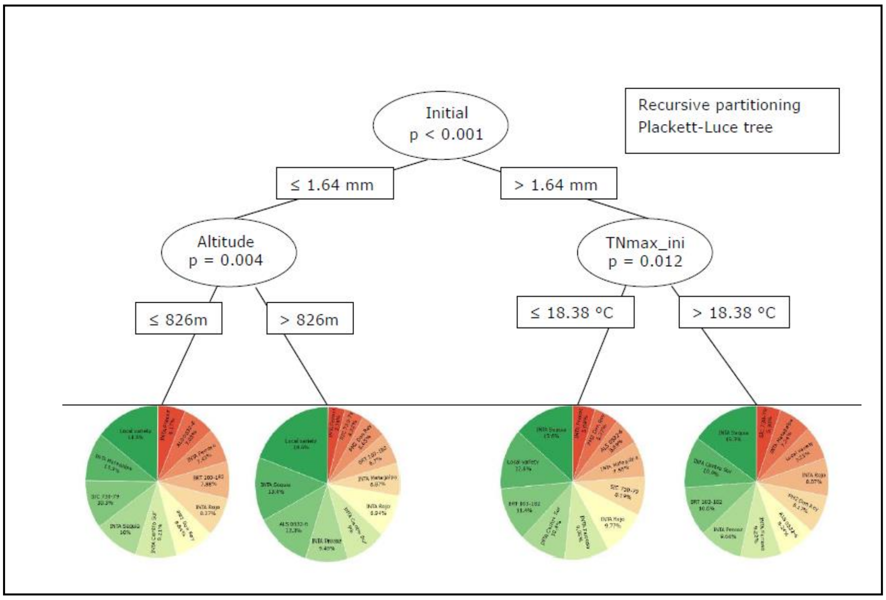
Geo-information Science and Remote Sensing

Thesis Report GIRS-2018-20

Spatial Analysis and Visualization of Triadic Crop Variety Trials in Central America

Elizabeth van Tilborg



2018

WAGENINGEN UNIVERSITY & RESEARCH

Spatial Analysis and Visualization of Triadic Crop Variety Trials in Central America

Elizabeth van Tilborg

Registration number: 95 08 23 832 010

Supervisors:

dr. ir. (S) Sytze de Bruin
dr. Ir. J (Jacob) van Etten, Bioversity International

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

26-04-2018
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2018-20
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Abstract

Climate change influences the yields of farmers. Yields can be sustained by developing and growing different varieties. The Seeds for Needs experiment of Bioversity International, includes on-farm trials in which farmers compare triplets of crop varieties from a larger set are compared against local varieties. This thesis is focused on a dataset of ten common bean varieties tested in Nicaragua. The experiment comprises 842 instances in which three randomly chosen varieties were compared to the local variety of farmers under different environmental circumstances An appropriate method for dealing with such incomplete rankings and environmental interaction is the Plackett-Luce model in combination with model based recursive partitioning. Environmental conditions were characterised using a simple water balance and temperatures in combination with several physiographic descriptors derived from a digital elevation model. The Plackett-Luce model was found able to predict the relative performances of the different bean varieties. The partitioning on environmental conditions, was found to influence the performance of different bean varieties compared to the local varieties of farmers. The results were visualised in maps representing physiographic circumstances, while water availability to and temperature influence on the crop is represented in scenarios.

Keywords: Bradley-Terry model, Plackett-Luce model, model-based recursive partitioning, blocked-cross-validation, incomplete rankings, grouped rankings

Table of Contents

List of tables

List of figures

1. Introduction

Long-term changes in the climate have been observed and are assumed to continue in the future. Temperatures have been increasing, resulting in sea level rise due to melting ice caps and melting of permafrost layers. Increased amounts of precipitation are occurring worldwide, while at the same time more extreme droughts occur (*Lemke, 2007*). Precipitation has always been highly variable over space and time but due to the increased frequency and intensity of the events, this variability is increasing. These developments have created a new pressure on agriculture (*Adams, 1990*). Higher temperatures, more frequent and/or intensive rainfalls and droughts increase production risks and decrease agriculture yields (*Fischer, 2005*). The pressure to sustain or even increase yields in agriculture is further increasing due to socio-economic trends. The population is growing, which results in a higher demand from agriculture. Additionally, urbanization and dietary shifts are driving further pressure on agriculture (*Van Etten, 2016*).

