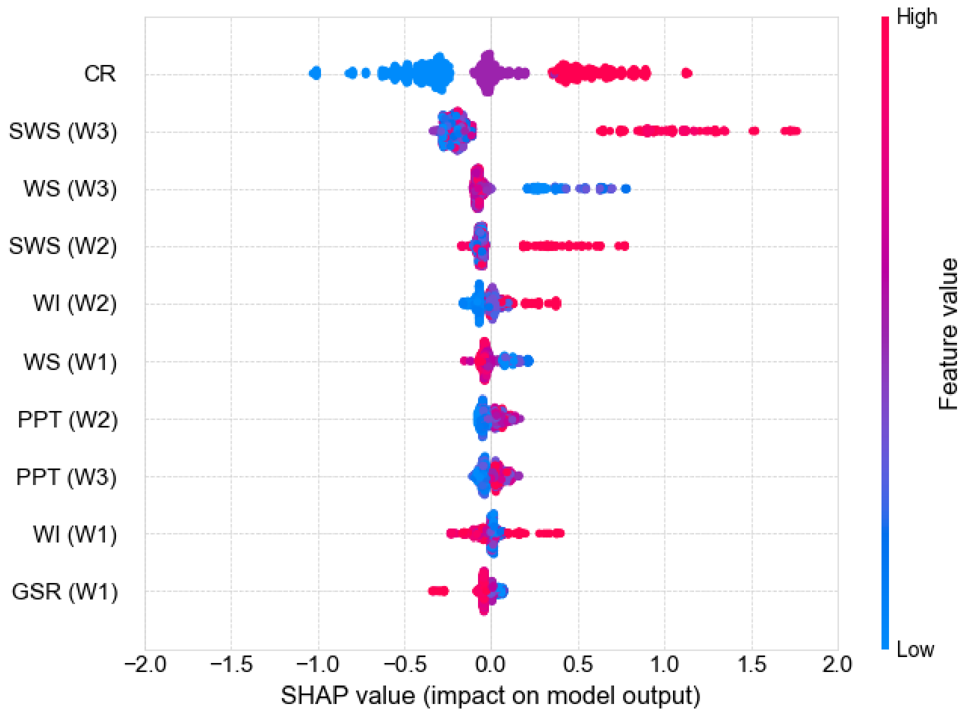


Fig. 14. SHAP summary plot showing the top ten predictors contributing to MWS severity prediction for the final XGBoost model. Each point represents an individual observation, with SHAP values indicating the contribution of each predictor to the model output. Point colors represent low (blue) and high (red) values of each predictor. Symbols: SWS, soil water storage; WS, wind speed; PPT, precipitation; WI, atmospheric water availability index; GSR, solar radiation at the Earth's surface; W1, window VE-V6/V8; W2, window V8-VT; W3, window VT-R2.

To provide a more detailed assessment of how the most influential predictors identified in the SHAP summary plot affect model predictions, SHAP dependence plots were analyzed (Fig. 15). These plots illustrate the relationship between the value of each predictor and its contribution to the predicted MWS severity, allowing an evaluation of how the model responds to changes in each predictor.

Cultivar resistance exhibited a clear and consistent effect on the predicted disease severity (Fig. 15A). Resistant cultivars were associated with negative SHAP values, indicating a reduction in predicted severity, whereas susceptible cultivars showed positive SHAP values, reflecting a greater contribution to increased disease severity. Moderately resistant cultivars displayed an intermediate response.

Soil water storage showed a phenology-dependent effect. During the VT-R2 period (W3), its influence was limited at low to moderate values but increased markedly at higher values (Fig. 15B). In this period, low to moderate SWS values contributed little or negatively to predicted severity, whereas high values resulted in a pronounced increase in SHAP values, indicating a strong positive contribution when soil moisture remained high during this critical phenological stage. Similarly, during the V8-VT period (W2), soil water storage exhibited a weaker but consistent positive contribution at higher values (Fig. 15D).

Wind speed during the VT-R2 period (W3), although showing a smaller contribution, indicated that low wind speeds were associated with positive SHAP values, reflecting an increase in predicted severity, whereas at moderate to high wind speeds its impact was limited or slightly negative (Fig. 15C). This pattern may reflect the role of wind in dispersing pathogen propagules at low to moderate speeds, while stronger winds may reduce leaf wetness and humidity, limiting pathogen survival.

Overall, the SHAP dependence plots indicate that cultivar resistance and soil moisture conditions during key phenological stages are the main factors shaping the model predictions of MWS.

