

Defining the target population of environments (TPE) for enviromics studies using R-based GIS tools

Demila D. M. Cruz¹, Alexandre B. Heinemann², Gustavo E. Marcatti^{3,4} and Rafael T. Resende^{1,4*}

Abstract: We present an R-based function for defining TPE as GIS-polygons, intended for use in enviromics studies. It offers customizable parameters, such as pixel size, buffer boundaries, and concavity, providing enhanced flexibility for G×E analysis. This tool optimizes genotypic, envirotypic, and spatial assessments, serving as a powerful resource for breeding research.

Keywords: 'TPEmap' function, precision breeding, envirotyping, G×E interaction, concave/convex hull, spatial analysis

INTRODUCTION

The concept of the target population of environments (TPE) has evolved significantly since its initial definition as the set of environments where genotypes are recommended based on performance in representative test environments (Comstock 1977, Allen et al. 1978). Chenu et al. (2011) stressed the importance of TPE in addressing genotype-by-environment (G×E) interactions, particularly in the context of water-deficit variability. Chapman et al. (2012) further broadened its scope by incorporating biotic and abiotic stresses, thus highlighting its role in improving crop adaptation to climate change. With advances in breeding technologies, Voss-Fels et al. (2018) integrated TPE into genomic selection, stressing the need for accurate environmental characterization. Crespo-Herrera et al. (2021) included socioeconomic factors in TPE considerations, while Cooper and Messina (2021) explored its application in multi-environment trials to improve management strategies and genotype predictions. Resende et al. (2021) applied TPE to enviromic-assisted selection. They used GIS tools for a detailed dissection of G×E interactions.

The concepts of TPE, mega-environments, and breeding zones are often conflated, yet they represent different scales of genotype adaptation. Mega-environments are broad regions characterized by relatively homogeneous environmental conditions, typically identified through environmental data analyses (Gauch and Zobel 1997, Gupta et al. 2013, Yan et al. 2022). In contrast, breeding zones are smaller sub-regions within mega-environments or TPEs, defined by specific environmental traits to reduce G×E interactions (Resende et al. 2021, Callister et al. 2024, Resende et al. 2024a). The integration of big data and advanced technologies has significantly improved the efficiency of

Crop Breeding and Applied Biotechnology
25(1): e50822519, 2025
Brazilian Society of Plant Breeding.
Printed in Brazil
<http://dx.doi.org/10.1590/1984-70332025v25n1s09>



*Corresponding author:

E-mail: rafael.tassinari@ufg.br

 ORCID: 0000-0002-7139-708X

Received: 12 September 2024

Accepted: 30 October 2024

Published: 10 January 2025

¹ Universidade Federal de Goiás, Avenida Esperança, s/n, Chácaras de Recreio Samambaia, 74690-900, Goiânia, GO, Brazil

² Embrapa - Arroz e Feijão, Rodovia GO-462, km 12, Zona Rural, 75375-000, Santo Antônio de Goiás, GO, Brazil

³ Universidade Federal de São João del-Rei, Rodovia MG-424, km 47, 35701-970, Sete Lagoas, MG, Brazil

⁴ TheCROP - Enviromics | Phenomics | Genomics. Avenida Esperança, 1533, Parque Tecnológico, FUNAPE, Campus Samambaia/ UFG, 74690-612, Goiânia, GO, Brazil

breeding programs by enabling the precise targeting of cultivars to specific environments, thereby increasing genetic gains (Costa-Neto and Fritsche-Neto 2021, Xu et al. 2022).

Clipping the TPE within a GIS framework involves integrating spatial data with environmental factors such as climate and soil characteristics, both of which have a significant impact on crop performance (Resende et al. 2024a). GIS tools play a critical role in mapping these factors and accounting for geographic variability. The complexity increases when this spatial data is combined with trials conducted in different environments to improve genotype performance and account for genotype-environment-management (G×E×M) interactions (Cooper and Messina 2021). The TPE may include continuous or discontinuous regions where environmental factors have unique effects on genotypes. Tools such as surrounding polygons derived from multi-environment trials using concave or convex hulls (Gombin et al. 2020) are instrumental in delineating the TPE. However, Olivoto and Lucio (2020) emphasize the challenges of conducting a comprehensive multi-environment trial (MET) analysis. These challenges could be mitigated through a GIS-based computational tool integrating phenotypic and geospatial data.

