NOTE

# The Enviromic marker

**Gustavo E. Marcatti**[1] **and Rafael T. Resende**[2*]

Crop Breeding and Applied Biotechnology
26(1): e54222611, 2026
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332026v26n1n1

**Abstract:** *We formalize "enviromic markers" as modeling units parallel to DNA markers, but herein for genotype–environment (G × E) prediction. Four operational premises (linearity; site potential; heterogeneous favorability; and envirotypic covariates (ECs)–genotype-dependence) are presented to enable their use in linear mixed models and also to motivate four construction strategies: (i) using raw environmental covariates as linear markers; (ii) applying transformations to capture mild nonlinearities; (iii) deriving ecophysiological functions; and (iv) engineering markers with Artificial Intelligence (AI) models which learn nonlinear environment → phenotype mappings for linear downstream use. Environmental data quality control is detailed, including checks of spatial coverage and resolution, variance within the TPE, collinearity control, and spatial/temporal validation without leakage. Envirome data are linked with GIS to compute environmental kernels, quantify covariate shifts, and deliver pixel-level predictions with uncertainty diagnostics. The framework clarifies assumptions and standardizes the use of enviromic markers for predictive breeding analyses.*

**Keywords**: *Envirotyping, genotype-by-environment interaction (G × E), environmental or envirotypic covariates (ECs), target population of environments (TPE), geographic information systems (GIS)*

## INTRODUCTION

A recurring hurdle for biometric plant breeding is figuring out how to incorporate environmental variability into prediction. Genotype response strongly depends on environment, so cultivars can perform differently across contexts (Fradgley et al. 2025). To address this, envirotyping (i.e., environmental-typing) was proposed: a framework that systematizes recording the environmental conditions of each trial or production site by analogy to genetic typing or DNA genotyping (Xu 2016). From this comes the idea of an enviromic marker, an index vector derived from environmental (or envirotypic) covariates (ECs) which reflects the potential expression of a phenotype in a given space and time (Resende et al. 2021). Each georeferenced pixel can be treated as carrying an environmental vector, functioning like an environmental genetic profile. Recent studies indicate that this approach can anticipate yield trends under future scenarios and support region-specific genotypic recommendations (Resende et al. 2024, Zhang et al. 2024, Araújo et al. 2024).

Enviromic markers complement genetic markers within the genotype-by-environment interaction (G × E) framework. While genetic markers describe each individual's allelic identity, enviromic markers describe the spatiotemporal evaluation context. Xu (2016) noted that systematic collection of climate, soil, and management provides a basis for understanding phenotypic variation,

***Corresponding author:**
E-mail: rafael.tassinari@ufg.br

**Received:** 15 October 2025
**Accepted:** 24 November 2025
**Published:** 01 December 2025

[1] Universidade Federal de São João Del Rei, Campus Sete Lagoas (CSL), Departamento de Engenharia Florestal, Rodovia MG-424, km 47, s/n, 35702-031, Sete Lagoas, MG, Brazil
[2] Universidade Federal de Goiás, Escola de Agronomia, Campus Samambaia, Avenida Esperança, s/n, 74690-900, Goiânia, GO, Brazil

with these covariates typically undergoing basic quality control (QC) before they are modeled. The comprehensive set of envirotyped variables forms the envirome, which modulates gene expression and crop performance (Heslot et al. 2014). According to Resende et al. (2021), structuring this information as environmental vectors equivalent to markers allows environmental data to be handled like SNPs, interacting with the genetic background in the model. Recent applied work shows that enviromic marker engineering is an upstream step in which ECs are structured and refined prior to model fitting (Montesinos-López et al. 2024, Resende et al. 2025). Nonlinear kernel methods have increased predictive performance at a distinct stage of the pipeline in multi-environment modeling (Costa-Neto et al. 2021).