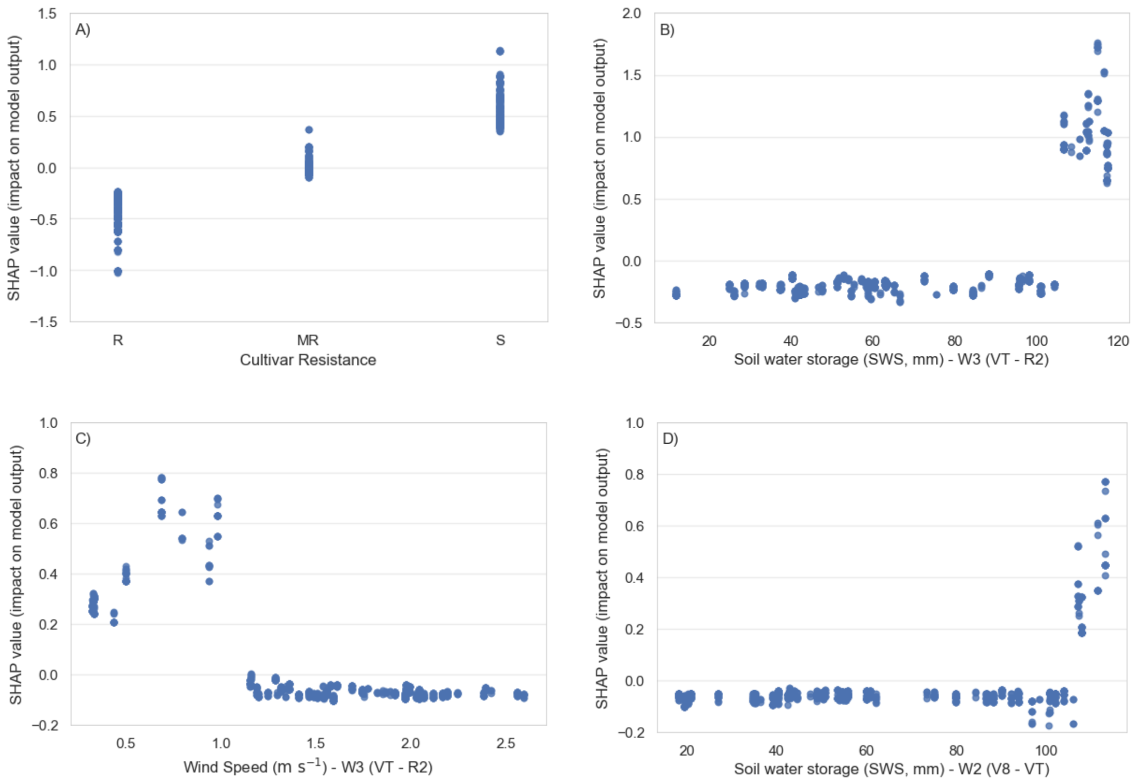


Fig. 15. SHAP dependence plots illustrating how the most influential predictors affect the predicted severity of MWS according to the XGBoost model. Each point represents an observation from the testing dataset. The x-axis shows the value of each predictor,

while the y-axis indicates its contribution (SHAP value) to the prediction. Positive SHAP values indicate an increase in predicted severity, whereas negative values indicate a decrease. Symbols: R, resistant; MR, moderately resistant; S, susceptible**.**

3.4 Discussion

3.4.1 Predictive performance and modeling strategy

The results indicate that machine learning-based models consistently outperformed multiple linear regression in both the training and testing datasets. This difference suggests that the relationship between MWS severity and the agrometeorological variables considered is predominantly nonlinear and mediated by complex interactions that cannot be adequately captured by traditional linear approaches. In this context, the good performance of algorithms such as XGBoost, Random Forest, SVR, and MLP highlights the ability of these models to integrate multiple sources of climatic and genetic information throughout the crop cycle.

From a methodological perspective, in most crop and plant disease prediction studies reviewed by Fenu and Malloci (2021), support vector machines have often been reported to outperform other machine learning approaches, such as artificial neural networks or random forest models. In contrast, the results of the present study showed very similar predictive performance across the evaluated machine learning algorithms. Within this group, tree-based ensemble methods exhibited a slight advantage, particularly XGBoost, a gradient boosting approach that has been increasingly adopted in agricultural applications in recent years. Previous studies have demonstrated the suitability of XGBoost for crop yield prediction (Mallikarjuna et al., 2022; Mariadass et al., 2022) and, more recently, for plant disease prediction, for example in the case of *Phoma* leaf spot in coffee (*Phoma spp.*) (de Oliveira et al., 2024), highlighting its capacity to represent interactions among crop, pathogen, and environmental conditions.

The use of agroclimatic time windows, defined based on the dynamics of daily correlations, enabled the identification of crop development periods during which environmental conditions exert a greater influence on final disease severity. In contrast to approaches based on fixed or cumulative averages over the entire crop cycle, this strategy facilitated the identification of relevant climatic signals associated with specific phenological stages, providing a more structured framework for analyzing the relationship between environmental conditions and MWS development.

Furthermore, training the models using agrometeorological information specifically from the early windows of the crop cycle (VE-R2) enabled the generation of predictions with sufficient lead time prior to disease assessment. This approach underscores the potential of the model to anticipate scenarios of increased disease severity based on environmental conditions occurring during early stages of crop development.

3.4.2 Role of host resistance in MWS severity prediction

The interpretability analysis showed that cultivar resistance (CR) was the most influential predictor of disease severity in the final model. This result underscores the central role of the host in determining MWS severity and is consistent with previous studies emphasizing the importance of genetic resistance as a key factor in foliar disease expression (de Souza et al., 2019; Kistner et al., 2021; Xing et al., 2023).

The use of resistance values derived from mixed models enabled the estimation of cultivar resistance adjusted for environmental variability associated with the different years and locations. In this context, the incorporation of CR as a predictor in machine learning models provided an integrated representation of the host effect, which contributed to distinguishing scenarios of higher and lower disease severity according to the susceptibility of the evaluated genetic material.

3.4.3 Influence of agroclimatic conditions across phenological windows

The influence of agrometeorological conditions on MWS severity exhibited a clear temporal dependence, with a progressive increase in the magnitude of associations from the intermediate windows of the crop cycle onward. In particular, the strongest correlations observed from W2 coincide with phenological stages starting at V8-V9 and extending into advanced reproductive phases, which is consistent with previous reports indicating that MWS typically appears from V9 and that its critical phase occurs between VT and R5, often persisting until the end of the crop cycle (Aegro, 2023; Getap, 2022).

The increased importance of agroclimatic variables from the V8-V9 stages onward may indicate that pathogen establishment and early infection processes begin during these phenological phases, even in the absence of visible symptoms. Under favorable environmental conditions, infection may already be occurring during advanced vegetative stages, with the pathogen colonizing host tissues while remaining asymptomatic. As a result, environmental conditions during these stages may influence disease development before symptom expression becomes evident later in the crop cycle.

