Using Explainable Machine Learning techniques to unpack farm-level management x climate interactions

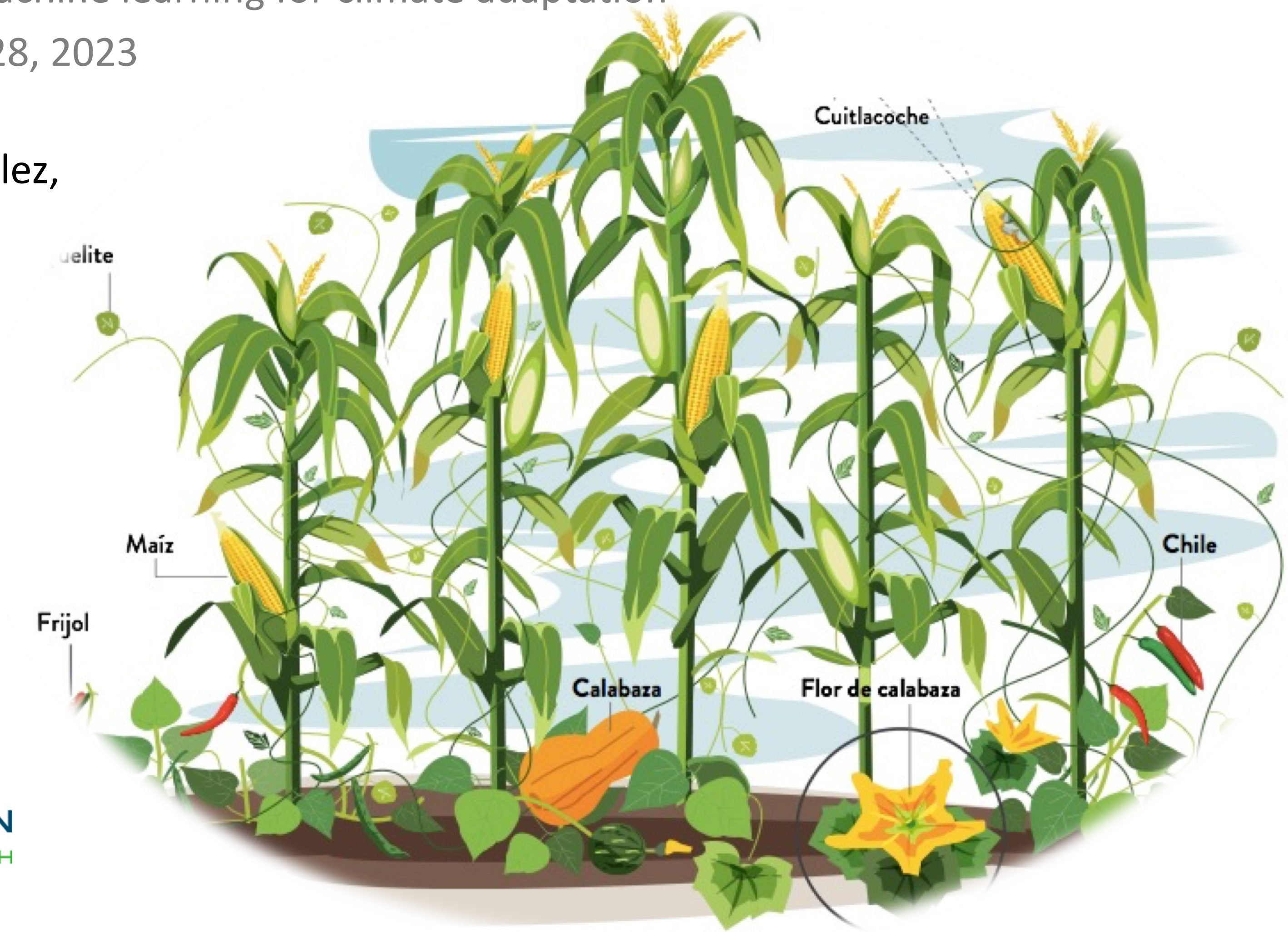
From impact to solutions, data, data science and machine learning for climate adaptation

Wageningen University & Research, November 26-28, 2023

Julian Ramirez-Villegas, Diana Jaimes, Arturo Gonzalez, Lizeth Llanos, Daniel Jimenez, Andrea Gardeazábal, Oscar Estrada, Daniel Nuñez

j.r.villegas@cgiar.org

@JulianCIAT



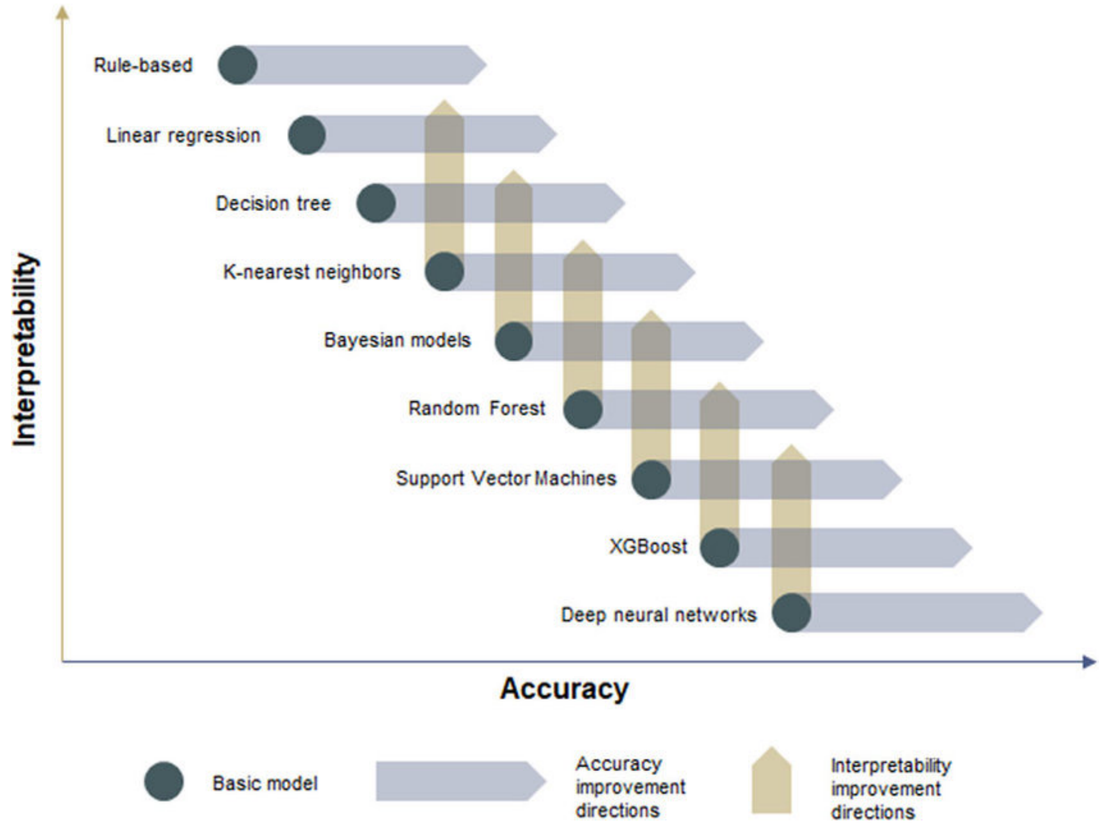
Alliance

Bioversity International

CIAT International Center for Tropical Agriculture Since 1967 Science to cultivate change

WAGENINGEN UNIVERSITY & RESEARCH

ML algoritm categories



Nesvijevskaia, Anna & Ouillade, Sophie & Guilmin, Pauline & Zucker, Jean-daniel. (2021). The accuracy versus interpretability trade-off in fraud detection model. Data & Policy. 3. 10.1017/dap.2021.3.