This approach streamlines the identification of environmental patterns within target areas, thereby ensuring that TPE accurately reflects real-world crop conditions (Resende et al. 2024b). The use of GIS makes it possible to cross-reference climatic, edaphic, and management data (Pebesma and Bivand 2023), providing an integrated perspective on genotype performance. The integration of GIS has been particularly effective in enviromics, where satellite and remote sensing data provide deeper insights into G×E interactions, enhancing strategies for crop improvement (Resende et al. 2024a). This paper introduces the R function ‘**TPEmap**’, designed to define precise TPEs by incorporating buffers and concavity adjustments. We demonstrate its application using a common bean dataset from the Brazilian Agricultural Research Corporation (Embrapa) Arroz & Feijão Breeding Program, as detailed by Heinemann et al. (2022).

METHODS

Herein we describe a function called ‘**TPEmap**’, designed to define geographic TPE polygons. This function delineates the entire area where experimental or on-farm trial data points serve as reference locations. The underlying premise is that any location where a crop is being bred, tested, or cultivated forms part of the target environment, while surrounding areas are included to account for environmental variability.

Geospatial data processing and analysis

Geospatial data processing began by loading the geographic coordinates of the MET and on-farm trial points. These shapefiles were processed using the ‘**sf**’ library in R (Pebesma and Bivand 2023). To ensure spatial compatibility with other datasets, the coordinates were transformed into the WGS 84 coordinate reference system (CRS; EPSG:4326). To facilitate the testing of different scenarios, we developed a ‘**generate_coordinates**’ function. This function allows for the simulation of any number of points within a specified area. While Brazil was used as the reference area in this study, the function can be applied to any location worldwide.

After processing and preparing the geospatial data, the next step was to apply buffers to the geographic coordinates (as shown in the Trial Buffers in Figure 1). These buffers represent phenotypic data collection points, such as breeding experiments or on-farm trial locations. By using the ‘**st_buffer**’ function from the ‘**sf**’ package, buffers were created at adjustable scales, measured in kilometers, around each data point. These individual buffer polygons were then merged using the ‘**st_union**’ function to form a single aggregated area of influence.

Building the TPE polygon using a concavity algorithm

The TPE polygon was constructed by using the ‘**concaveman**’ algorithm in R (Gombin et al. 2020), which generates a concave polygon adaptable to the geospatial points of interest. Following the creation of buffers around the trial points, these polygons were merged and converted into a set of points, forming the foundation for the concave polygon. The ‘**TPEmap**’ function provides flexibility for users by allowing adjustments to parameters such as buffer distance, concavity, and length threshold, enabling customization of the polygon to meet the specific needs of a study. The ‘**generate_coordinates**’ and ‘**TPEmap**’ functions, along with detailed instructions on how to use them, are available on GitHub [<https://github.com/Enviromics/TPE-mapping>].