Although these innovations expand what is possible, practical and conceptual limits remain. Many genetic selection models do not explicitly capture environmental variation across space and time, which weakens extrapolation to new sites (Cooper and Messina 2021). Cruz et al. (2025) stress defining the Target Population of Environments (TPE) before analysis, since this shapes prediction validity. Adding environmental covariates can improve accuracy, yet how to structure them as linear markers in mixed models remains unsettled. Tolhurst et al. (2022) suggest that regressions on latent covariates can reveal hidden environmental patterns, but they do not specify how large-scale matrices should be built. Standard nonlinear methods also struggle to represent complex trait-by-environment responses (Montesinos-López et al. 2024). Evidence shows that specific environmental descriptors account for important portions of phenotypic variation (Rogers and Holland 2022), and EC-informed decompositions indicate that covariate definition directly affects G×E interpretation (Mumford et al. 2023). Altogether, these observations reinforce the need for more flexible strategies to interpolate and model environmental structure.

In this context, this note revisits the concept of the *enviromic marker*, sets clear operational premises, and compares it with genetic markers. Recent cases show that Engineered Enviromic Markers (EEMs) yield substantial prediction gains in external validations (Resende et al. 2025), indicating that sharpening the concept can directly benefit breeding programs.

## RELATIONSHIP BETWEEN THE ENVIROMIC MARKER AND GENETIC MARKERS

Genetic and enviromic markers differ in concept and use. A traditional genetic marker refers to the allelic state at a specific genomic locus, typically discrete (0/1/2) and heritable (Park et al. 2025). In contrast, an enviromic marker is the value of an envirotyped covariate, or a function of it, indexed to a specific spatial point and time window (Resende et al. 2021, Resende et al. 2025). While the genome is static for each individual, enviromic markers vary across evaluations due to environmental fluctuations and seasonality. Cooper et al. (2014) and Xu (2016) introduced envirotyping, showing that climate and soil factors can be measured at multiple scales, from field trials to remote sensing, and used in G × E analyses analogous to genotyping. Thus, each trial site can be described by a continuous vector of environmental covariates. Each covariate assumes a specific value at that site, an envirotype, in analogy to how each genetic marker assumes an allelic state within a genotype.

A genetic marker has a defined physical position on the genetic map and exhibits allelic polymorphism within a population. An enviromic marker is tied to a georeferenced raster or shapefile, with defined spatial and temporal resolution (Resende et al. 2025). In turn, enviromics uses kernels or distances between environmental vectors to represent environmental relatedness across sites (Jarquín et al. 2014, Costa-Neto et al. 2021). Informationally, genetic markers encode inherited polymorphisms that are generally stable and depend on accurate genotyping, whereas environmental measurements are continuous or discretized signals subject to sensing noise, missingness, cloud cover and cross-sensor calibration issues, requiring careful envirotyping quality control (Resende et al. 2024, Cruz et al. 2025). Moreover, while genetic mutations are rare, climate trends and land-use change can rapidly shift environmental distributions, which demands continuous envirotyping updates (Cooper and Messina 2021, Zhang et al. 2024).

Both marker types drive genetic reactions in G×E models, but in complementary ways. Genetic markers encode the genomic contrasts which shape each genotype's baseline potential, whereas enviromic markers specify the environmental signals that trigger different reaction norms. For example, this helps identify genotypes that maintain performance under stress or yield more in limiting conditions (Resende et al. 2022). Structuring data as environmental vectors enables affinities among trial sites to be computed, reflecting patterns like spatial autocorrelation, where nearby areas share similar conditions. GIS-based G × E studies already use this to extrapolate performance to untested locations (Marcatti et al. 2017). Finally, reaction norms arise in phenotypic prediction from genetic-marker effects interacting with enviromic markers, which encode environmental variation and structure the phenotype surface within the G × E framework.

## THE FOUR PREMISES OF AN ENVIROMIC MARKER

We adopt four operational premises for practical use in predictive models which place enviromic markers on comparable footing with genetic markers for fitting and effect estimation.