Black Box Algorithm

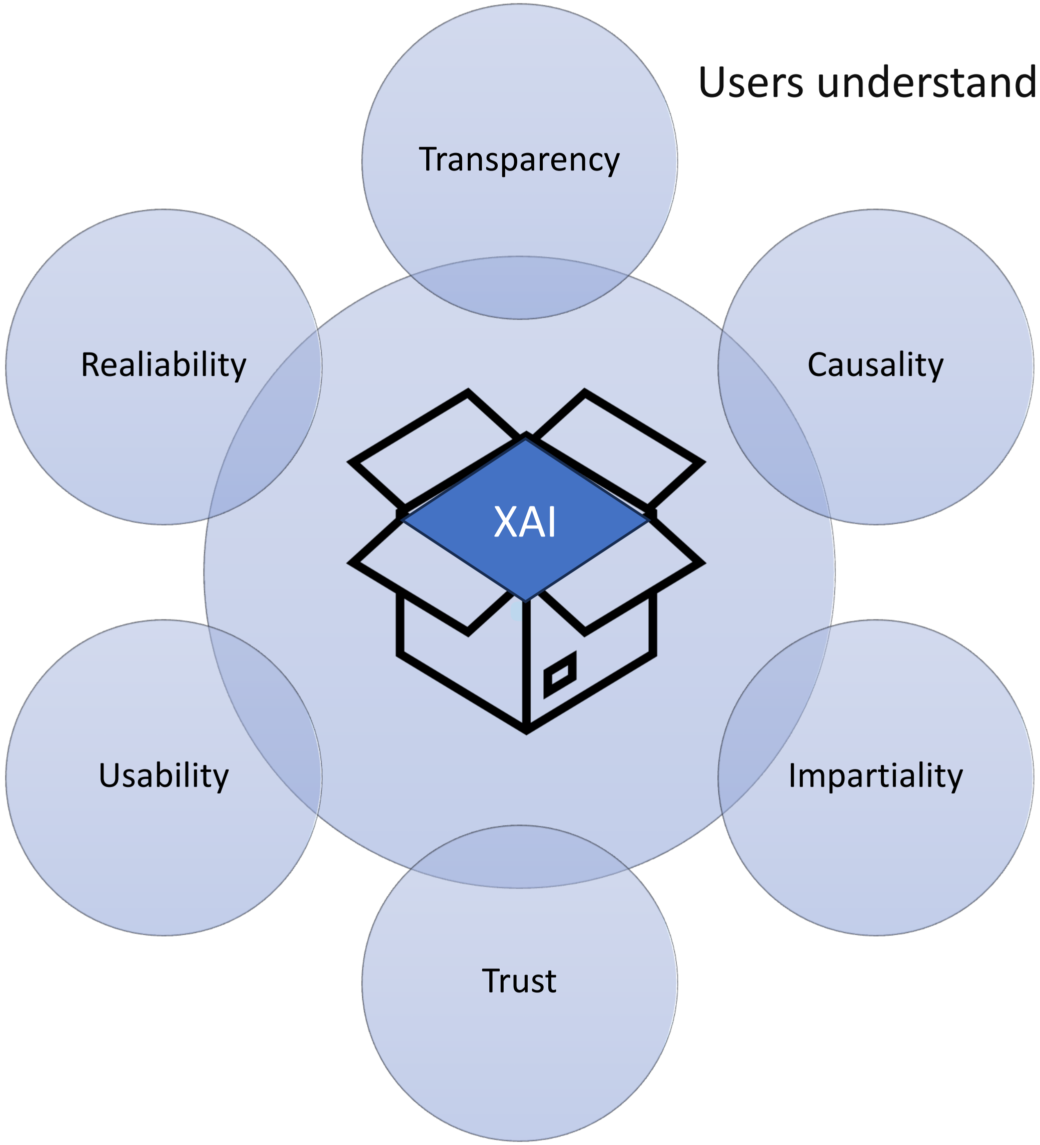
- Hard to interpret
- Needs a large volume of data
- Usually more powerful and effective than White box models

White Box Algorithm

- High interpretability and explainability.
- Lower performance in some applications
- Require less data for training

Algorithm Transparency

Why use XAI methods?



Users understand the decisions made by the model

It can determine how robust the model is when new variables are included, or existing ones are changed.

"Cause-and-effect" relationships can be established, as the change in the model's prediction can be measured when certain input variables are perturbed.

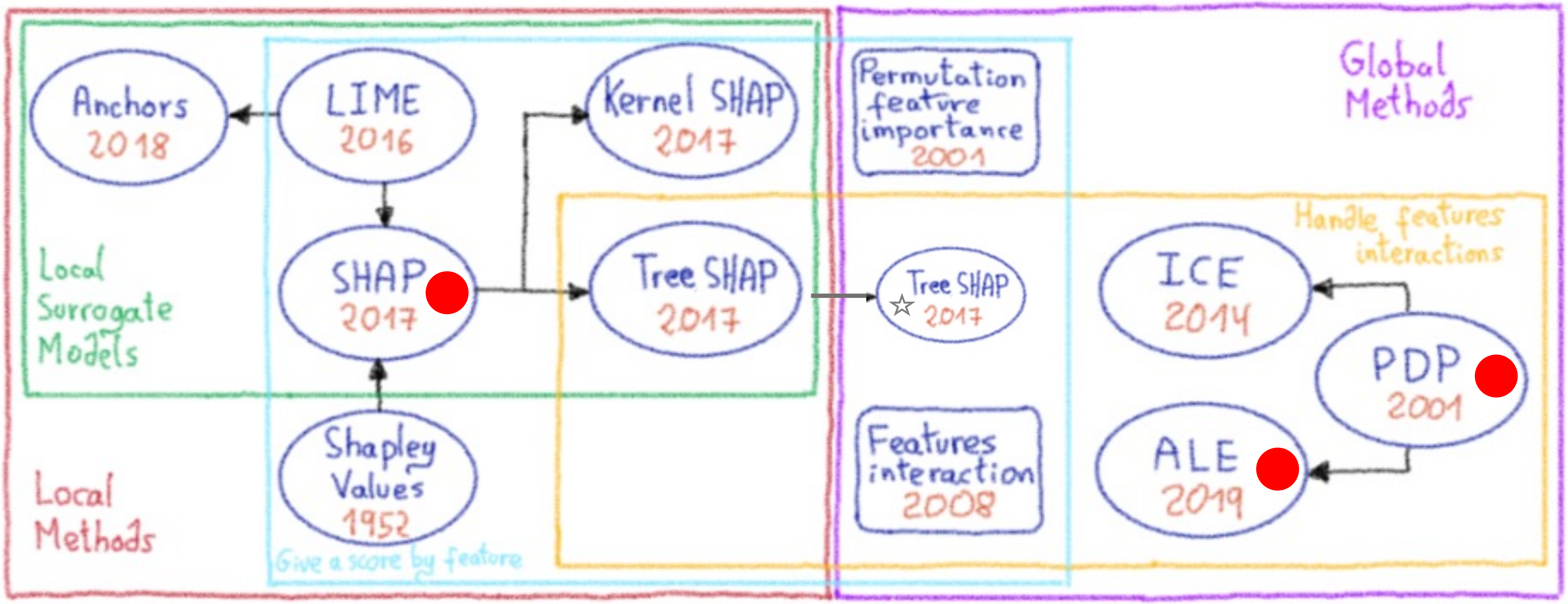
Unveiling how the model operates enhances user confidence in the results