Adaptation of farmers is the key for agriculture to cope with climate change and socio-economic trends, but this requires investments by both farmers and governments (*Lobell, 2008; Muthoni, 2017*). Adapting to the current situation in agriculture can be done with sustainable and relatively inexpensive innovative changes, such as switching to more suitable crop varieties. The increase in yield after using novel varieties might be proven under design conditions, but this is not the sole contributor to adaptation. Soft innovation focusses on the interaction of the technology with the users (*Silva, 2017*). In case of switching to other varieties, this includes introducing the different varieties to the farmers in order to make adaptation of the other varieties easier. It is actually a prerequisite that farmers are well informed about available varieties and their performance before accepting change (*Bentley, 1989; Fielding, 1997*). Proving the technological improvement of different varieties to farmers can be done by introducing farmers to these new varieties by providing seeds and knowledge (*Mwongera, 2014*). The performance of different varieties can be demonstrated to farmers by on-farm trials, resulting in lowering the moral decline towards accepting change and adaptation (*Van de Gevel, 2013*).

One of the soft innovations aiming to help solving the issues surrounding food security may come from an initiative introduced by Bioversity International, a global research-for-development organization. The program, called Seeds for Needs, aims to reduce the consequences of climate change in agriculture by introducing farmers to various varieties of a certain crop using on-farm trials (*Beza, 2017*). The viability of implementing on-farm trials is supported by *Atlin (2001)*, whom states maximal gains in local cropping systems are obtained using on-farm trials and the experience of farmers themselves to maximize yields. The initiative is a micro-level experiment across 14 countries (*Beza, 2017*), founded on the proposal for crowdsourcing crop improvement by *Van Etten, (2011)* in Africa. According to the latter proposal, farmers receive small packages of seeds with different crop varieties for free, which they were asked to evaluate and report their findings using mobile phones. An evaluation consists of comparing one variety to another, based on the yield, consumption quality and marketability. These aspects of performance are combined into an overall performance (*Bioversity, 2016*). The evaluations are subjective, while they depend on each of the farmers' own evaluation. Additionally, these trials were performed as a blind experiment, to reduce the bias of the farmers having prior knowledge on crop varieties, while it simultaneously motivates farmers to complete the experiment. As soon as farmers hand in their evaluation of the relative performance of the crop varieties, the name and type of the evaluated crop varieties are returned. This idea is also supported by *Van Etten (2011)* for similar experiments in Africa and supports the soft innovation as highlighted by *Silva (2017)*.

The evaluations made by the farmers are collected using the digital platform ClimMob (https://climmob.net) via a web platform or a mobile application. The evaluation is done on a series of randomly selected triplets of the to-be-tested crop varieties in a relative scoring of performance. To simplify the experiment and also the evaluation, not all of the to-be-tested varieties are distributed over all farmers, but a triplet (*Bioversity, 2016*). For this triplet they had to assign the 'best' performing variety and the 'worst' performing variety. A relative order could be made per observation for three varieties, excluding the other varieties. An additional comparison between each of the improved varieties from the triplet is made with the local variety of the farmer. Each observation provides thus an incomplete categorical ranking (*Agresti, 2011*). Although the accuracy of the trial observations of each farmers' individual observation may be low, *Steinke (2015)* concluded a sufficient validity is achieved when a large set of combined observations is used.

The incomplete categorical rankings based on the overall performance, are combined per crop and per country. The dataset used in this thesis contains the results of the experiment Seeds for Needs regarding the common bean in Nicaragua. The locations of the trial observations were recorded, which makes it possible to link the observations to environmental variables. The environmental variables might give information as explanatory variables on the evaluated performances of different bean varieties. Information about environmental interaction of different varieties is extracted, while farmers are directly involved in innovation, using different varieties. *Bioversity* and the contributing farmers of the Seeds for Needs initiative are interested in the relative crop performance of different varieties, resulting in the following research question: '*Which crop variety performs best under given environmental conditions?*'