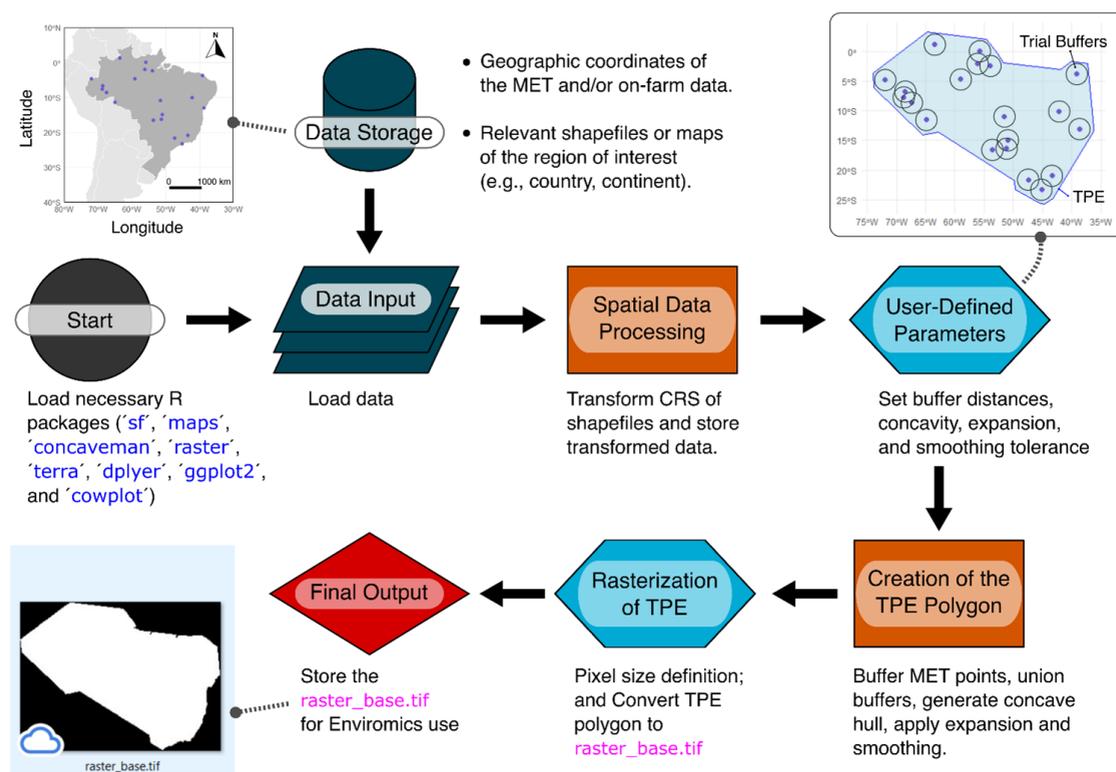


Figure 1. Geographical information system (GIS) flowchart for target population of environments (TPE) construction by using the 'TPEmap' function. The process includes: 1) loading and transformation of geospatial data; 2) definition of user-defined parameters (e.g., buffers, concavity, and smoothing); 3) creation of TPE polygon; and 4) rasterization. The output is the 'raster_base.tif' file, which serves as a foundational layer for enviromics analyses. This file integrates data from multi-environmental trials (MET) or on-farm trials with environmental information.

An optional buffer can be applied during TPE polygon generation, allowing the user to expand the final area to include all desired influence zones. The 'st_simplify' function can be used to smooth out irregularities and small protrusions to improve the usability of the polygon, with the smoothing tolerance adjustable to the user's preference. Breeders may choose to define an external TPE without incorporating phenotypic trial data, with the goal of obtaining predictive results. However, such an approach cannot be validated by using appropriate cross-validation schemes (Resende et al. 2024b). An alternative validation strategy involves planting in the target region in the next cycle or season to confirm the model's predictions. For example, if data points from previous trials define the region, the model can be trained and validated using data from the same region (see Rogers and Holland 2022 for possible validation schemes). If the goal is to predict outcomes for a different region, such as another country, similar TPEs may be identified, but the predictions cannot be validated without new trials. This would necessitate additional cycles to generate and confirm the required validation data.

The final polygon is converted into a raster format, with the pixel size determined based on the specifications of the G×E study. The resulting file ('raster_base.tif') is prepared for use in Enviromics analyses, allowing for the integration of environmental data with the geographical locations of the trials. This ensures that the TPE accurately represents the environmental conditions of the regions of interest. For examples of potential enviromics results achievable using the base raster, refer to Resende et al. (2024a). Increase in pixel resolution (i.e., decrease in pixel granularity) can enhance spatial detail. However, this comes at the cost of a quadratic increase in the total number of pixels within the TPE, which requires more physical computational memory.

