Spatiotemporal Variability of Apple Tree Vegetative Vigor Using Proximal Sensing

Luciano Gebler[1], Andrea de Rossi[1], Jardel Talamini de Abreu[2], Eduardo Antonio Speranza[3], Alessandra Sessi[1], Lucas de Ross Marchioretto[1]

A paper from the Proceedings of the
17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture
13-15 July 2026
Porto Alegre, Brazil

Abstract.

Apple production in southern Brazil often suffers from excessive vegetative growth due to the region's subtropical climate. Currently, farmers rely on empirical, subjective monitoring to make crucial management decisions regarding pruning, thinning, and the application of growth regulators. Coupled with increasing labor shortages and costs, this empirical approach often leads to operational inefficiencies and compromised fruit quality. To address this, there is a critical need for quantitative, data-driven methods to monitor vegetative vigor and generate precise spatial maps for targeted orchard management. This study investigates the use of multispectral proximal sensors to measure the Normalized Difference Vegetation Index (NDVI) as an objective indicator of plant biomass, replacing traditional empirical assessments. Linear and non-linear regression models were used to correlate NDVI with bud growth rates in two apple orchards in Vacaria-RS across six growing seasons. The analysis incorporated various time windows (ranging from 5 to 11 days) alongside meteorological variables such as accumulated precipitation, solar radiation, and thermal sum. The results indicate that the optimal periods for data collection to accurately identify vegetative vigor variability are windows of 7 to 11 days, specifically occurring between 55 and 67 days after dormancy break. Identifying these crucial periods enables the generation of accurate temporal vigor maps and the delineation of specific management zones. Ultimately, this methodology allows producers to objectively adjust the intensity of vigor control practices according to specific zones, optimizing resource allocation, reducing labor dependency, and ensuring optimal fruit yield and quality.

Keywords.
NDVI, branch growth rate, management zones

Introduction

Over the last few decades, Brazilian apple production has expanded significantly, primarily driven by the cultivation of the 'Gala', 'Fuji', 'Pink Lady', and 'Eva' cultivars. According to the Food and Agriculture Organization (FAO, 2022), Brazil currently ranks among the top ten apple-producing countries globally. Domestically, apples are the fourth most consumed fruit, generating approximately 1.97 billion BRL in revenue and yielding an annual harvest of over 1.04 million tons (IBGE, 2022).

Historically, research regarding fruit tree nutrition in Brazil has largely focused on maintenance fertilization, particularly nitrogen application. However, crop yield responses to added nitrogen have frequently proven non-significant in the deep soils of traditional growing regions, such as Fraiburgo and Vacaria (Basso & Suzuki, 1992; Ernani et al., 1997; Proença et al., 2002).

Regional climatic conditions profoundly influence apple tree phenology and vegetative development. Under the climatic conditions of Southern Brazil, the vegetative growth period for 'Gala' and 'Fuji' apple trees is notably longer than that observed in typical temperate regions. This extended period, when combined with high temperatures and substantial rainfall during the growing cycle, frequently results in excessive vegetative vigor, particularly in years with low fruit set (Hawerroth & Petri, 2014). Consequently, implementing cultural practices that promote a proper balance between vegetative growth and fruiting is crucial for enhancing both productive efficiency and fruit quality. Crop load management, for instance, is an essential strategy to ensure this vegetative-reproductive equilibrium and to properly control overall canopy vigor (Pereira & Petri, 2006).

Given the rising costs and scarcity of agricultural labor, highly precise, phenology-based crop monitoring is essential. By generating temporal vigor maps throughout the growing cycle, farmers can implement targeted, site-specific management to optimize resource allocation without compromising fruit quality or yield.

Vegetation indices derived from optical sensors have become important tools for monitoring spatial and temporal variability in perennial crops. Among them, the Normalized Difference Vegetation Index (NDVI) has been widely used as an indicator of canopy development, biomass accumulation, and plant vigor due to its sensitivity to changes in photosynthetically active vegetation. In vineyards and fruit orchards, spectral indices have been successfully employed to estimate canopy characteristics, yield variability, and fruit quality attributes, supporting the delineation of management zones and site-specific interventions (Hall et al., 2010; Zarco-Tejada et al., 2005; Santos et al., 2022). Consequently, proximal sensing technologies represent a promising alternative for transforming qualitative field observations into objective and spatially explicit information for decision-making.

To overcome the limitations of subjective empirical assessments, this study evaluates the use of the Normalized Difference Vegetation Index (NDVI) from proximal sensors to quantify apple tree vegetative vigor. By applying linear and non-linear regression models to data collected over six growing seasons in Vacaria-RS, we correlated NDVI with bud growth rates and key climatic variables. Ultimately, this research identifies the optimal temporal windows for data collection—specifically 55 to 67 days post-dormancy break—to accurately map spatial variability. Establishing these critical periods enables the creation of precise management zones, allowing decision-makers to implement site-specific vigor control in commercial orchards.