This interpretation is consistent with physiological evidence reported by Godoy et al. (2001), who showed that reductions in leaf photosynthesis caused by *Phaeosphaeria maydis* were substantially greater than the visually assessed disease severity. In their study, leaves exhibiting only 10-20% disease severity showed reductions of approximately 40% in net photosynthetic rate. These findings indicate that pathogen infection can impair host physiological processes well before extensive symptom development, reinforcing the hypothesis that disease impact begins prior to visible lesion expansion.

Among the agrometeorological variables, those associated with water availability showed a particularly strong influence on MWS severity, especially during periods V8-VT (W2) and VT-R2 (W3). Both soil water storage (SWS) and the water index (WI) exhibited recurrent positive associations, suggesting that sustained moisture conditions during advanced vegetative and early reproductive stages favor disease progression. This pattern has been previously reported in studies highlighting the importance of soil moisture and the soil-plant-atmosphere system as key factors in MWS severity, as they promote environmental conditions conducive to pathogen development (A. O. Rodriguez-Roa et al., 2025).

Regarding precipitation, in contrast to previous findings reported by Rodriguez-Roa et al. (2025), where accumulated precipitation did not show a clear association with disease severity when evaluated in relation to the period close to the assessment date, accumulated precipitation in the present study emerged as a relevant variable during periods V8-VT (W2) and VT-R2 (W3). This difference suggests that the effect of precipitation on MWS is not determined solely by total rainfall amounts, but also by its occurrence during specific phenological stages of the crop, when it can contribute to sustaining favorable moisture conditions for disease development. This interpretation is consistent with the findings of Sawazaki et al. (1997), who emphasized the importance of rainfall uniformity as a relevant factor for MWS.

In contrast to humidity-related variables, wind speed showed predominantly negative associations with MWS severity across the analyzed windows. This behavior suggests that higher wind speeds may limit disease development by promoting canopy drying and reducing the duration of wetness conditions on the leaf surface.

In the modeling framework, air temperature did not emerge as an influential predictor of maize white spot severity. However, air temperature is widely recognized as an important factor conditioning disease occurrence and development (Codognoto et al., 2017; Moreira, 2004; Rodriguez-Roa et al., 2025). This result may be related to the characteristics of the dataset, which was derived exclusively from the second season maize crop (safrinha), when air temperature typically remains within moderate and relatively stable ranges. Under such conditions, temperature may act as a prerequisite for disease occurrence but contribute little to explaining differences in disease severity once favorable thermal conditions are met. In contrast, variables such as soil water storage exhibited greater variability and therefore greater predictive importance. Temperature effects might become more evident if disease evaluations from contrasting cropping seasons, such as “safra” and “safrinha”, were jointly considered.

4.4.4 Implications for disease forecasting and crop management

The final model demonstrates potential as a decision-support tool for MWS management. Integrating agrometeorological, phenological, and crop resistance information enables the identification of key factors influencing disease severity and provides an early estimate of MWS risk. Beyond predicting a severity value, the incorporation of agroclimatic windows offers an interpretive framework that facilitates translating agrometeorological information into management-relevant indicators for crop management.

A particularly relevant aspect of the proposed approach is its ability to generate predictions based on agrometeorological data from the early windows of the crop cycle, up to W3, allowing MWS risk to be anticipated before advanced reproductive stages. This time frame is consistent with standard management practices in the region, where the final fungicide application is typically performed around the R2 stage (De Cól et al., 2025), which coincides with the end of window W3 and the beginning of W4. In this way, the model provides information sufficiently in advance to support decision-making, aligns with crop operational timelines, and avoids reliance on data collected close to the moment of disease assessment.

Overall, the results suggest that the proposed approach can contribute to the development of decision-support systems by accounting for the temporal dynamics of MWS. Future studies could focus on the validation and extrapolation of the model in other producing regions, as well as on incorporating management variables and additional sources of information, such as remote sensing data, which may further enhance model performance through the combined use of meteorological and remotely sensed information (Fenu and Malloci, 2021).

3.5 Conclusions

The results indicate that MWS severity can be predicted with a satisfactory level of accuracy using machine learning models that integrate agrometeorological, phenological, and genetic resistance information. Compared with multiple linear regression, these models showed better predictive performance in both the training and testing datasets, with very similar results across the different algorithms evaluated. Within this group, XGBoost showed a slight advantage and emerged as the most prominent alternative for predicting disease severity.

The use of agroclimatic windows defined based on the temporal dynamics of correlations enabled the objective identification of crop cycle periods during which environmental conditions exert the greatest influence on MWS severity. This approach allowed climatic signals to be linked to specific phenological stages, providing an objective methodological basis for analyzing the relationship between climate, crop development, and disease severity.

The model interpretability analysis showed that cultivar resistance was the factor with the greatest influence on MWS severity, confirming the central role of the host in disease development. In addition, variables associated with water availability and higher humidity conditions, particularly during advanced vegetative and early reproductive stages of the crop, made a relevant contribution, highlighting the combined role of genetic and environmental factors in disease severity.

The model's ability to generate MWS severity predictions based on agrometeorological information from early stages of the crop cycle enables the anticipation of higher-risk scenarios before advanced reproductive phases. This level of anticipation is consistent with key decision points in crop health management and reinforces the potential of the proposed approach as a decision support tool.

References

Aegro (2023) Doenças do milho: sintomas e controle. Available at: https://blog.aegro.com.br/doencas-do-milho/.

Agrios GN (2005). Plant pathology. Elsevier.

Allen RG, Pereira LS, Raes D, Smith M (1998) **Crop Evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56**. FAO, Rome: 300(9) p. D05109.

Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G (2013) Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift** 22(6): 711–728. https://doi.org/10.1127/0941-2948/2013/0507

ANA (2021) Atlas Irrigação: uso da água na agricultura irrigada - 2. ed.

Codognoto LC, Conde TT, Cipriano LP, Cipriano TL, Azevedo Silva BA, Ferreira LR (2017) Mudanças climáticas e doenças na cultura do milho. Revista **Edutec** 2(1).

Cól MC, Braga K, Custódio AAP, Canteri, MG, Fantin LH, Utiamada CM, Campos HD, Silva-Araújo DD, Costa RV, Yada IFU (2025) Quantitative summary of nine years of fungicide efficacy data for managing white spot and protecting yield in maize. **Crop Protection** 107339.

Ageitec (2021) Doenças Foliares. Embrapa. Available at: https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/milho/producao/pragas-e-doencas/doencas/doencas-foliares.

FAO (2022) World Food and Agriculture – Statistical Yearbook 2022. Rome. https://doi.org/10.4060/cc2211en

Fenu G, Malloci FM (2021) Forecasting plant and crop disease: an explorative study on current algorithms. **Big Data and Cognitive Computing** 5(1): 2. https://doi.org/10.3390/bdcc5010002

Getap (2022) Mancha-branca no milho. Available at: https://getap.agr.br/2022/08/11/mancha-branca-no-milho/

González-Domínguez E, Caffi T, Rossi V, Salotti I, Fedele G (2023) Plant disease models and forecasting: changes in principles and applications over the last 50 years. **Phytopathology®** 113(4): 678–693. https://doi.org/10.1094/PHYTO-10-22-0362-KD.

Jensen A, Brown P, Groves K, Morshed A (2025) Next generation crop protection: A systematic review of trends in modelling approaches for disease prediction. **Computers and Electronics in Agriculture** 234: 110245. https://doi.org/10.1016/j.compag.2025.110245.

Killick R, Fearnhead P, Eckley IA (2012) Optimal detection of changepoints with a linear computational cost. **Journal of the American Statistical Association** 107(500): 1590–1598.

Kistner MB, Galiano-Carneiro AL, Kessel B, Presterl T, Miedaner T (2021) Multi-parental QTL mapping of resistance to white spot of maize (*Zea mays*) in southern Brazil and relationship to QTLs of other foliar diseases. **Plant Breeding** 140(5): 801–811.

Lima RF, de Oliveira Aparecido LE, Torsoni GB, Chiquitto AG, Moraes JR, de Souza Rolim G (2022) Agroclimatic zoning for the incidence of brown eye spot on coffee under climate change scenarios. **Theoretical and Applied Climatology** 149(3): 1471–1496. https://doi.org/10.21203/rs.3.rs-1277245/v1.

Mallikarjuna Rao GS, Dangeti S, Amiripalli SS (2022) An efficient modeling based on XGBoost and SVM algorithms to predict crop yield. In.: Borah S, Balas VE, Polkowski Z (Eds.) **Advances in Data Science and Management**. Springer Singapore, p. 565-574.

Mariadass DAL, Moung EG, Sufian MM, Farzamnia A (2022) EXtreme gradient boosting (XGBoost) regressor and shapley additive explanation for crop yield prediction in agriculture. In: 12th International Conference on Computer and Knowledge Engineering (ICCKE 2022), Mashhad, Iran, Islamic Republic of, p. 219–224.

Morkeliūnė A, Rasiukevičiūtė N, Valiuškaitė A (2021) Meteorological conditions in a temperate climate for Colletotrichum acutatum, strawberry pathogen distribution and susceptibility of different cultivars to anthracnose. **Agriculture** 11(1): 80. https://doi.org/10.3390/agriculture11010080

Nleya T, Chungu C, Kleinjan J (2016) Corn growth and development. Grow Corn Best Manag. Pract 722: 2019-09.

Nunes LH, Gabriel GH, Marengo JA (2023) More Erratic and More Extreme: Trends in Precipitation in the State of São Paulo, Brazil. **American Journal of Climate Change** 12(1): 140–171. https://doi.org/10.4236/ajcc.2023.121008

Oliveira Aparecido LE, Lorençone PA, Lorençone JA, Torsoni GB, de Lima RF, Padilha F, de Souza PS, de Souza Rolim G (2024) Addressing coffee crop diseases: forecasting Phoma leaf spot with machine learning. **Theoretical and Applied Climatology** 155(3): 2261–2282. https://doi.org/10.1007/s00704-023-04739-z.

Piepho HP, Büchse A, Emrich K (2003) A hitchhiker's guide to mixed models for randomized experiments. **Journal of Agronomy and Crop Science** 189(5): 310–322.

Rodriguez-Roa AO, de Lima RF, Fantin GM, Duarte AP, de Souza Rolim G (2025) Influence of agrometeorological variables on the severity of maize white spot and its effect on yield in tropical and subtropical regions. **International Journal of Biometeorology** 69(12): 3613-3626.

Rossi RL, Gimenez Pecci MP, Guerra FA, Plazas MC, Brücher E, Guerra GD, Torrico Ramallo AK, Camiletti BX, Maurino MF, Barontini JM (2017) Enfermedades del maíz de siembra tardía causadas por hongos. **El mismo maíz, un nuevo desafío: Compendio primer congreso de maíz tardío**. Dow AgroSciences Argentina, Buenos Aires, pp 109-127.

Sawazaki E, Dudienas C, Paterniani M, Galvao JCC, Castro JL, Pereira J (1997) Reação de cultivares de milho à Mancha de Phaeosphaeria no estado de São Paulo. **Pesquisa Agropecuaria Brasileira** 32(6): 585-589.