- LINEARITY: The enviromic marker should have a linear relationship with the phenotype, analogous to treating SNPs in linear reaction models.
- SITE POTENTIAL: A site with a high value for a given enviromic marker (a high envirotype) tends to offer greater potential for expressing the trait, aligning with the environmental-productivity concept of Finlay and Wilkinson (1963).
- HETEROGENEOUS FAVORABILITY: A site favorable for one genotype or group is not necessarily favorable for all, genotype responses to the same environmental marker are heterogeneous.
- EC–GENOTYPE-DEPENDENCE: An environmental (or envirotypic) covariate (EC) influential for one genotype or group may not be influential for others, indicating that envirotypic polymorphism does not imply uniformity of genetic response.

Scope notes. Linearity in Premise 1 must hold in the final step of the enviromic model; causal EC→Y relations may be nonlinear before marker engineering. Premise 2 is trait- and time-window specific. Premises 3-4 justify random slopes by genotype and parsimonious EC selection. Use simple transformations or markers derived from nonlinear models for sigmoidal or asymptotic responses, then keep the final fit linear.

## STRATEGIES FOR WORKING WITH ECS AND ENVIROMIC MARKERS

Enviromic markers can be incorporated into several modeling frameworks. Linear mixed models are a common option for including ECs or EEMs as regressors interacting with the genetic background (Resende et al. 2021, Resende et al. 2025). Bayesian mixed models provide an alternative through shrinkage, hierarchical priors, and uncertainty propagation. Reaction-norm formulations integrating environmental covariates into kernel-based genomic prediction were formalized by Jarquín et al. (2014) and later adapted to enviromic contexts with nonlinear kernels by Costa-Neto et al. (2021). Araújo et al. (2024) combined thematic maps, envirotyping, and factor-analytic modeling, using partial least squares (PLS) to predict environmental loadings within a GIS-FA workflow. Feature-engineering procedures (Montesinos-López et al. 2024) adjust the covariate signal before modeling and influence marker construction. Enviromic information can also feed AI-based predictors and crop-growth modeling (CGM) pipelines that integrate environmental descriptors with mechanistic or hybrid schemes (Xu et al. 2022).

There are different ways to incorporate ECs as enviromic markers in genetic prediction or inference models, and four core strategies are:

**(1) Raw covariate as marker:** use the EC itself as the enviromic marker. The value that this marker takes at each site, its envirotype, is modeled as a linear predictor. The EC–phenotype relationship herein is assumed global and linear. ECs enter in mixed models as linear explanatory variables (Jarquín et al. 2014, Trevisan et al., 2025). Studies at macroscopic scales show that high-resolution envirotyping with thousands of ECs can yield consistent, usable predictions. Bahia et al. (2025) illustrate this in upland rice, while Resende et al. (2024) show how remote sensing enlarges the spatial representation of the envirome. These results support the view that directly using ECs as enviromic markers is a practical way to delineate recommendation zones and guide breeding decisions.

**(2) Transformations of ECs:** apply mathematical functions (e.g., square, cubic, log, exponential, splines) to the original ECs to capture simple nonlinearities. Each transformed marker produces a different envirotype value at each site, expanding the set of candidate enviromic markers. This strategy captures part of the curvature seen in field responses but may still be limited for strongly asymptotic or sigmoidal behavior. Recent feature-engineering pipelines can automate transformations, integrate derived covariates into genomic models, and improve predictive performance (Montesinos-López et al. 2024).

**(3) Construction of ecophysiological variables derived from ECs:** combine environmental covariates into physiological or ecophysiological functions which represent crop processes (e.g., growing degree days, photothermal units, thermal or water-stress metrics, and radiation- or evapotranspiration-based combinations). These mechanistic functions translate raw ECs into functionally interpretable measures more directly linked to plant physiology and field performance. Nonlinear and time-dependent formulations can represent thresholds, saturation, and phase-specific responses. Stress indicators that integrate variation across the crop cycle often outperform raw climatic descriptors by capturing cumulative environmental effects (Rincent et al. 2019). Such indices reduce noise and stabilize EC-phenotype relations within enviromic models (Costa-Neto and Fritsche-Neto 2021).