Materials and Methods

Study Area

The field experiments were conducted in two commercial apple orchards, designated as "Fitotecnia" and "Gestão", located at the Temperate Climate Fruit Forestry Experimental Station of the Brazilian Agricultural Research Corporation (Embrapa Grape and Wine) in Vacaria, Rio Grande do Sul, Brazil (28°30'58.2"S, 50°52'52.2"W). Both orchards consisted of 15-year-old 'Maxi Gala' apple trees (*Malus domestica Borkh*).

According to the Köppen climate classification, the regional climate is categorized as Cfb, characterized as a humid subtropical climate with mild summers and cold winters (Peel et al., 2007). Historical meteorological data from the National Institute of Meteorology (INMET) indicate a well-distributed annual precipitation averaging 1,400 mm, with mean maximum and minimum temperatures of 25°C and 15°C, respectively, during the growing season.

NDVI from proximal sensing and manual measurement of branch length

To assess the spectral response of the crop, canopy Normalized Difference Vegetation Index (NDVI) was measured using a FLEXUM® active optical sensor (Falker, Brazil). Continuous georeferenced readings (1 Hz) were performed laterally at a constant speed of 7 km/h. The sensor was positioned approximately 1 m from the canopy, capturing a 1 m diameter area at the mid-height of the plants to ensure high-density spatial data collection. Emphasizing the temporal monitoring of the crop, NDVI data were collected weekly throughout the vegetative cycle over six consecutive growing seasons (from 2020-2021 to 2025-2026).

For the ground-truth assessment of vegetative growth, 92 georeferenced apple trees were continuously monitored (42 in the 'Fitotecnia' orchard and 50 in the 'Gestão' orchard). On each tree, four current-year branches were tagged in the middle section of the canopy, strictly aligning with the optical sensor's field of view. Branch length was measured weekly, from the annual growth scar to the apical bud at the base of the newest expanding leaf. In total, 368 branches were tracked weekly across the six seasons, providing a robust temporal dataset of crop vegetative vigor to be correlated with the spatial sensor readings.

Climate data

Meteorological data was obtained from a weather station located 100 meters north from the experimental orchards, belonging to the National Institute of Meteorology (INMET). Data was organized in R software version 4.5.2 (R Core Team, 2025) through the package dplyr version 1.2.0 (Wickham et al., 2026). The accumulated precipitation, solar radiation and thermal sum data was accounted from the day of dormancy breaking until each NDVI reading. Accumulated precipitation and solar radiation consisted of the sum of the values logged each day, whereas thermal sum consisted of the sum of the daily mean temperature above 4.4 °C. This basal temperature was considered as for apples it presents tissue growth (Greene et al., 2013).

Data preprocessing

All collected data were compiled into a single database, containing, in each record, the following attributes: crop year, days after dormancy break (DADB), NDVI, branch growth rate (BGR, in cm), thermal summation (TS, in degree-days), accumulated precipitation (AP, in mm) and accumulated radiation (AR, in Mj/m^2), in addition to attributes related to geolocation (latitude, longitude and plant ID). Due to the use of climatic data from a single station, this data is repeated for all plants in each DADB. In total, considering the two monitored orchards and the six crops, we have 5550 records available with this structure.

Data from the 2025-2026 crop season (a total of 900 records) were selected for validation of the regression models and were not used for the training and testing phases. Data from the 2020-2021 to 2024-2025 crop seasons (a total of 4650 records) were initially aggregated into different time windows ranging from 5 to 78 days, to initially identify the most suitable time period for the models. After identifying this period, only the selected data were reused to generate new regression models, considering in this case 70% of the data for training and 30% for testing phase.

Linear and non-linear regression models

In this work, three different linear and non-linear regression methods were evaluated to obtain models that allow the generation of vegetative vigor maps from proximal NDVI and climatic data, considering the branch growth rate as a target variable and characteristic for this measure. Ridge Regression, a penalized linear model, was employed as a robust baseline due to its capacity to mitigate multicollinearity and prevent overfitting through L2 regularization (Hoerl & Kennard, 1970). To capture non-linear dynamics while maintaining high interpretability, a Generalized Additive Model (LinearGAM) was implemented; this non-linear framework models the response variable as a sum of smooth functions of the predictors, offering greater flexibility than standard linear methods without rigid structural assumptions (Hastie, 2017). Finally, Random Forest, a powerful non-linear ensemble learning technique, was utilized to map complex, high-dimensional interactions; by aggregating multiple decision trees, this algorithm provides exceptional predictive accuracy and inherent resilience to outliers and non-parametric data distributions (Breiman, 2001).