Silva RS, Campos HD, Ribeiro LM, Braz GBP, Magalhães WB, Bueno JN (2021) Danos na cultura do milho em função da redução de área foliar por desfolha artificial e por doenças. **Summa Phytopathologica** 46: 313–319.

Souza Almeida K, da Silva AR, Lemos NM, Coutinho WBG, Silva JM, Peixoto CC, do Carmo Leite M, da Paz-Lima ML (2019) Temporal progression of foliar plant diseases in corn hybrids. **Australian Journal of Crop Science** 13(11): 1731–1739.

Souza RG, Camargo MBP, Lania DG, Moraes JFL (2007) Classificação climática de Köppen e de Thornthwaite e sua aplicabilidade na determinação de zonas agroclimáticas para o estado de São Paulo. **Bragantia** 66: 711–720.

Tanumihardjo SA, McCulley L, Roh R, Lopez-Ridaura S, Palacios-Rojas N, Gunaratna NS (2020) Maize agro-food systems to ensure food and nutrition security in reference to the Sustainable Development Goals. **Global Food Security** 25: 100327. https://doi.org/10.1016/j.gfs.2019.100327

Thornthwaite C, Mather J (1955) The Water Balance. Drexel Institute of Technology, Laboratory of Climatology. Publications in Climatology VIII (1).

Wang G, Zhang X (2024) Maize in Global Food Security: Role and Challenges. **Field Crop** 7.

Xing E, Fan X, Jiang F, Zhang Y (2023) Advancements in Research on Prevention and Control Strategies for Maize White Spot Disease. **Genes** 14(11): 2061. https://doi.org/10.3390/genes14112061

CHAPTER 4 - Final considerations

The integration of agrometeorological, agronomic, and plant protection information, supported by historical data collected over multiple years under real production conditions, contributes to the epidemiological understanding and the development of predictive approaches for Maize White Spot. The quantification of the relationship between disease severity and yield, together with the identification of agroclimatic thresholds and critical phenological windows associated with disease progression, complements descriptive approaches through quantitative and predictive modeling. Furthermore, the application of machine learning models, complemented by explanatory analytical methods, strengthens the methodological rigor of the study by integrating predictive performance with a biologically grounded interpretation of the results.

The results obtained have direct relevance for plant protection risk management in maize production, providing technical foundations for decision-making in the management of Maize White Spot. The identification of predisposing agrometeorological conditions and critical phenological periods enables more precise targeting of monitoring strategies and fungicide applications, favoring more timely and efficient interventions. From a biological perspective, the study reinforces the understanding of the disease as the outcome of interactions among environmental conditions, crop phenology, and resistance level, generating relevant information for breeding programs and for the design of adaptation strategies under climatic variability.

The evaluation of disease behavior under climate change scenarios and the validation of the developed models in other production regions constitute research directions that may broaden their applicability and robustness. Analyzing epidemiological dynamics under potential changes in temperature and water regimes will allow the anticipation of shifts in disease intensity and distribution, while external validation will support model refinement and consolidation under diverse agroenvironmental conditions.

In addition, the results lay the groundwork for the development of digital decision-support tools integrating epidemiology, agrometeorology, and artificial intelligence for the sustainable management of maize diseases. The articulation of these approaches may facilitate the implementation of operational predictive systems aimed at improving the timeliness and efficiency of phytosanitary management strategies under conditions of climatic variability.

APPENDICES

Appendix A - Supplementary material of chapter 2

Table S1 Slope coefficient and its standard deviation for the relationships between MWS severity and agrometeorological variables across resistance groups. Symbols: * indicates significance at the 0.05 level. Blue denotes positive associations; orange denotes negative correlations

	Resistant	Moderately resistant	Susceptible
Relative Humidity	0.00 ± 0.00	0.02* ± 0.01	0.10* ± 0.02
Surface Soil Moisture	0.34* ± 0.08	0.00 ± 0.00	0.00 ± 0.00
Root Zone Soil Moisture	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Precipitation	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Dew Point Temperature	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Water Index	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Diurnal temperature range	0.00 ± 0.00	-0.12* ± 0.02	-0.46* ± 0.08
Maximum Temperature	-0.01* ± 0.01	-0.01* ± 0.02	-0.30* ± 0.04
Wind Speed	-0.05* ± 0.01	-0.18* ± 0.05	-0.70* ± 0.20
Mean Temperature	0.00 ± 0.00	-0.09* ± 0.02	-0.42* ± 0.10
Vapor Pressure Deficit	0.00 ± 0.00	0.00 ± 0.00	-2.85* ± 1.46
Solar Radiation	0.00 ± 0.00	-0.11* ± 0.03	-0.42* ± 0.09
Minimum Temperature	0.00 ± 0.00	-0.08* ± 0.02	0.00 ± 0.00
Reference Evapotranspiration	0.00 ± 0.00	-0.01* ± 0.00	-0.03* ± 0.01

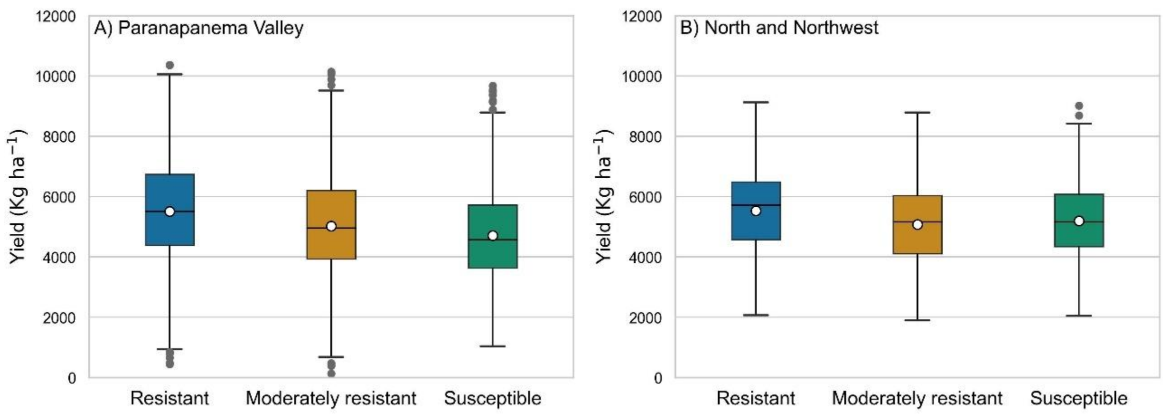


Fig. S1 Maize yield by severity range for the two regions of the state of São Paulo. Symbols: ○mean, — median, □ 50% of data, ┬┴ 99% of data, ● outliers.