**(4) Construction of enviromic markers via AI:** generate synthetic markers from nonlinear models that learn the EC→phenotype relation. This marker engineering is analogous to deep-learning uses in genomics but applied to environment. Resende et al. (2025) exemplify the strategy by training AI models, such as random forests, with thousands of covariates to interpolate phenotypes across the study area, producing panels with about 10,000 synthetic enviromic markers. These markers capture nonlinear field patterns, including exponential, sigmoidal, and multimodal responses, and increase predictive ability. Zhang et al. (2024) show that AI-based models allow broader exploration of environmental space, though at higher computational cost and with strong dependence on data quality.

Strategy 1 provides an efficient baseline when the EC → phenotype relation is approximately linear (Figure 1-A) and using the EC directly as a marker. Implementation in mixed models is simple, interpretation is direct, and cost is relatively low. You must standardize scale, check collinearity, and avoid extrapolation outside the training environmental envelope.

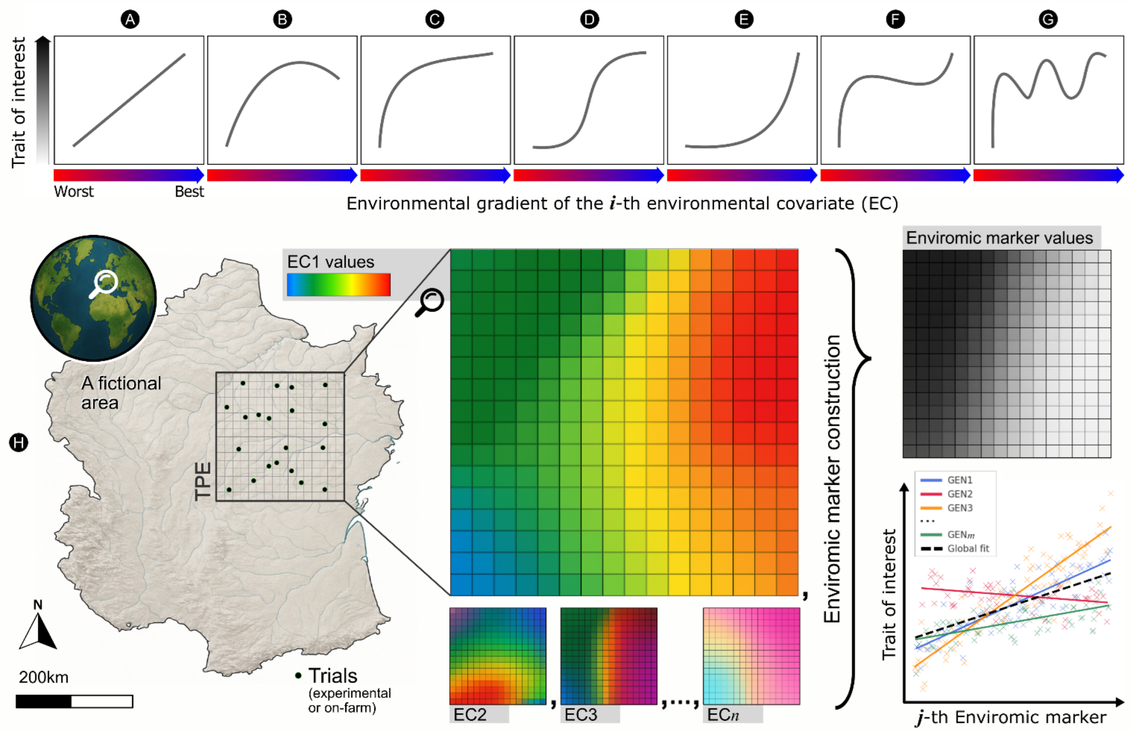

***Figure 1.*** Conceptual representation of enviromic marker construction. Panels (A–G) illustrate possible relationships between an environmental covariate (EC, x-axis: gradient “worse → better”) and phenotypic expression (y-axis): (**A**) increasing linear (model: linear term); (**B**) concave quadratic (2nd-degree polynomial); (**C**) increasing asymptotic (e.g., asymptotic or spline); (**D**) sigmoid (e.g., logistic, Gompertz, or spline); (**E**) exponential (e.g., exponential or log transform); (**F**) non-monotonic with inflections (e.g., spline or piecewise); (**G**) multimodal, reflecting complex responses or interaction among ECs. Panel (**H**) shows a fictional region with trial locations (dots) within a Target Population of Environments (TPE). Environmental covariates (EC1, ..., $EC_n$) are represented as spatial rasters whose values are extracted at each trial location. These EC layers are then combined to construct enviromic markers (right panels), representing environmental similarity among locations and enabling linear modeling of genotype-by-environment (G × E) responses.

This choice already delivers operational recommendations and consistent macro-scale predictions in wide networks when envirotyping covers the environmental space well (Xu et al. 2022, Bahia et al. 2025). Remote sensing coverage extends this utility by providing continuous, comparable ECs across sites, which anchors predictions to the target more firmly (Resende et al. 2024). As a practical rule, use strategy 1 as the initial reference and move to strategies 2–4 only when residual diagnostics reveal systematic deviations from linearity, thereby constituting patterns which are consistent with the nonlinear responses illustrated in Figure 1B–G.