The goal of this research is to contribute to enhanced statistical analysis of the on-farm trial data of the Seeds for Needs experiment. A spatial analysis is part of the result after combining the current available statistical methods with environmental data to highlight spatial variation. The Seeds for Needs experiment is an on-farm trial, which means the seeds are tested on their performance in natural circumstances, which change over space and time. This also makes the experiment non-repeatable and it is harder to predict the (relative) performance of the different varieties. When an experiment is repeatable, the adaptation to certain stresses or production constraints can be characterized and will not be limited to the influence of (unique) local environmental conditions (*Atlin, 2001*). Despite the non-repeatability, the analysis is deemed able to characterize the performance of certain crop varieties over other varieties based on the relative performance linked to environmental conditions. Some environmental factors such as topography can be assumed to be fixed, and can help with describing the performance of a crop variety. Other environmental conditions, such as the weather, are not fixed over time and space but knowledge on variety performance under different weather conditions can help decision making under risk management strategies, such as varietal diversification. Varietal diversification is the use of several varieties of the same crop by a farmer and it can reduce the yield reduction caused by climate variability (*Sukcharoen and Leatham, 2016*). The tricot approach can produce important information about the actual performances of the crop varieties combined with environmental interactions.

The statistical methods used in order to deal with incomplete categorical rankings are the Bradley-Terry model (*Strobl, 2011; Caron, 2012; Bioversity, 2016*) and the Plackett-Luce model (*Caron, 2012; Turner, 2018*). Currently, the experimental approach and the existing data that comes with it are not used to their full potential, due to a lack of the spatial component in the analysis. The Bradley-Terry model covered the performance of the crop varieties explained by extreme precipitation events (*Bioversity, 2016*) and the Plackett-Luce model explained the performance using the planting season and year and the maximum temperature at night during the growing period (*Turner, 2018*). However, the crop in question, the common red bean, also has other optimal environmental conditions (*Steduto, 2012*), which were not taken into account in previous research. The environmental conditions, which are defining this optimal environment, are included in this research and based on work of *Gómez and Blair (2004)* describing the interaction of the common bean with environmental conditions in Nicaragua.

In order to answer the main research question '*Which crop variety performs best under given environmental conditions?*', the following sub-research questions were formulated:

1. What methods are suitable for the statistical analysis of ranked sets of variety trials in combination with explanatory environmental variables?
2. Which environmental conditions influence the crop variety scores, and can these conditions be characterised in Central America to support the analysis?
3. What are the relative scores of tested varieties under different environmental conditions?
4. Which (geo)graphical presentation methods are appropriate for these kinds of analysis?

The methodological background in Chapter 2 is the result of the first sub-question, in which the data type and models to deal with this data are discussed. An explorative analysis is done for Chapter 3 in order to investigate the environmental conditions influencing crop varieties in general, zooming in on the common bean in Nicaragua. The actual methodology, describing the chosen model and the inputs (Chapter 4) result in the relative scores of the tested varieties under different environmental conditions (Chapter 5). The (geo-)graphical presentation is visualised in Chapter 5.

2. Methodological background

This chapter gives a methodological background on the statistical model used for dealing with incomplete ranking datasets. It seeks to explain what incomplete rankings are and how they can be analysed using the Bradley-Terry model and the Plackett-Luce model. Additionally, a method for analysing incomplete rankings in combination with explanatory variables is explained.

2.1 Incomplete rankings

This section tries to define the term 'incomplete ranking' using practical situations. A situation with so-called complete pairwise comparisons is explained using a chocolate example while a situation of an incomplete pairwise ranked evaluations is explained using a soccer pool.

First, assume two types (A and B) of chocolate given to a test panel of an odd number of randomly selected subjects who are asked to give their preference for either type A or type B. In this simple example each member of the panel evaluates the same pair of chocolate types. However, since chocolate preference depends on personal taste, it is likely that the responses of the panel are non-uniform. Nevertheless, the winner and the loser of the test can be determined, along with a measure about the uncertainty related to the observed outcome.

This winning-losing principle can also be applied in, for example, a soccer pool. Assume there are four teams (A, B, C and D) in a pool, playing games against each other. Each soccer team will face the three other teams and will either win, lose or tie. To simplify the example, the teams are either winning or losing, excluding the possibility of ties. Team A will for example win against team B and C, but lose against team D, creating a matrix as shown in *Table 1*. Team D wins in total three games, team A wins two games, team B wins one game and team C loses all games. In this example six games are played, resulting in the order D > A > B > C. This dataset is considered to be a complete ranking dataset, because all teams played against all other teams.