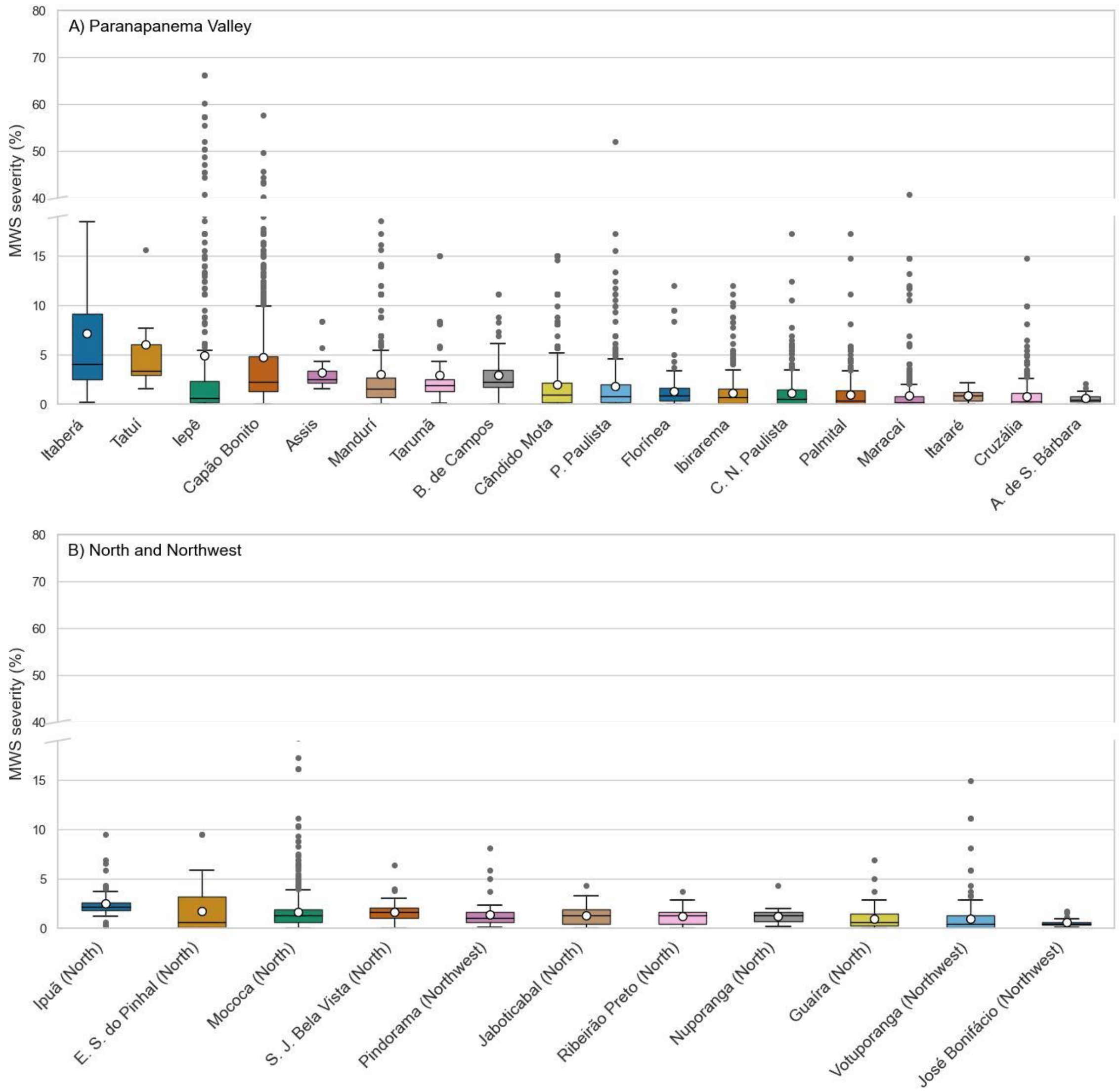


Fig. S2 MWS severity for locations in the two regions of the State of São Paulo. Symbols: ○mean, ─ median, □ 50% of data, ┬ ┴ 99% of data, ● outliers

Appendix B - Supplementary material of chapter 3

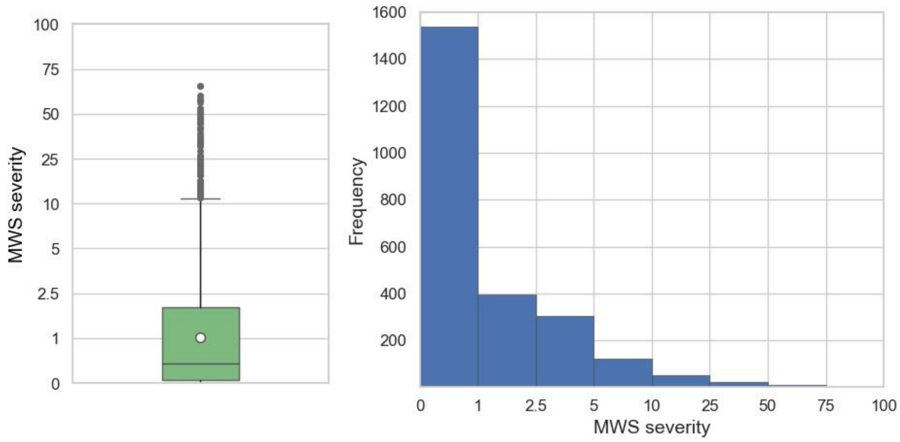


Fig. S3 Boxplot (left) and frequency histogram (right) of MWS severity in the study area. The severity axis is expressed as the equivalent percentage of leaf area affected.