The user, analyzing the model's decisions, can identify biases in the decisions early on

Unveiling how the model operates enhances user confidence in the results

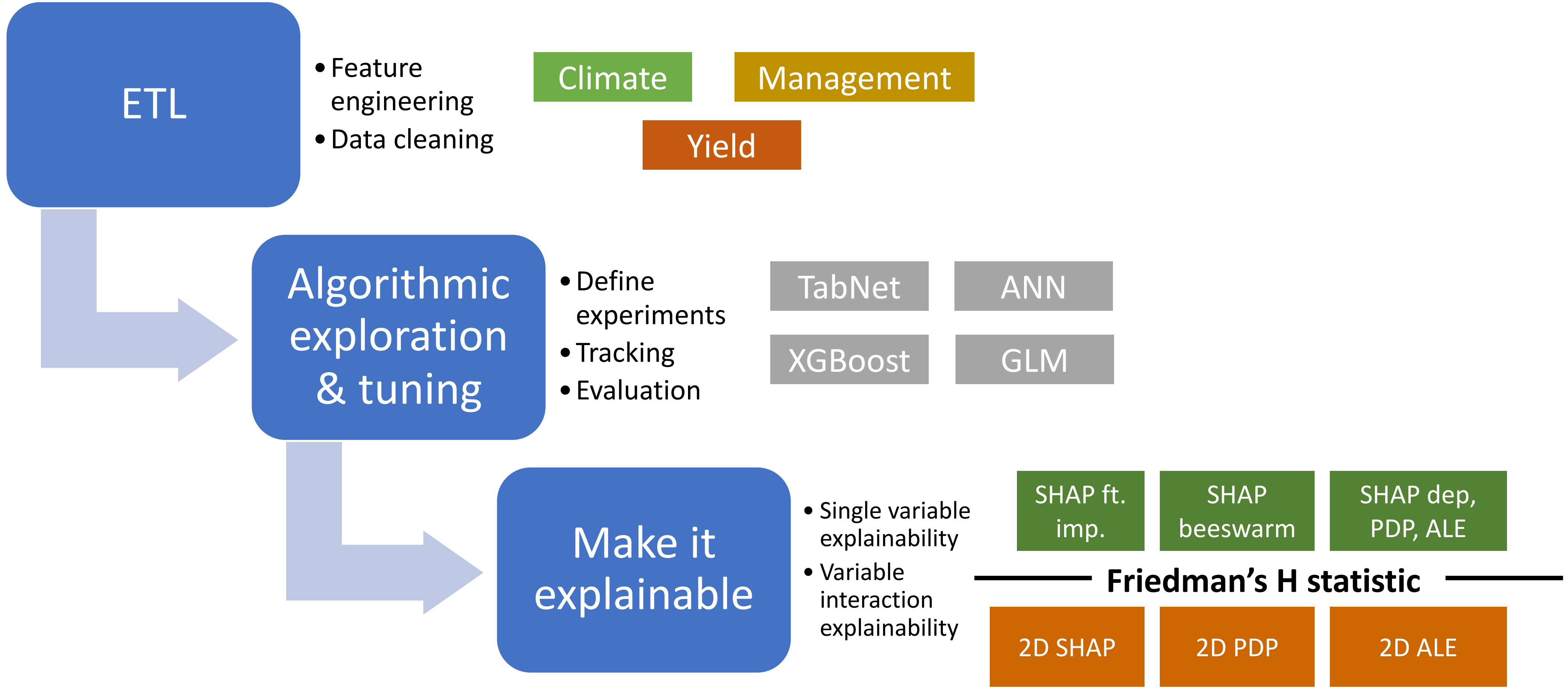
Approaches to interpret a model:

- the local level, the aim is to explain why a specific instance of the model has had a particular prediction.
- the global level, the goal is to understand the drivers of predictions across all instances.



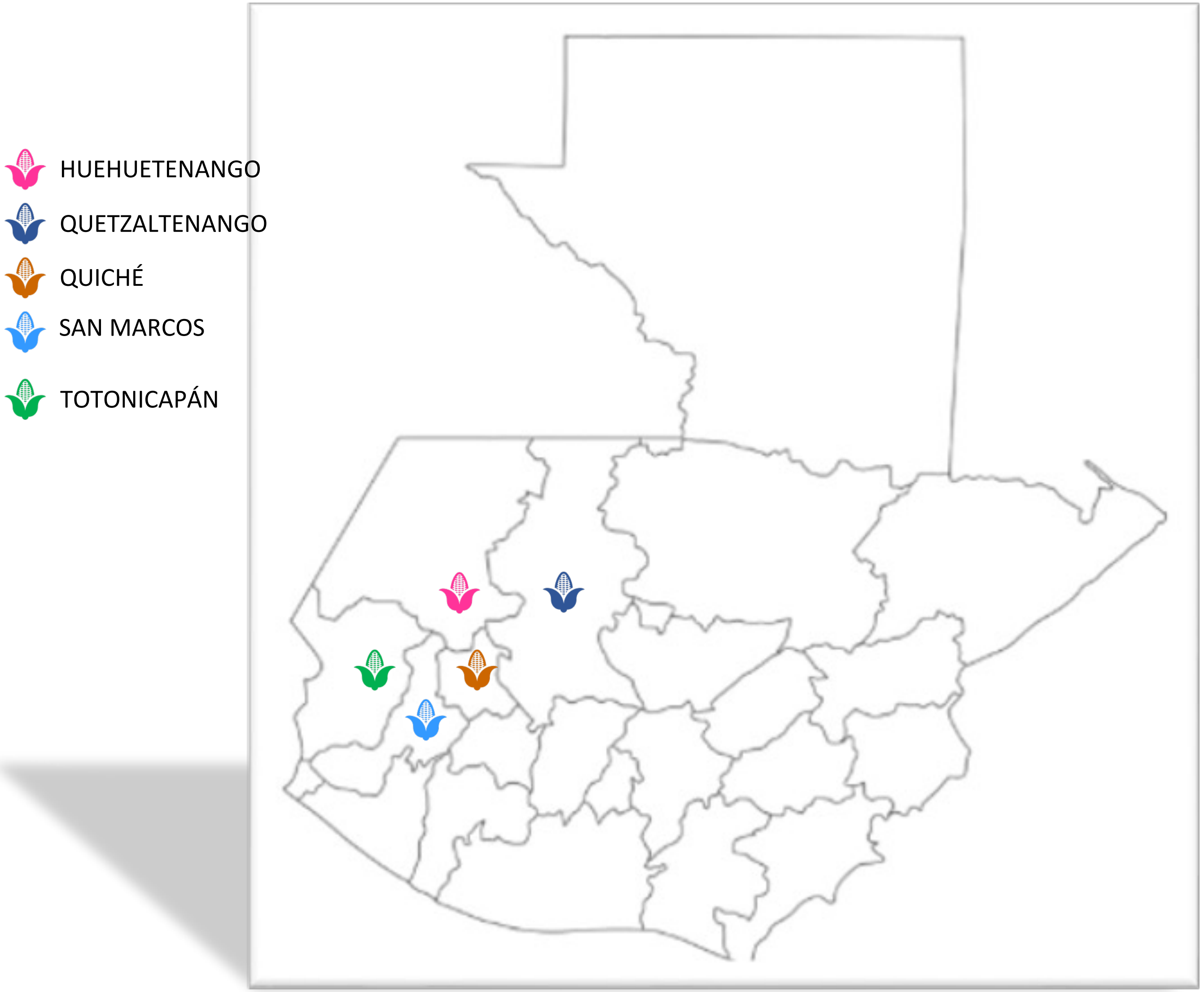
https://blog.ovhcloud.com/interpretability-engine-an-open-source-tool-to-interpret-your-models-in-ml-serving/