THE 'TPEMAP' FUNCTION ARGUMENTS

The **'TPEmap'** function was developed to facilitate the creation of a TPE for plant breeding, leveraging geospatial data from MET or on-farm trials. This function includes several user-defined parameters (Figure 1, "User-Defined Parameters"). Key arguments, detailed below, include adjustments for buffers, polygon concavity, and pixel size for rasterization:

- **'coordinates'**: A data frame containing the geographic coordinates of MET or on-farm trial points. The data frame must include two columns, **x** (longitude) and **y** (latitude), representing the location of each trial point.
- **'point_buffer'**: A numeric value specifying the buffer distance (in kilometers) to be applied around each point. This argument allows the user to adjust the TPE to include nearby regions of interest.
- **'concavity'**: A numeric value that controls the degree of polygon concavity. Lower values create a more detailed polygon, while higher values (approaching infinity) result in a convex hull.
- **'length_threshold'**: A numeric value that sets the edge length threshold in the concavity algorithm. Edge segments shorter than this threshold are excluded from the polygon's extra detail. This parameter helps balance the level of detail with computational efficiency.
- **'expansion_buffer'**: A numeric value defining an additional buffer distance (in kilometers) applied after generating the initial concave polygon, expanding the final TPE.
- **'simplify_tolerance'**: A numeric value that sets the tolerance level for simplifying the final polygon. This parameter allows users to smooth the polygon by removing excessive detail and minor irregularities.
- **'pixel_size'**: A numeric value that specifies the pixel size for rasterizing the TPE. This parameter is crucial during the conversion of the final polygon into a raster format.

FEATURES OF THE 'TPEMAP' FUNCTION

To demonstrate the functionality of the **'TPEmap'** function with a realistic example, we adapted the **'generate_coordinates'** function to simulate geographic coordinates for 100 points. These points represent locations such as experiments or plantings, ensuring their suitability for soybean cultivation (Silva et al. 2021). Protected and natural reserve areas were excluded from the simulations to ensure compliance with environmental regulations. Figure 2 showcases various configurations possible with **'TPEmap'**. Users can customize key parameters, including 1) buffer sizes, 2) concavity (concave hull), 3) length threshold (**'length_threshold'**), and 4) simplification tolerance (**'simplify'**).

Concavity and length threshold parameters are crucial for fine-tuning the TPE clipping process, allowing users to create polygons with varying levels of detail. These parameters enable the generation of polygons ranging from highly detailed, intricate shapes (e.g., Figure 2A and 2C) to simpler, more generalized, and convex shapes (e.g., Figure 2D). The **'TPEmap'** function also facilitates the visualization of different configurations by enabling adjustments to concavity values, such as varying it from 1 to 5. For cases requiring a convex hull, the concavity can be set to very high values (e.g., 1e5 in this context), providing a straightforward way to encapsulate all trial points within a broad, enclosing polygon.

APPLICATION TO COMMON BEAN MET DATA SET

The success of breeding strategies relies heavily on accurate environmental characterization and the alignment of MET data with the environments within the TPE (Cooper et al. 2023). In defining the TPE for common bean (*Phaseolus vulgaris* L.), we utilized MET coordinates from Embrapa Arroz & Feijão, encompassing three distinct crop season types: Rainfed (Figure 3A), Dryland (Figure 3B), and Winter (Figure 3C) (Heinemann et al. 2022).

This dataset includes 423 trials conducted across 71 unique geographic locations from 2011 to 2023. Depending on the regionalization of production areas, trials covered up to three different crop seasons (Figure 3D). Specifically, 26 trials were conducted in the Rainfed season; 12 in rotations of Dryland and Rainfed; 9 in Rainfed and Winter; 9 exclusively in Winter; 7 exclusively in Dryland; 5 in Dryland and Winter; and 3 locations encompassed all three seasons. Trials included 87 commercial cultivars, with 52 classified as "Carioca" type and 35 as "Preto" type.