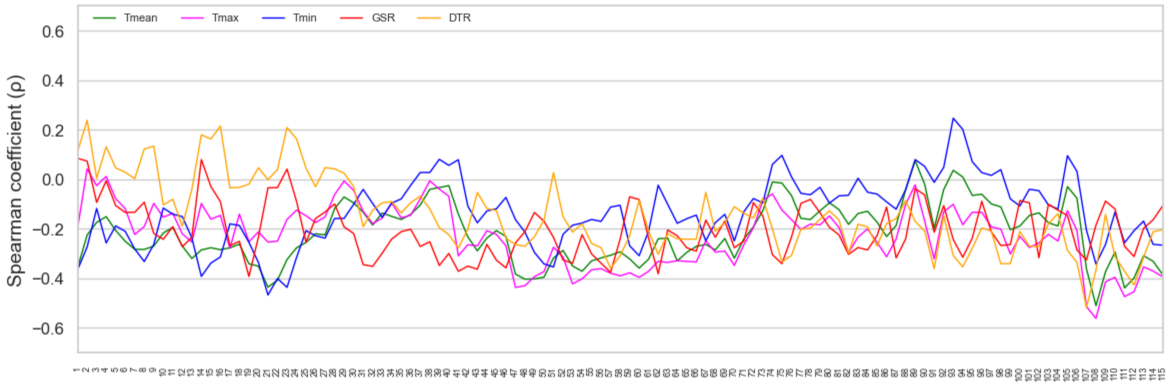


Fig. S4 Daily Spearman correlations between MWS severity and thermal agrometeorological variables: mean (Tmean), maximum (Tmax) and minimum (Tmin) temperatures, diurnal temperature range (DTR) and solar radiation (GSR).

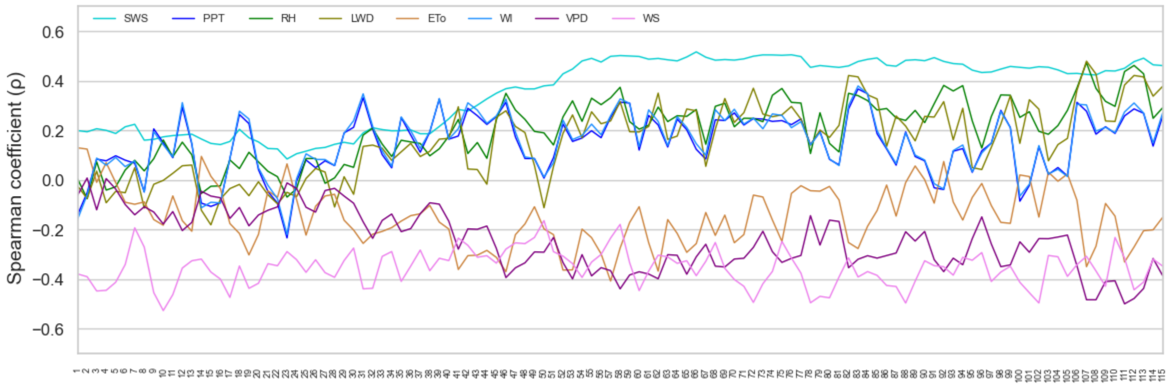


Fig. S5 Daily Spearman correlations between MWS severity and Water agrometeorological variables: Soil water storage (SWS), Precipitation (PPT), Relative humidity (RH), Leaf wetness duration (LWD), Reference evapotranspiration (ETo), Water Index (WI), Vapor pressure deficit (VPD), wind speed (WS).

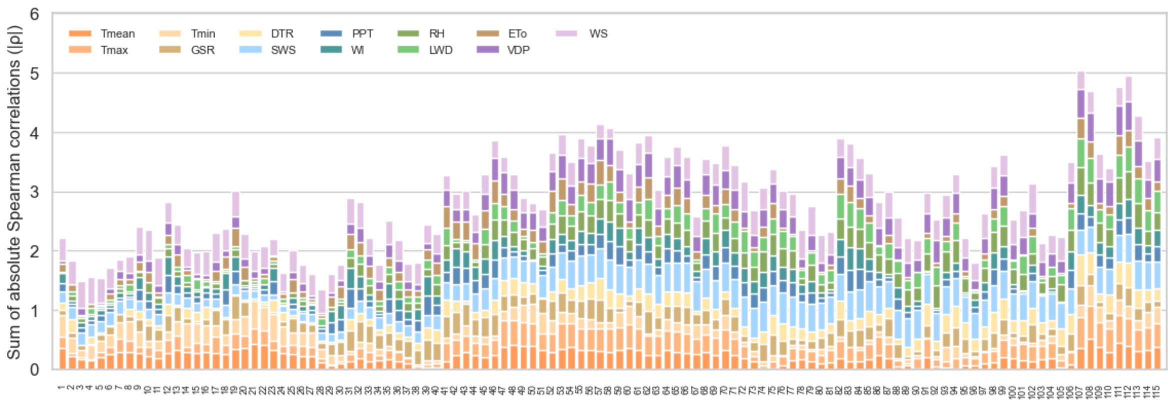


Fig. S6 Sum of the magnitudes of the daily correlations of the agrometeorological variables. Symbols: Tmean, mean temperature; Tmax, maximum temperature; Tmin, minimum temperature; PPT, precipitation; RH, relative humidity; GSR, solar radiation at the Earth's surface; WS, wind speed; ETo, reference evapotranspiration; VDP, vapor pressure deficit; DTR, diurnal temperature range; SWS, soil water storage; LWD, leaf wetness duration; WI, atmospheric water availability index.