Explainable Machine Learning workflow



*SHapley Additive exPlanations

Use case Project:



Log books

During the years 2016-2019, information was collected:

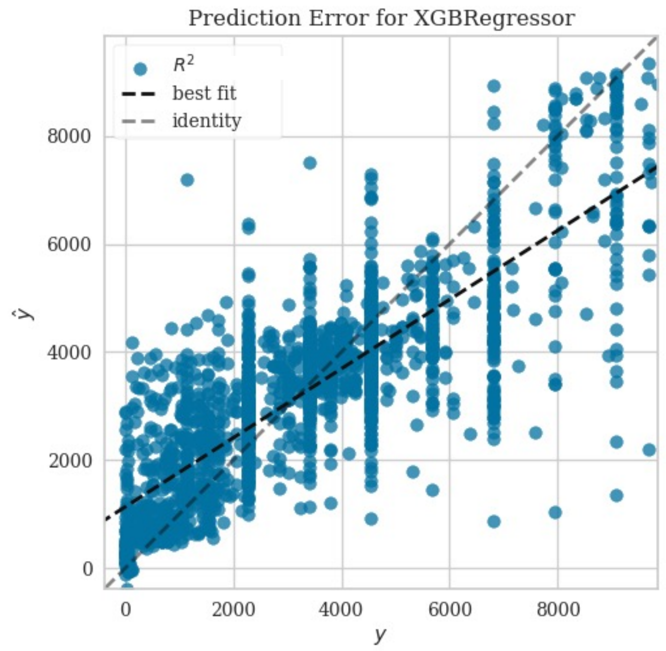
- *Sociodemographic*
- *Crop management such as: Planting dates, agronomic cultivation management, forms of cultivation management, size, class, etc.*

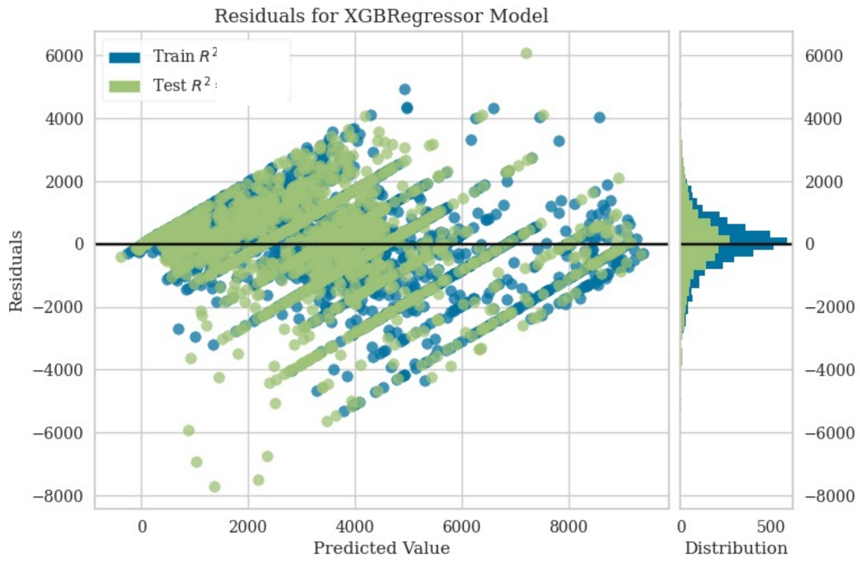
Provided information

*Bitácora escalamiento
*Fitomejoramiento
*Zoomejoramiento
*Área de impacto
*Formato visita

About the model

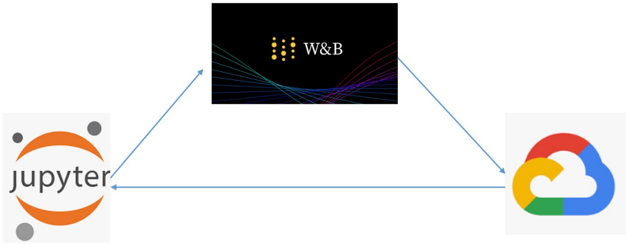
- Gradient Boosting Model (xgboost) with categorical variables
- R-squared as evaluation metric.
- Bayesian hyperparameters search
- Explicit use of methods to reduce overfitting
- We ”search” for the ‘fittest’ model based on performance (R-squared) and explainability.