As local curvature and moderate nonlinear effects appear (Figure 1B to G), expanding the function space with transformations and splines (strategy 2) corrects much of the deviation from linearity. However, the gain comes with more terms and collinearity, which calls for selection and penalization. Feature-engineering pipelines with cross-validation and regularization tend to yield gains while preserving EC interpretability (Montesinos-López et al. 2024). In parallel, ecophysiological indices (strategy 3) use agronomic knowledge to encode stage-specific sensitivity and cumulative stress into fewer, more interpretable markers, often stabilizing G × E patterns and reducing noise in reaction norms. Unsupervised reductions, such as PCA and factor analysis, can stabilize the matrix before fitting strategies 1–3, but they do not use the response and should be treated as auxiliary tools, not substitutes for supervised selection.

For responses with plateaus, multiple thresholds, or multimodal shapes, learning the EC→phenotype mapping with nonlinear models and converting local predictions into synthetic markers (strategy 4) expands predictive reach. Benefits include capturing saturation, combining complex effects, and producing pixel-level maps for adaptation zones (Figure 1H). This approach depends on reliable labels, broad environmental coverage, control of spatiotemporal bias, and substantial computation. Evidence points to advantages for kernels and ensembles when G × E is strong and the environmental space is complex (Costa-Neto et al. 2021, Resende et al. 2025). Strategies 1 and 2, complemented by carefully chosen ecophysiological indices (strategy 3), are often preferable with limited data and resources. In turn, strategy 4 tends to deliver larger gains with a large and diverse set of environmental covariates with adequate spatial and temporal resolution, strong prior knowledge, and a dense spatial mesh, provided that the generated markers undergo supervised selection and validation.

The choice across these four strategies depends on how linear, curved, or fully nonlinear the EC→phenotype relation appears in the data and on the available biological and computational resources. Enviromic markers contribute to both prediction and inference, but the main target in routine breeding is the phenotype to be anticipated for each genotype–environment combination. Following the *prediction-based breeding* view, model adequacy is first judged by out-of-sample accuracy in new sites or seasons, and only afterwards used to interpret environmental drivers (Fritsche-Neto et al. 2025). Raw ECs, ecophysiological indices, and engineered markers can coexist in the same model, including multiple transformations of a given EC, as long as, induced collinearity is controlled through shrinkage, kernels, or dimensionality reduction. Influential ECs or groups of envirotypes can be examined after marker engineering by comparing model scenarios, effect patterns, or contribution profiles, always respecting the transformations applied. Inference remains feasible, but it should be derived from models selected under predictive validation rather than from isolated inspection of regression coefficients.