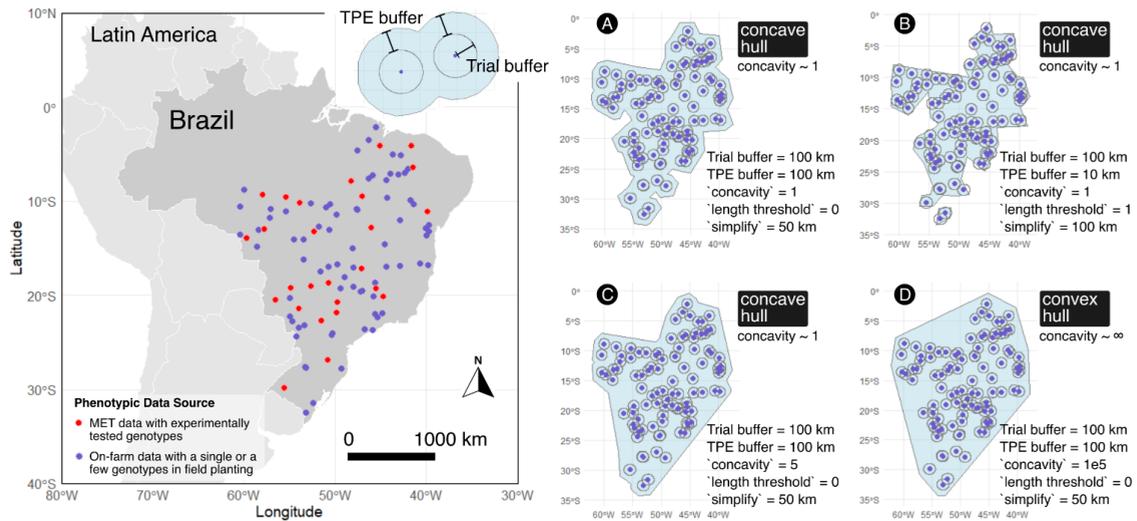


Figure 2. Definition of the target population of environments (TPE) polygon using different concavity and buffer parameters. (A) TPE polygon generated with a concavity of ~ 1 , a trial buffer of 100 km, a TPE buffer of 100 km, and a simplification tolerance of 50 km. This configuration includes 100 geographic points, 26 of which are breeding experiments (shown in red). (B) TPE polygon with a concavity of ~ 1 , a trial buffer of 100 km, a TPE buffer of 10 km, and a simplification tolerance of 100 km. (C) TPE polygon generated with a concavity of ~ 1 , a trial buffer of 100 km, a TPE buffer of 100 km, but with a concavity adjusted to 5. (D) TPE polygon produced as a convex hull (concavity $\sim \infty$), using a trial buffer of 100 km and a TPE buffer of 100 km. These configurations demonstrate how changes in concavity and buffer parameters affect the shape and fit of the polygon, highlighting areas of interest for enviromics analysis.

Figure 3E demonstrates the process of generating the surrounding polygons (concave hull) for the TPE using a concavity value of 2 and a 'length_threshold' of 10. Trial buffers ('point_buffer') of 100 km were applied around each data point. An additional 100 km buffer was then added to the TPE polygon for a total expansion of 200 km. The final polygon was then smoothed with a simplification tolerance of 50 km to create a more refined shape. This methodology is essential for identifying and mapping representative production areas that consider the different environmental conditions that common beans face in Brazil due to the country's bioclimatic diversity (Elias et al. 2021).

The success of a breeding program often hinges on achieving a balance between broad adaptation and targeted regional adaptation (Piepho and Möhring 2005). The different responses of common beans to climatic conditions during different growing seasons must be accounted for to ensure that the TPE reflects actual growing conditions in Brazil (Heinemann et al. 2022). The results reveal a wide distribution of experimental points across Brazil, encompassing diverse climatic zones and soil types. Using the 'TPEmap' function, we constructed a TPE polygon that integrates data from all three crop seasons, offering a comprehensive representation of common bean cultivation areas. The overlay of trials (Figure 3D) shows that incorporating data from multiple seasons is a practical approach to defining TPE that captures broad environmental variability. However, TPEs can be designed with different focuses: 1) broad, spanning multiple seasons; 2) season-specific, tailored to individual growing seasons; or 3) based on other stratification strategies used by breeding programs.