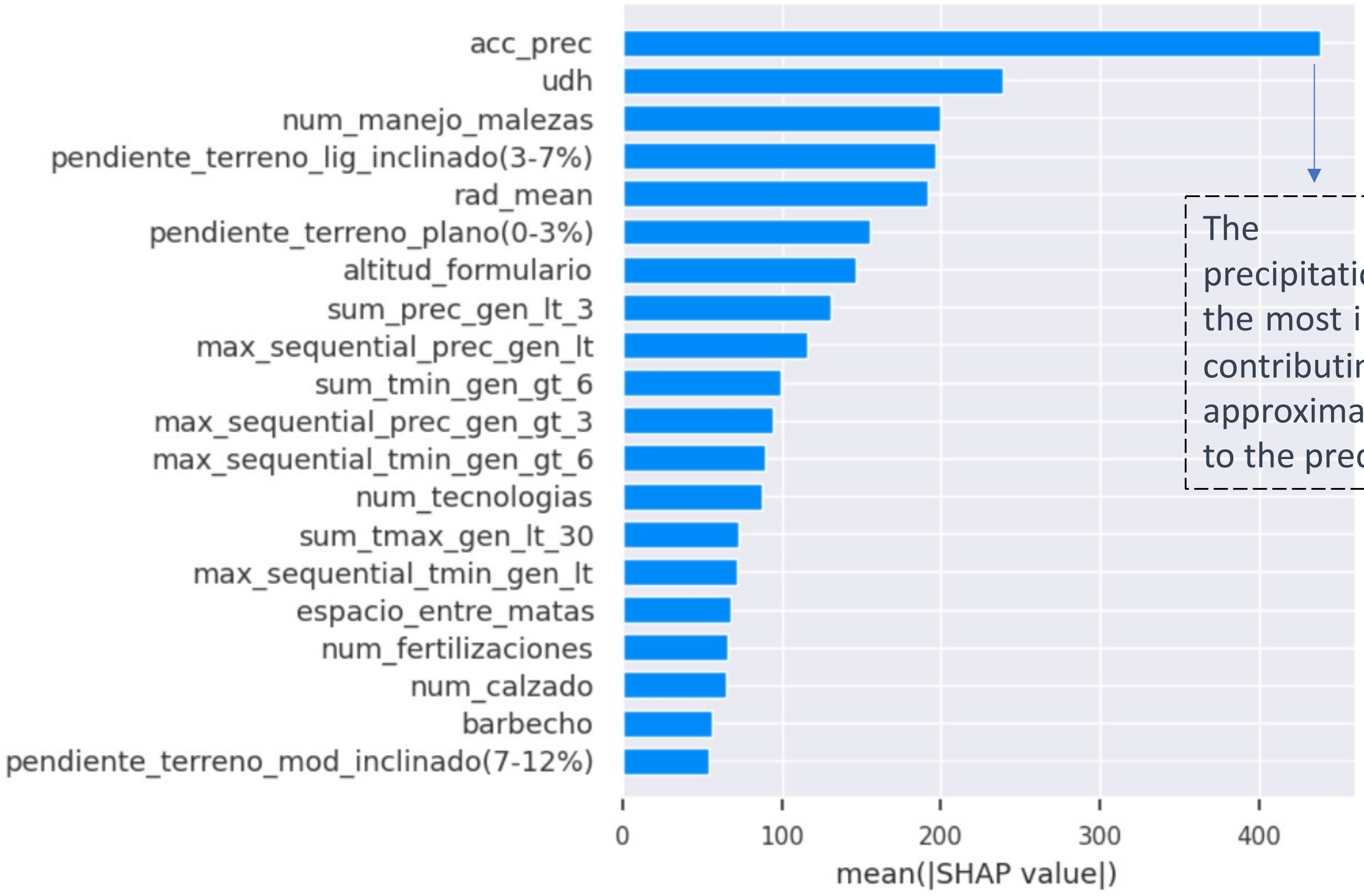
- https://lookerstudio.google.com/reporting/df2beca4-3070-41da-835d-c467d8fb5ceb/page/TYRXD

Model tracking



Bar plot quantifies, on average, the magnitude() positive or negative of each feature’s contribution towards the predicted crop yield.

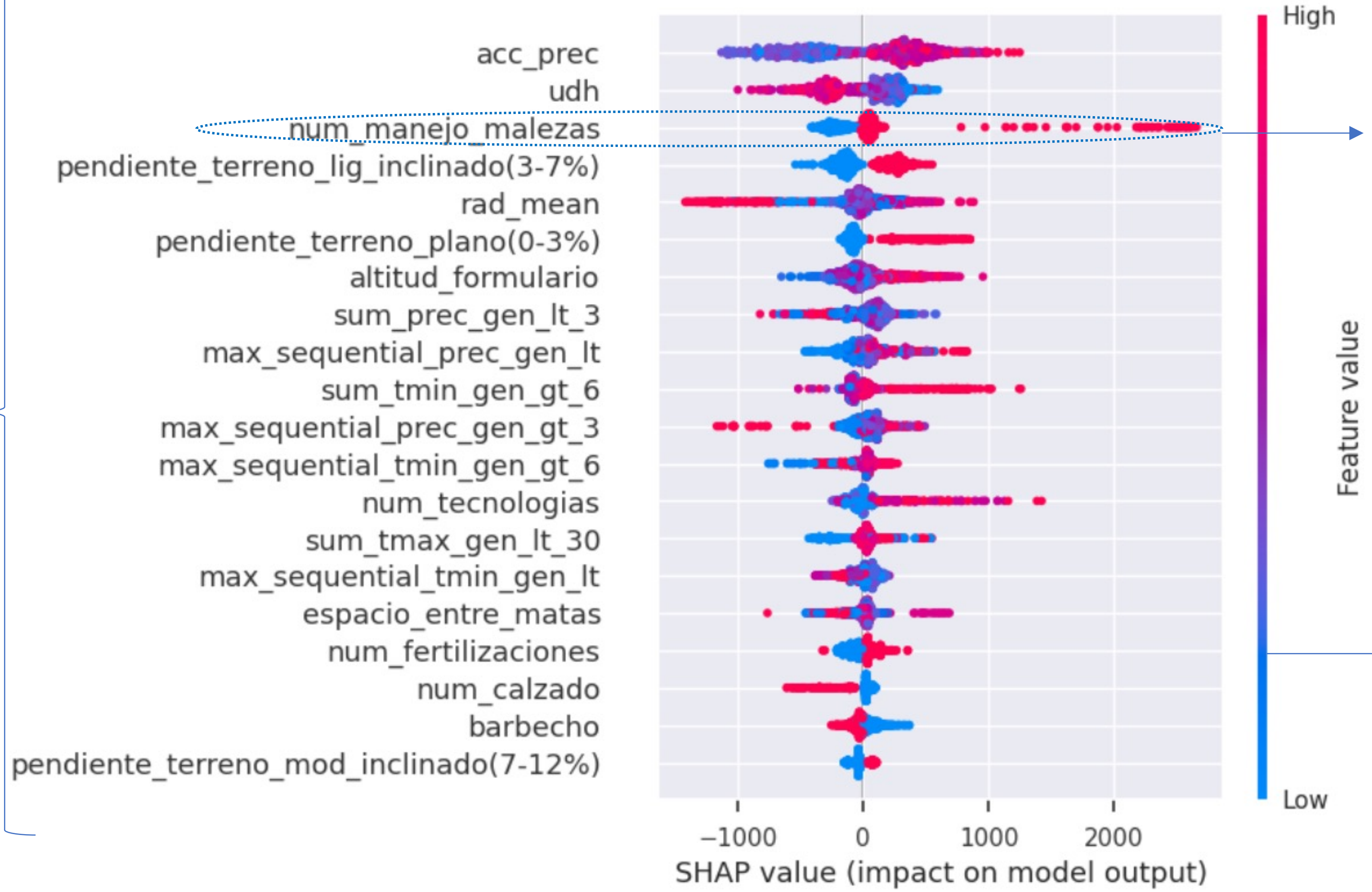
These are the input variables, ranked from top to bottom by their mean absolute SHAP values for the entire dataset. i.e. the average magnitude of each variable’s impact on the predicted crop yield across all instances