## QUALITY CONTROL OF ECS AND EEMS

We suggest adopting a SNP-like workflow for environmental covariate (EC) quality control, including: assessing completeness by variable and by pixel, call rate and coverage, standardizing geodetic reference and time window, checking spatiotemporal gradient coherence, identifying ECs that show near-constant values within the Target Population of Environments (TPE) and removing or aggregating them using variance thresholds or cluster-wise diagnostics, and considering collinearity reduction via clustering or dimensionality reduction. When missing data are extensive or the signal-to-noise ratio is low, imputation can rely on neighborhood-assisted interpolation such as inverse-distance weighting (IDW), k-nearest neighbors (k-NN), or spatiotemporal kriging, together with correlated ECs; pixel size can also be increased to stabilize the signal (Yao et al. 2013). We also recommend validating overlays and resolution sensitivity during the environmental data extraction–transformation–loading (ETL) steps (Resende et al. 2024). Sampling plans that expand relevant envirotypic variation tend to strengthen predictive ability and can be prioritized (Resende et al. 2021).

Beyond these procedures, data quality conditions both EC-based modeling and the construction of Engineered Enviromic Markers (EEMs). Inconsistent coordinates, reference mismatches, gaps in time series or noisy layers propagate instability to derived markers and weaken prediction and inference. Recent studies show that open climate, soil and remote-sensing products already provide resolution compatible with breeding trials and recommendation domains (Resende et al. 2025), so enviromic modeling is feasible with public data when data pipelines and quality control are handled explicitly. Under this view, the main constraint is less access to environmental information and more how datasets are curated, aligned and transformed into markers that are evaluated under realistic validation schemes.

If target ECs fall outside the training range, the covariate shift should be characterized. Extrapolation can be measured per EC and in multivariate space using Mahalanobis distance or kernel similarity. Predictions may be retained but uncertainty widened within the same TPE and with mild extrapolation, with shrinkage applied to EC-linked effects and decisions restricted to interpolative ranges. Partial abstention is advisable in cases of strong extrapolation: prioritize or filter stable genotypes, simplify ECs to the observable envelope, or refit models using only valid predictors. ECs should be standardized in the target domain, environmental hyperparameters re-estimated with a small local sentinel set, and domain adaptation implemented via density reweighting across TPEs. When possible, calibration can be expanded with target-domain data or mechanistic or climate scenarios consistent with the phenological window (Cooper and Messina 2021, Rogers and Holland 2022).

Validation schemes should mirror real use and prevent leakage to validate EEMs, defined herein as phenotypes interpolated and then entered linearly in the model (Resende et al. 2025). Spatial and temporal partitions can be used, such as leave-one-region-out (LORO), pixel or municipality blocks, year-by-year splits and, when relevant, partitions across TPEs. Markers must only be built in the training set, with hyperparameters frozen before generating blind predictions in the test set. Predictive gain should be reported against four baselines: (i) no ECs; (ii) raw ECs; (iii) transformed and ecophysiological ECs; and (iv) AI-engineered enviromic markers. Additional diagnostics include uncertainty calibration, prediction-versus-observation curves, resolution and pixel-size stability, and sensitivity to input noise. Interpolation uncertainty should be propagated to the final metrics. Randomization or permutation tests can verify that gains exceed noise and estimated marker effects should be checked for consistency with the assumed linearity and expected G×E heterogeneity. External validation is recommended when independent data exists, together with an environmental confidence index per sample at recommendation time.

## FINAL CONSIDERATIONS

Systematic integration of environmental data into genetic models via enviromic markers is an active front in quantitative genetics. Evidence shows it is feasible to treat raw environmental covariates as linear markers and to run envirotyping at scale (Jarquín et al. 2014, Bahia et al. 2025). Advances in remote sensing and AI indicate that enviromics-enabled selection can anticipate phenotypic trends without costly direct measurements (Montesinos-López et al. 2024, Araújo et al. 2024, Resende et al. 2024, Resende et al. 2025). Taken together, recent work points to connecting the envirome with the genome in a single predictive framework (Costa-Neto et al. 2021, Resende et al. 2024).