Statistical validation

The coefficient of determination (R^2) was employed to quantify the proportion of variance in the field-measured branch growth real rates explained by the NDVI and climatic predictors. Across both linear and non-linear model architectures, R^2 serves as a standardized, practical metric to assess how effectively these combined environmental variables capture the observed variations in crop vigor (Kvålseth, 1985).

Results and Discussion

Considering the data from the first 5 harvests, regressions were performed using different combinations for aggregated values of mean, maximum, and sum (cumulative) of the attributes NDVI, TS, AP and AR as predictor variables and BGR as the predicted variable, regarding windows of days of different sizes (ranging frDADBom 5 to 78 days), with an initial DADB of 34 and a final DADB of 112. Results obtained by the three regression methods pointed out that the period between 56 and 67 DABD obtained average R^2 values above 0.95, showing itself to be a promising period for the generation of specific models that can actually indicate the variability of vegetative vigor in the studied orchards as a function of NDVI and climatic attributes.

In the next step, the dataset was filtered by the aforementioned period, limiting it to a total of 300 samples, with 210 for training and 90 for testing. Tab. 1 shows the results obtained by the three regressors, for training and testing.

Tab. 1. Coefficient of determination (R^2) between observed and predicted values of BGR as a function of NDVI and climatic attributes.

Regression Method	Training (R^2)	Testing (R^2)
Ridge	0.6072	0.6423
LinearGAM	0.9598	0.9675
RandomForest	0.9876	0.9597

The significant performance gap between the LinearGAM and RandomForest models compared to Ridge model strongly indicates that the dataset contains complex, non-linear relationships that a strictly linear approach cannot capture. RandomForest, as an ensemble of decision trees, naturally maps intricate non-linearities and high-order feature interactions without requiring explicit mathematical transformations, while LinearGAM relaxes strict linearity by applying flexible, smooth non-linear functions to individual features. In this way, LinearGAM and RandomForest succeed because their structural flexibility aligns with the actual non-linear dynamics of data, whereas Ridge fails due to its rigid linear limitations.

Following these results, validation was performed only for the LinearGAM and RandomForest models, considering unknown data from these models regarding the 2025-2026 crop season (50 samples obtained between 56 and 67 DADB). In this case, the RandomForest-based model performed much better than the LinearGAM-based model, with R^2 values of 0.69 and 0.47, respectively. Figure 1 shows a graph containing the predicted and measured BGR values

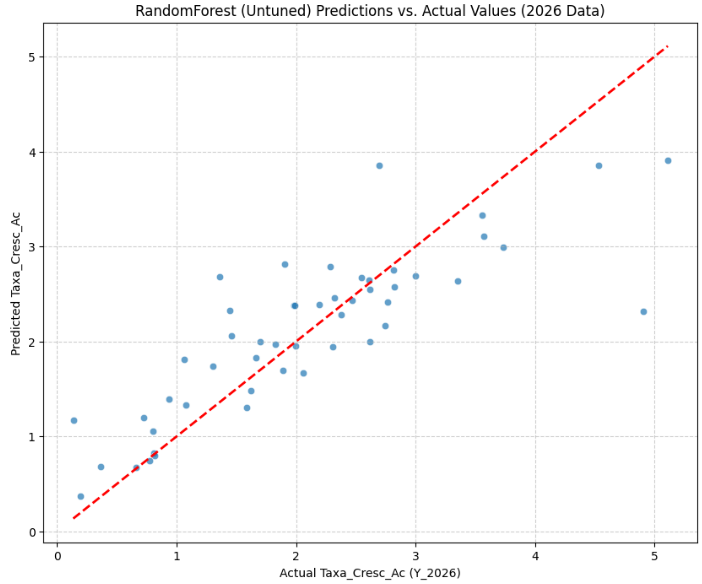

Fig. 1. Measured (X) and predicted BGR values regarding RandomForest model (Y).

Conclusion

This study successfully demonstrates the viability of utilizing proximal NDVI and climatic data to predict the vegetative vigor of 'Maxi Gala' apple trees, pinpointing a critical temporal window of 56 to 67 days after dormancy break for optimal spatial variability mapping. By evaluating different regression architectures, the research established that the underlying relationship between environmental variables and branch growth rate is highly complex and non-linear. Consequently, the RandomForest and LinearGAM models significantly outperformed the baseline Ridge regression during the initial training and testing phases. Crucially, when applied to unseen validation data from the 2025-2026 growing season, the RandomForest model proved to be the most robust approach, achieving a notably higher predictive performance than LinearGAM and underscoring the efficacy of ensemble decision trees in capturing intricate crop dynamics.