The cumulative precipitation (acc_prec) is the most influential feature, contributing on average approximately ±400 kg/ha to the predicted yield

Beeswarm plots reveal not just the relative importance of features, but their actual relationships with the predicted outcome

These are the input variables, ranked from top to bottom by their mean absolute SHAP values for the entire dataset. i.e. the average magnitude of each variable's impact on the predicted crop yield across all instances

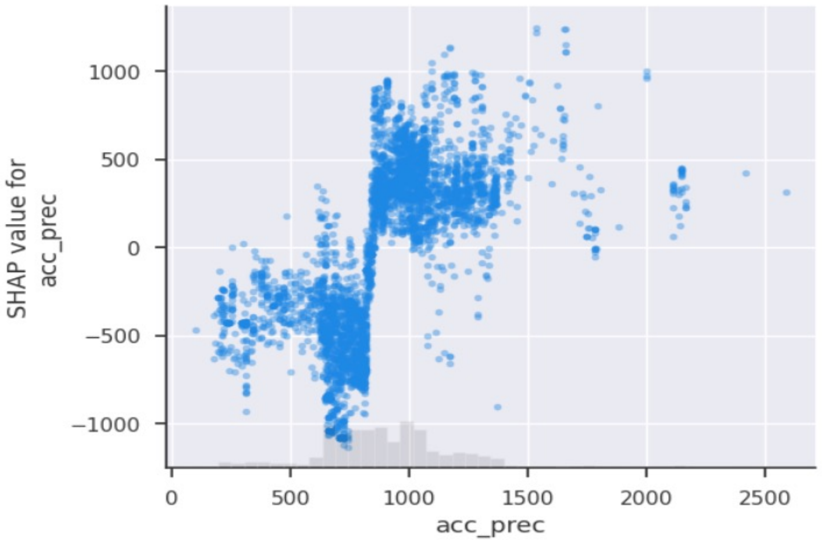


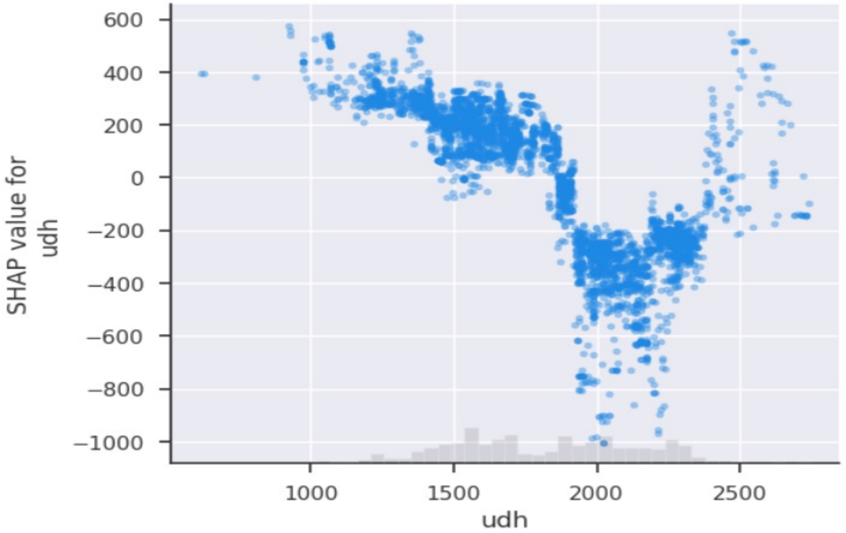
The variable representing the number of times weed management was carried out (num_manejo_malezas), it is observed that the most significant volume has an impact on the model below a SHAP value of 1000. Instances with a positive impact (located to the right) have higher values, suggesting that the more practices are performed, the greater the benefit for the crop.

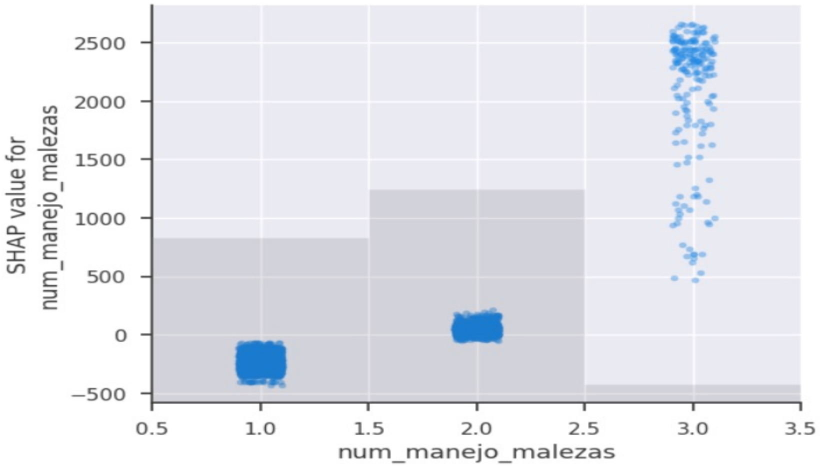
The color bar corresponds to the raw values of the variables for each instance on the graph. If the value of a variable for an instance is high, it appears as a red dot. Low variable values appear as blue dots. Examining the color distribution horizontally along the x-axis for each variable provides insights into the general relationship between the raw values of a variable and its SHAP value

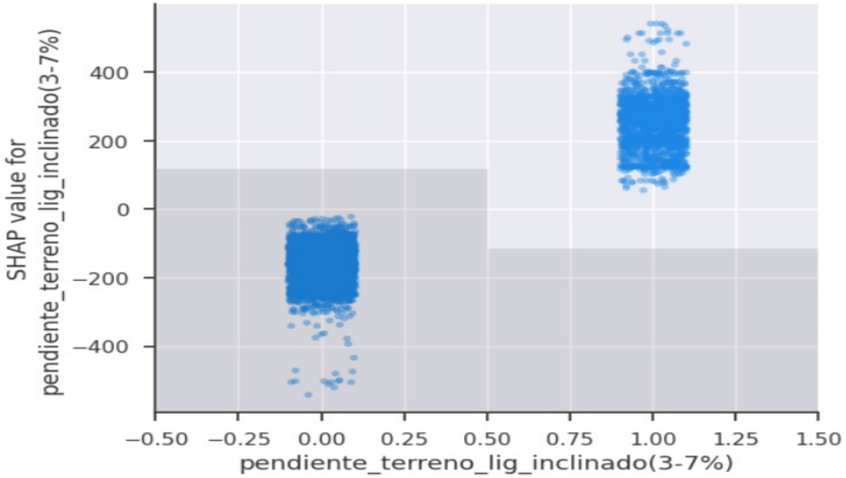
SHAP Dependence Plots

To understand the actual (real) relationship between a feature and the predicted output