Considering modern envirotyping techniques, as described by Xu et al. (2022) and Resende et al. (2024a), the rasterized base with refined pixels can be used to extract environmental covariates essential for Enviromics studies and G×E interaction analyses at various spatial scales (Resende et al. 2021). Envirotyping data can be sourced from platforms such as WorldClim, Planet, NASAPower, ERA5, SRTM, MODIS, and SoilGrids (for details, see Resende et al. 2024a). Additionally, the 'EnvRtype' package facilitates access to NASAPower data (Costa-Neto et al. 2021). The resolution of the raster base ('raster_base') must match the resolution of the data sources, as the base may have a high level of refinement, while some platforms provide coarser pixel resolutions. For example, when the TPE was converted to a raster base (Figure 3F), a pixel size of 0.1° (~ 11.1 km) was used, resulting in 33,465 pixels within the TPE. However, refining the pixel size to 0.01° dramatically increases the number of pixels to 3,347,285, which presents significant computational challenges when downloading and processing environmental data.

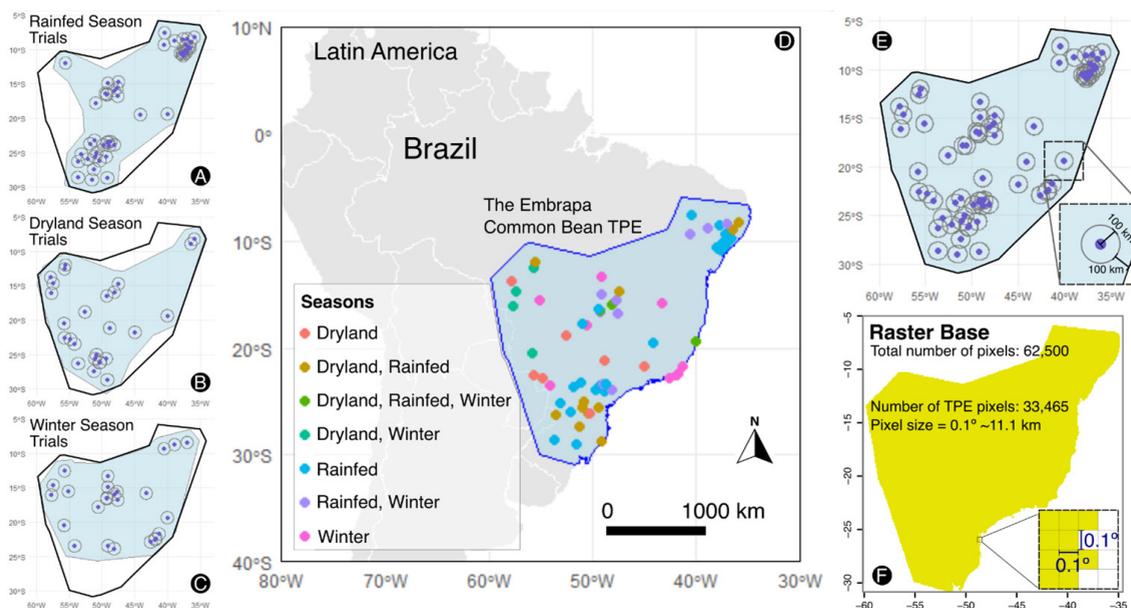


Figure 3. Definition of the target population of environments (TPE) for common beans across different trial types (seasons). (A) TPE polygon generated for rainfed season trials, applying the specified trial buffer. (B) TPE polygon generated for dryland season trials. (C) TPE polygon generated for winter season trials. (D) Final TPE polygon integrating all common bean trials across the three seasons. Different colors indicate the distribution of trials by season: dryland, rainfed, and winter. (E) Visualization of the combined trial buffers, representing all three seasons, along with the resulting TPE polygon. The buffer distances applied around trial points are highlighted. (F) Rasterized TPE base created with a pixel size of 0.1° (~ 11.1 km). The total number of pixels and the pixels contained within the TPE are displayed.