Herein we aim to offer operational definitions and a unified guide to apply enviromic markers in plant breeding. Practically, genomic markers (typically biallelic SNPs) encode who each genotype is, while environmental kernels encode where and when genotypes express their potential (Costa-Neto et al. 2021). High-resolution knowledge of the envirome enables pixel-level reasoning about G×E, mapping adaptation zones, and building virtual trials under climate scenarios (Heslot et al. 2014, Resende et al. 2024, Resende et al. 2025). Breeding programs should explicitly delimit their TPE using GIS and regional data (Cruz et al. 2025) and incorporate region-relevant ECs into the predictive pipeline (Costa-Neto et al. 2021, Resende et al. 2022).

The enviromic marker adds a complementary dimension to genomics for climate change and prediction in new environments (Figure 1-H). Leveraging enviromics can accelerate genetic gain and support zone-based selection, provided data are standardized and modeling is integrated (Cooper and Messina 2021, Xu et al. 2022, Rogers and Holland 2022). In practice, this requires training in environmental data analysis and multidisciplinary approaches with remote sensing and AI integrated into the predictive pipeline (Resende et al. 2024, Montesinos-López et al. 2024, Cruz et al. 2025). Enviromic markers will likely become as common in vocabulary as SNPs or QTLs, extending A/T/G/C toward space-time precision breeding.

## DATA AVAILABILITY

The datasets generated and/or analyzed during the current research are available from the corresponding author upon reasonable request.

## REFERENCES

Araújo MS, Chaves SF, Dias LAS, Ferreira FM, Pereira GR, Bezerra AR, Alves RS, Heinemann AB, Breseghello F, Carneiro PCS, Krause MD, Costa-Neto G and Dias KO (2024) GIS-FA: an approach to integrating thematic maps, factor-analytic, and envirotyping for cultivar targeting. **Theoretical and Applied Genetics 137**: 80.

Bahia MAM, Marcatti GE, Breseghello F, Melo PG, Dias KO, Xu Y and Resende RT (2025) Predictive ability of enviromic modeling in G×E interactions for upland rice site recommendations. **Euphytica 221**: 1-29.

Cooper M and Messina CD (2021) Can we harness “enviromics” to accelerate crop improvement by integrating breeding and agronomy? **Frontiers in Plant Science 12**: 735143.

Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, Wright D and Graham G (2014) Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. **Crop and Pasture Science 65***:* 311-336.

Costa-Neto G and Fritsche-Neto R (2021) Enviromics: Bridging different sources of data, building one framework. **Crop Breeding and Applied Biotechnology 21**: e003721210.

Costa-Neto G, Fritsche-Neto R and Crossa J (2021) Nonlinear kernels, dominance, and envirotyping data increase the accuracy of genome-based prediction in multi-environment trials. **Heredity 126**: 92-106.

Cruz DD, Heinemann AB, Marcatti GE and Resende RT (2025) Defining the target population of environments (TPE) for enviromics studies using R-based GIS tools. **Crop Breeding and Applied Biotechnology 25**: e50822519.

Finlay KW and Wilkinson GN (1963) The analysis of adaptation in a plant breeding programme. **Australian Journal of Agricultural Research 14**: 742-754.

Fradgley NS, Gerard GS, Govindan V, Nicol JM, Singh A, Tadesse W, Zwart AB, Trethowan R, Trevaskis B, Whan A and Hyles J (2025) Prediction of Australian wheat genotype by environment interactions and mega-environments. **Theoretical and Applied Genetics 138**: 241.

Fritsche-Neto R, Resende RT, Olivoto T, Garcia-Abadillo J, Nascimento M, Bahia MAM, Jarquín D and Vieira RA (2025) Prediction-based breeding: Modern tools to optimize and reshape programs. **Crop Science 65**: e70175.

Heslot N, Akdemir D, Sorrells ME and Jannink JL (2014) Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. **Theoretical and Applied Genetics 127**: 463-480.

Jarquín D, Crossa J, Lacaze X, Cheyron P, Daucourt J, Lorgeou J, Piraux F, Guerreiro L, Pérez P, Calus M, Burgueño J and Campos G (2014) A reaction norm model for genomic selection using high-dimensional genomic and environmental data. **Theoretical and Applied Genetics 127***:* 595-607.