Despite these promising findings, the study presents specific limitations, most notably the substantial drop in predictive accuracy during the final validation phase compared to the near-perfect scores of the testing phase, which suggests a degree of overfitting or a challenge in generalizing the models to entirely new harvest years. Furthermore, the reliance on climatic data from a single meteorological station represents a structural flaw, as it forces the same environmental values onto all plants and fails to account for the micro-climatic spatial variability that naturally exists within commercial orchards. Future work should directly address these constraints by integrating high-resolution, localized IoT climatic sensors to refine the dataset and exploring advanced hyperparameter tuning to improve year-over-year model generalization. Additionally, extending this methodology to other economically significant cultivars and evaluating the practical impact of these predictive vigor maps on variable-rate management strategies in the field will be essential next steps.

Acknowledgments

This research was funded by FAPESP – São Paulo Research Foundation (Proc. 2022/09319-9 for funding and Proc. 2025/28583-7 for scholarship grant), under the CCD-AD/SemeAr Project. The authors also acknowledge Embrapa Grape and Wine for their institutional support.

References

Basso, C.; Suzuki, A. Resposta da macieira cv. Golden Delicious à adubação nitrogenada. Revista Brasileira de Ciência do solo, Campinas, v.16, n.2, p.223-227, 1992.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Ernani, P.R.; Dias, J.; Vanz, L. Application of nitrogen to the soil after fruit harvest has not increased apple yield. Revista Brasileira de fruticultura, Cruz das Almas, v. 9, n.1, p.33-37, 1997.

FAO – Food and Agriculture Organization of the United Nations. (2022). FAOSTAT Statistical Database. Available at: https://www.fao.org/faostat/.

Greene, D. W., Lakso, A. N., Robinson, T. L. & Schwailler, P. (2013). Development of a fruitlet growth model to predict thinner response on apples. HortScience, 48(5), 584-587

Hall, A., Lamb, D. W., Holzapfel, B. P., & Louis, J. P. (2010). Within-season temporal variation in correlations between vineyard canopy and winegrape composition and yield. Precision Agriculture, 12(1), 103–117.

Hastie, T. J. (2017). Generalized additive models. Statistical models in S, 249-307.

Hawerroth, F. J.; Petri, J. L. Crescimento vegetativo de macieiras 'Fuji Suprema' sob influência da época de aplicação de proexadiona calcica. Revista Brasileira de Fruticultura, Jaboticabal, v. 36, n. 2, p. 373-380, 2014.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67.

IBGE – Instituto Brasileiro de Geografia e Estatística. (2022). Produção Agrícola Municipal: Maçã. Available at: https://www.ibge.gov.br/.

Kvålseth, T. O. (1985). Cautionary note about R 2. The American Statistician, 39(4), 279-285.

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen–Geiger climate classification. Hydrology and Earth System Sciences, 11, 1633–1644.

Pereira, A. J.; Petri, J. L. Poda e condução da macieira. In: EPAGRI. A cultura da macieira. Florianópolis, 2006. p. 391-418.

Proença, M. M.; Ernani, P. R.; Dias, J. Rendimento e qualidade de frutos de macieira influenciados pela aplicação de N mineral em pomar com porta-enxerto anão. In: REUNIÃO SUL BRASILEIRA DE CIÊNCIA DO SOLO, 4., 2002. Porto Alegre. Solos e qualidade ambiental: resumos... Porto Alegre: Universidade Federal do Rio Grande do Sul e Núcleo Regional Sul da Sociedade Brasileira de Ciência do Solo, 2002. CD-ROM

Wickham H, François R, Henry L, Müller K, Vaughan D (2026). _dplyr: A Grammar of Data Manipulation_. doi:10.32614/CRAN.package.dplyr <https://doi.org/10.32614/CRAN.package.dplyr>, R package version 1.2.0, <https://CRAN.R-project.org/package=dplyr>.

R Core Team (2025). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. <https://www.R-project.org/>.

Santos, M. S., Gebler, L., & Sebem, E. (2022). Correlation between vegetation indexes generated at Vitis vinifera L. and soil, plant and production parameters for emergency application in decision making. Ciência Rural, 52(2).

Zarco-Tejada, P., Berjón, A., López-Lozano, R., Miller, J. R., Martín, P., Cachorro, V., González, M. R., & De Frutos, A. (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. Remote Sensing of Environment, 99(3), 271–287.

Authors and Affiliations

[1] Brazilian Agricultural Research Corporation (Embrapa), Bento Gonçalves, RS, Brazil.

[2] Agriexata, Santa Maria, RS, Brazil.

[3] Brazilian Agricultural Research Corporation (Embrapa), Campinas, SP, Brazil.

Corresponding author: Eduardo Antonio Speranza (eduardo.speranza@embrapa.br)