Detailed characterization of environmental types within the TPE is critical for directing breeding programs toward specific targets for adaptation to environmental stresses (Cooper et al. 2023). Results from Resende et al. (2024a) highlight the importance of environmental variables such as soil properties, radiation, and temperature in predicting genotypic performance. The resolution of pixels used to define the TPE plays a pivotal role in the quality of its representation. Smaller pixels provide greater spatial detail but achieving a balance between precision and computational efficiency is essential (Piepho 2022). This study highlights the importance of using GIS tools for the precise definition of TPEs. Such tools enable breeders to account for geographic variation and optimize G×E interaction analyses in envirotyping studies. By doing so, they enhance the adaptation and performance of crop varieties across diverse environmental conditions.

FINAL REMARKS

This article introduces a practical methodology for defining TPE using GIS-based tools and the ‘**TPEmap**’ R function. By incorporating spatial data, particularly the geographic coordinates of phenotypic data sources such as breeding trials and on-farm data, the method allows for the precise delimitation of TPE polygons. The core principle underlying this approach is that areas hosting breeding trials or planting sites, along with their surrounding regions, hold direct relevance for breeders. By focusing on these areas, this methodology optimizes breeding program efficiency, specifically targeting G×E interactions to achieve breeding goals. Through customizable parameters, including buffer distances, concavity, and other spatial adjustments, the ‘**TPEmap**’ adapts models to meet the specific needs of trials. This was effectively demonstrated using Embrapa’s common bean dataset. The tool improves the accuracy of mapping, facilitates informed decision-making in breeding programs, and generates raster to facilitate future data integration.

The future of TPE mapping has significant potential, including expanding its application to regions beyond the Americas and incorporating additional land-use data, such as agricultural areas outside protected zones and urban regions, with the ability to update in real time. Developing methods to subdivide TPEs into sub-TPEs, or breeding zones, could enhance

precision in targeting specific breeding goals. Current efforts aim to refine tools for determining the optimal pixel size and to enable efficient large-scale downloads of envirotypic data. Applying this methodology across a variety of crops and breeding programs will help validate its robustness and adaptability in diverse agricultural systems. Extending the accessibility of 'TPEmap' to Python, in addition to the existing R implementation, could broaden its adoption among breeders and researchers.

ACKNOWLEDGMENTS

The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for granting a master's scholarship to the first author (DDMC). We also thank the Programa de Pós-Graduação em Genética e Melhoramento de Plantas (PPGGMP) at the Universidade Federal de Goiás (UFG) for institutional support. Special thanks to Embrapa Arroz e Feijão for providing the data used in this study and to the Laboratório de Melhoramento de Precisão (LAMP) at the Escola de Agronomia of UFG for supporting the analyses and discussions.

DATA AVAILABILITY

The Embrapa geo coordinates datasets analyzed during the current research are available from the corresponding author upon reasonable request. The simulated datasets are available in: [<https://github.com/Enviromics/TPE-mapping>]¹⁹.