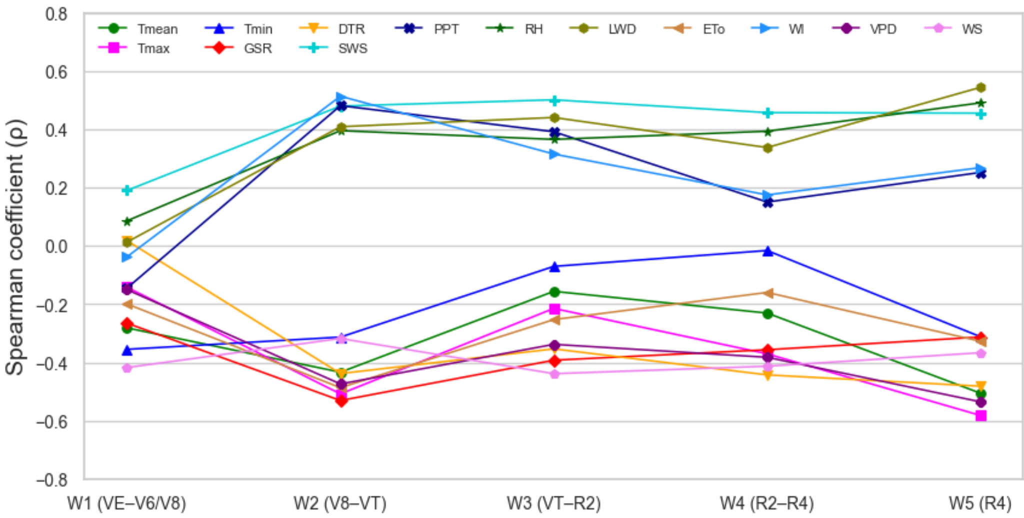


Fig. S7 Spearman correlation coefficients between agrometeorological variables and MWS severity aggregated within the defined agroclimatic time windows (W1–W5). Symbols: Tmean, mean temperature; Tmax, maximum temperature; Tmin, minimum temperature; PPT, precipitation; RH, relative humidity ; GSR, solar radiation at the Earth's surface; WS, wind speed; ETo, reference evapotranspiration; VDP, vapor pressure deficit; DTR, diurnal temperature range; SWS, soil water storage; LWD, leaf wetness duration; WI, atmospheric water availability index. W1, emergence–mid vegetative growth; W2, advanced vegetative growth–male flowering; W3, flowering–kernel blister onset; W4, kernel blister–kernel dough onset; W5, kernel dough stage.

Table S2 Descriptive statistical values of the agrometeorological variables calculated from the average of the 115 days prior to the assessment of MWS severity. The mean

(mean), standard deviation (std), minimum (min), maximum (max), and 25%, 50%, and 75% percentiles are presented.

	mean	std	min	25%	50%	75%	max
Tmean (°C)	20.25	1.42	16.89	19.56	20.41	21.05	23.87
Tmax (°C)	26.21	1.64	21.70	25.52	26.41	27.31	29.70
Tmin (°C)	15.45	1.27	13.16	14.73	15.43	16.05	19.01
GSR (MJ m^{-2} d^{-1})	15.32	1.59	11.22	14.77	15.37	16.16	18.72
DTR (°C)	10.76	1.04	8.37	10.34	10.69	11.26	14.10
SWS (mm)	63.50	24.59	23.55	46.04	52.55	86.36	112.32
PPT (mm)	327.40	105.97	93.45	264.57	317.06	410.88	564.28
RH (%)	78.88	4.34	65.41	76.34	79.40	81.00	87.78
LWD (h d^{-1})	6.86	1.14	3.52	6.37	7.12	7.62	8.98
Eto (mm)	313.33	33.82	244.56	293.43	313.30	331.91	393.16
WI	2.65	1.76	0.52	1.34	1.88	4.07	6.47
VPD (kPa)	0.52	0.13	0.24	0.44	0.50	0.61	0.83
WS (m s^{-1})	1.55	0.56	0.36	1.32	1.64	1.99	2.29

Table S3 Optimal hyperparameter configurations for the evaluated machine learning models identified through grid search optimization with cross-validation during the training phase.

Model	Optimal hyperparameters
Random Forest	*n_estimators* =1100, *criterion* = poisson, *max_depth* = None, *min_samples_split* =10, *min_samples_leaf* =4, *max_features* = None, *bootstrap* =True
Support Vector Regression	*kernel* = rbf, *gamma* = auto, *C* = 15, *epsilon* = 0.2
Multilayer Perceptron	*hidden_layer_sizes* = (128, 64, 32), *activation* = tanh, *solver* = adam, *alpha* = 1e−06, *learning_rate_init* = 0.001, *max_iter* = 500, *early_stopping* = True, *validation_fraction* = 0.1, *random_state* = 42
XGBoost	*learning_rate* = 0.01, *max_depth* = 6, *subsample* = 0.8, *colsample_bytree* = 0.8, *reg_lambda* = 10, *reg_alpha* = 1.0, *n_estimators* = 800, *objective* = reg:squarederror, *tree_method* = hist, *random_state* = 42
Multiple Linear Regression	Not applicable