Marcatti GE, Resende RT, Resende MDV, Ribeiro CAA, Santos AR, Cruz JP and Leite HG (2017) GIS-based approach applied to optimizing recommendations of Eucalyptus genotypes. **Forest Ecology and Management 392**: 144-153.

Montesinos-López OA, Crespo-Herrera L, Pierre CS, Cano-Paez B, Huerta-Prado GI, Mosqueda-González BA, Ramos-Pulido S, Gerard G, Alnowibet K, Fritsche-Neto R, Montesinos-López A and Crossa J (2024) Feature engineering of environmental covariates improves plant genomic-enabled prediction. **Frontiers in Plant Science 15**: 1349569.

Mumford MH, Forknall CR, Rodriguez D, Eyre JX and Kelly AM (2023) Incorporating environmental covariates to explore genotype× environment× management (G× E× M) interactions: A one-stage predictive model. **Field Crops Research 304**: 109133.

Park TC, Chamara P, Lubberstedt T and Scott PM (2025) Beyond the genome: The role of functional markers in contemporary plant breeding. **Frontiers in Plant Science 16**: 1637299.

Resende RT, Chenu K, Rasmussen SK, Heinemann AB and Fritsche-Neto R (2022) Editorial: Enviromics in plant breeding. **Frontiers in Plant Science 13**: 935380.

Resende RT, Hickey L, Amaral CH, Peixoto LL, Marcatti GE and Xu Y (2024) Satellite-enabled enviromics to enhance crop improvement. **Molecular Plant 17**: 848-866.

Resende RT, Piepho HP, Rosa GJM, Silva-Junior OB, Silva FFE, Resende MDV and Grattapaglia D (2021) Enviromics in breeding: Applications and perspectives on envirotypic-assisted selection. **Theoretical and Applied Genetics 134**: 95-112.

Resende RT, Xavier A, Silva PIT, Resende MP, Jarquin D and Marcatti GE (2025) GIS-based G×E modeling of maize hybrids through enviromic markers engineering. **New Phytologist 245**: 102-116.

Rincent R, Malosetti M, Ababaei B, Touzy G, Mini A, Bogard M, Martre P, Le Gouis J and van Eeuwijk F (2019) Using crop growth model stress covariates and AMMI decomposition to better predict genotype-by-environment interactions. **Theoretical and Applied Genetics 132**: 3399-3411.

Rogers AR and Holland JB (2022) Environment-specific genomic prediction ability in maize using environmental covariates depends on environmental similarity to training data. **G3: Genes, Genomes, Genetics 12**: jkab440.

Tolhurst DJ, Gaynor RC, Gardunia B, Hickey JM and Gorjanc G (2022) Genomic selection using random regressions on known and latent environmental covariates. **Theoretical and Applied Genetics 135**: 3393-3415.

Trevisan BA, Junqueira VS, Florêncio BM, Coelho ASG, Marcatti GE and Resende RT (2025) A framework for building enviromics matrices in mixed models. **Brazilian Journal of Biometrics 43**: e-43865.

Xu Y (2016) Envirotyping for deciphering environmental impacts on crop plants. **Theoretical and Applied Genetics 129**: 653-673.

Xu Y, Zhang X, Li H, Zheng H, Zhang J, Olsen MS, Varshney RK, Prasanna BM and Qian Q (2022) Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. **Molecular Plant 15**: 1664-1695.

Yao X, Fu B, Lü Y, Sun F, Wang S and Liu M (2013) Comparison of four spatial interpolation methods for estimating soil moisture in a complex terrain catchment. **PloS ONE 8**: e54660.

Zhang B, Hauvermale AL, Zhang Z, Thompson A, Neely C, Esser A, Pumphrey M, Garland-Campbell K, Yu J, Steber C and Li X (2024) Harnessing enviromics to predict climate-impacted high-profile traits to assist informed decisions in agriculture. **Food and Energy Security 13**: e544.