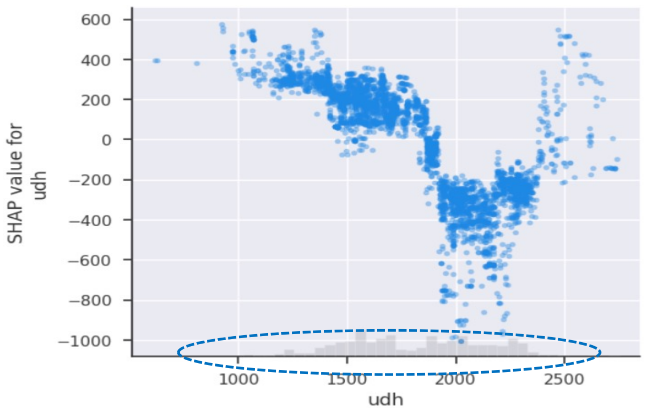




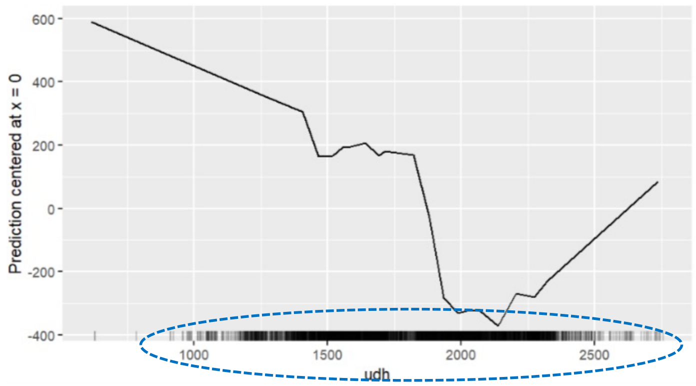


Influence of the variable heat units (udh) on crop yield using Shap, PDP, and ALE

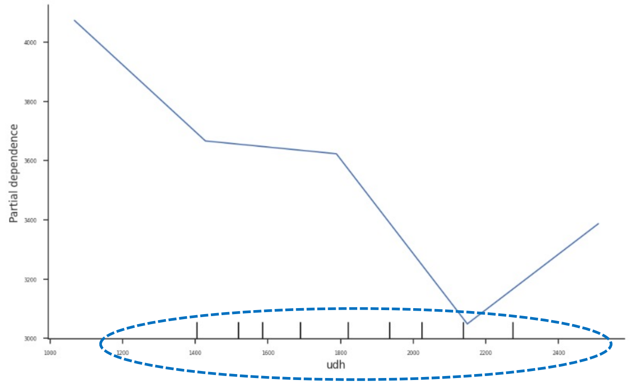
SHAP



ALE



PDP



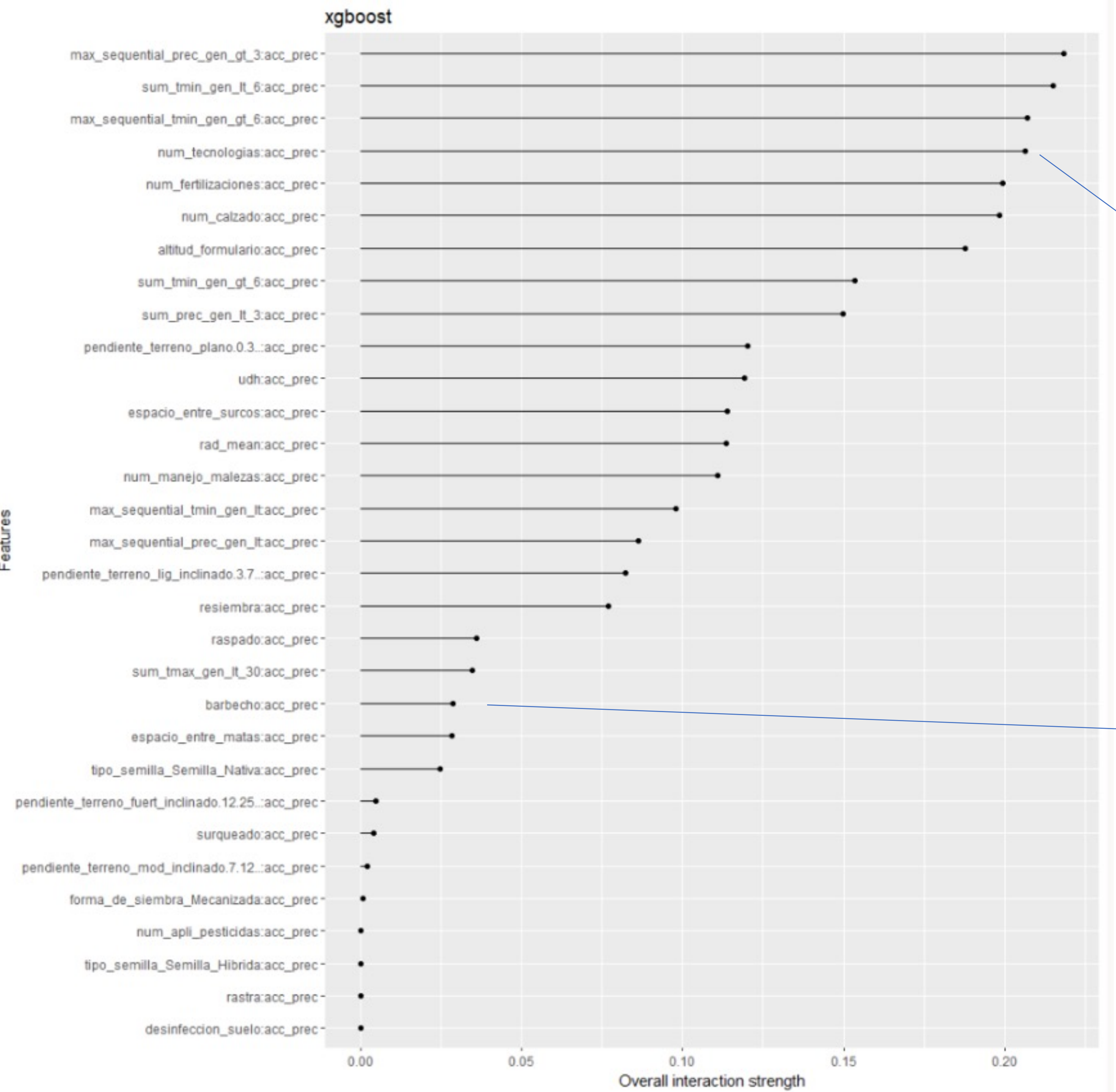
The inset histograms just above the x-axis display the distributions of raw variable values. We should be cautious not to overinterpret regions of the dependence plot where the underlying data is sparse

Shap values above the y=0 line lead to predictions of higher yield crops, whereas those below it are associated with lower yield crops predictions.