REFERENCES

- Allen FL, Comstock RE and Rasmusson DC (1978) Optimal environments for yield testing. *Agronomy Journal* **70**: 561-567.
- Callister AN, Costa-Neto G, Bradshaw BP, Elms S, Crossa J and Brawnner JT (2024) Enviromic prediction enables the characterization and mapping of *Eucalyptus globulus* Labill breeding zones. *Tree Genetics & Genomes* **20**: 3.
- Chapman SC, Chakraborty S, Dreccer MF and Howden SM (2012) Plant adaptation to climate change - opportunities and priorities in breeding. *Crop and Pasture Science* **63**: 251-268.
- Chenu K, Cooper M, Hammer GL, Mathews KL, Dreccer MF and Chapman SC (2011) Environment characterization as an aid to wheat improvement: Interpreting genotype-environment interactions by modelling water-deficit patterns in North-Eastern Australia. *Journal of Experimental Botany* **62**: 1743-1755.
- Comstock RE (1977) Quantitative genetics and the design of breeding programs. In Pollak E, Kempthorne O and Bailey TB (eds) *Proceedings of the international conference on quantitative genetics*. Iowa State University, Ames, p. 705-718.
- 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, Powell O, Gho C, Tang T and Messina C (2023) Extending the breeder's equation to take aim at the target population of environments. *Frontiers in Plant Science* **14**: 1129591.
- Costa-Neto G and Fritsche-Neto R (2021) Enviromics: bridging different sources of data, building one framework. *Crop Breeding and Applied Biotechnology* **21**: e393521S12.
- Costa-Neto G, Galli G, Carvalho HF, Crossa J and Fritsche-Neto R (2021) EnvRtype: a software to interplay enviromics and quantitative genomics in agriculture. *G3* **11**: jkab040.
- Crespo-Herrera LA, Crossa J, Huerta-Espino J, Mondal S, Velu G, Juliana P and Singh RP (2021) Target population of environments for wheat breeding in India: definition, prediction and genetic gains. *Frontiers in Plant Science* **12**: 638520.
- Elias JCF, Gonçalves-Vidigal MC, Ariani A, Valentini G, Martiniano-Souza MDC, Vaz Bisneta M and Gepts P (2021) Genome-environment association analysis for bio-climatic variables in common bean (*Phaseolus vulgaris* L.) from Brazil. *Plants* **10**: 1572.
- Gauch HG and Zobel RW (1997) Identifying mega-environments and targeting genotypes. *Crop Science* **37**: 311-326.
- Gombin J, Vaidyanathan R and Agafonkin V (2020) concaveman: A very fast 2D concave hull algorithm. R package version 1.1.0. Available at <<https://CRAN.R-project.org/package=concaveman>>. Accessed on September 21, 2024.
- Gupta S, Rathore A, Yadav OP, Rai KN, Khairwal IS, Rajpurohit BS and Mahala RS (2013) Identifying mega-environments and essential test locations for pearl millet cultivar selection in India. *Crop Science* **53**: 607-619.
- Heinemann AB, Costa-Neto G, Fritsche-Neto R, da Matta DH and Fernandes IK (2022) Enviromic prediction is useful to define the limits of climate adaptation: a case study of common bean in Brazil. *Field Crops Research* **286**: 108628.
- Olivoto T and Lúcio AD (2020) Metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution* **11**: 783-789.
- Pebesma E and Bivand R (2023) *Spatial data science: With applications in R*. Chapman and Hall/CRC, New York, 324p.
- Piepho HP (2022) Prediction of and for new environments: What's your model? *Molecular Plant* **15**: 581-582.
- Piepho HP and Möhring J (2005) Best linear unbiased prediction of cultivar effects for subdivided target regions. *Crop Science* **45**: 1151-1159.

- Resende RT, Hickey L, Amaral CH, Peixoto LL, Marcatti GE and Xu Y (2024a) Satellite-enabled enviromics to enhance crop improvement. **Molecular Plant** **17**: 848-866.
- Resende RT, Piepho HP, Rosa GJ, Silva-Junior OB, Silva FF, de 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 (2024b) GIS-based G×E modeling of maize hybrids through enviromic markers engineering. **New Phytologist** **245**: 102-116.
- 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.
- Silva EHFd, Antolin LAS, Zanon AJ, Junior ASA, de Souza HA, dos Santos Carvalho K and Marin FR (2021) Impact assessment of soybean yield and water productivity in Brazil due to climate change. **European Journal of Agronomy** **129**: 126329.
- Voss-Fels KP, Cooper M and Hayes BJ (2018) Accelerating crop genetic gains with genomic selection. **Theoretical and Applied Genetics** **132**: 669-686.
- Xu Y, Zhang X, Li H, Zheng H, Zhang J, Olsen MS and Qian Q (2022) Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. **Molecular Plant** **15**: 1664-1695.
- Yan W, Nilsen KT and Beattie AD (2022) Mega-environment analysis and breeding for specific adaptation. **Crop Science** **62**: 59-70.