The ALE values that are above the y=0 line indicate that the effect of the variable at that value contributes positively to higher corn yield. For example, in the graph, a threshold is observed to determine positive and negative impacts, which is when the heat units (Udh is 2000)

The partial dependence plot shows that as the amount of heat increases, there may be an adverse impact on corn yield. This could be due to extreme weather conditions or changes in the environment that are not favorable for corn cultivation

On the Y-axis, the pairs of features that interact more strongly, with a higher H statistic value in the model prediction, are presented in descending order. In this graph, the interaction of the cumulative precipitation variable with the others stands out.

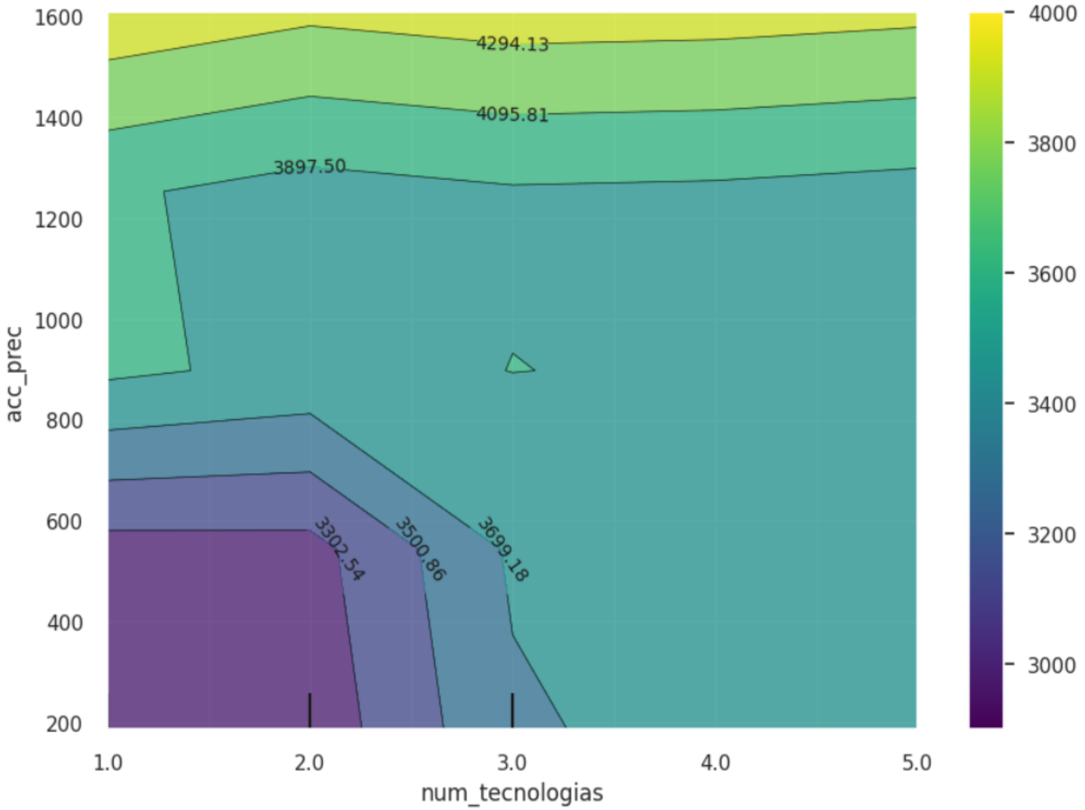


The strength of the interaction between cumulative precipitation(acc_prec) and the number of applied technologies(num_trcnologias) is significant exceeding 20%

The “barbecho” practice does not exhibit a significant interaction strength with cumulative precipitation

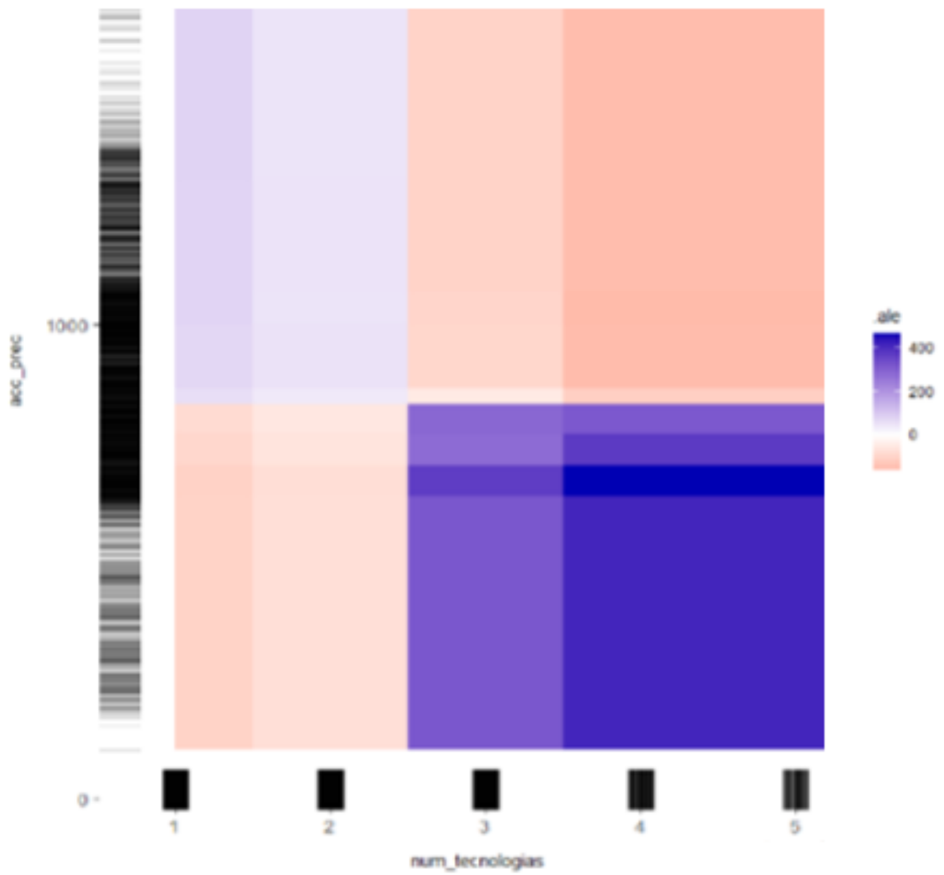
Interaction between the number of technologies and cumulative precipitation in the impact on performance

PDP (10-90%)



The interpretation of the 2D partial dependence plot varies depending on the orientation of its analysis. When examined horizontally, for example, with acc_prec <= 600 mm, it is revealed that as more technologies are employed, the impact on corn yield may experience an increase of approximately 12%. Conversely, when reading the plot vertically, specifically when num_technologies = 3, it is observed that as water availability increases, the impact on yield can reach up to an 18% higher production

ALE



The 2D partial dependence plots assess the strength of the interaction between variables. In the graph, it is observed that the variables acc_prec and num_tecnologia interact positively in two specific scenarios: first, when there is an adequate amount of water (more than 800 mm) and less than 2 technologies are employed; and second, when the water supply does not reach the required threshold, but there is the possibility of applying more technologies.

… a method (still) in the making

Jaimes et al. (in prep.)