Table 1. Paired comparison matrix: example soccer teams

	Team A	Team B	Team C	Team D
Team A	-	B < A	C < A	D > A
Team B	A > B	-	C < B	D > B
Team C	A > C	B > C	-	D > C
Team D	A < D	B < D	C < D	-

In order to win a soccer game, the rules are very straightforward: the team who made the most goals wins and the other team loses. It does not matter who is observing the game: every observer would see the same result. Sometimes, there are no straightforward rules and the 'win' or 'loss' is based on the observers' taste. In the case of the chocolate example, each observer can make its own preference. The four teams in the matrix in *Table 1* can be replaced by four chocolate types. This matrix could be different from observer to observer, depending on taste. Each observer has its own ranking of the four chocolate types. The resulting dataset is still considered to be a complete ranking dataset, because all types are evaluated against all other types, but now for more observations. In order to find out which type is most preferred, the number of 'wins' per type is simply added and compared to the other amount of 'wins' per type.

In the case of the soccer pool example, a ranking becomes incomplete as soon as not all the games between the teams are played. A game on itself is still a paired comparison, because one team is playing and compared to another team. As soon as one of the paired comparisons is missing, the ranking becomes incomplete. It is possible team A is never compared to team C, because they did not play a game against each other. In a soccer pool with an incomplete ranking, it is not possible to assign a winner. This is different in the case of the chocolate example. A ranking becomes, in the case of the chocolate example, incomplete as soon as an observer is not comparing each type against each other type. When a person is comparing two chocolate types out of the four types, it is not possible to fill in the whole matrix of *Table 1*. It is still possible to find out which sample is most preferred by adding the 'wins' per sample, because there were more observations done by different observers. However, it is important to keep in mind the distribution of the paired samples over the different observers in order to be able to say something about the result.

In the case of the triadic crop variety trial in Nicaragua, there are ten different varieties. The varieties are evaluated based on their performance. When comparing one variety to another, one of them is 'winning', while the other is 'losing'. The same matrix as in *Table 1* can be created for all varieties, resulting in a 10x10 matrix. In a case of a complete data ranking set, each variety would be compared to each other variety and the 'winner' or 'loser' would be determined per observation. If we consider the comparison between one variety to another as a game, there would be 45 games to compare each variety to each other variety. This variety trial results in a final ranking per observer in which the best variety wins against all other varieties based on performance, while the worst variety loses to all other varieties. When combining all observations, the varieties can be ranked, based on their total preference.

The chocolate and the soccer pool example are both paired comparisons, while one thing is compared to another, resulting in a win or a loss. It is also possible to do comparisons between triplets or quartets. The triadic crop variety trial is an example of comparisons between triplets, in which per observation three varieties are relatively compared. For each observation are three paired comparisons between one of the improved varieties from the triplet and the local variety of the farmer added. Combining all observations results in an incomplete ranking dataset, with only the relative performance of three varieties compared to each other out of the ten varieties and paired comparisons between these three varieties and the local variety of the farmer.

2.2 Bradley-Terry model

The Bradley-Terry model (BT-model) (*Van Etten, 2016; Firth and Turner, 2012*) is suitable for dealing with paired comparisons. In the example with the four soccer teams each 'winner' or 'loser' is the result of a paired comparison between two teams. The Bradley-Terry model is an approach to assess the probability of a possible team as 'winner' based on the outcomes of previously played games (*Caron, 2012*). It is possible to account for influencing factors in combination with the outcomes of the previously played games, such as 'home-field advantage' (*Agresti, 2011*). This advantage increases the probability a team wins if playing on their home-field. These factors are called covariates that can be taken into account in the Bradley-Terry model. Estimation of the win probabilities are obtained by a maximum likelihood approach (*Caron, 2012*).

The Bradley-Terry model can be used to assess the probabilities of preferences for a certain variety in a triadic variety trial using pairwise comparisons (*Caron, 2012; Bioversity, 2016*). It is assumed that one variety is performing 'better' than another variety, leading to multiple paired comparisons between two varieties. Since the data ranking consists of three improved varieties and a local variety, a complete single observation consists of six paired comparisons. Within a triplet consisting of the varieties A, B and C, all varieties are compared to each other, resulting in a 'winner' and a 'loser' per paired comparison. In each observation are also three paired comparisons with the local variety of the farmer (local). The six different pairwise comparisons per observation are (A-B), (A-C), (B-C), (A-local), (B-local) and (C-local). However, it should be kept in mind that because of the paired comparisons, a variety is considered to be more often a 'winner' or 'loser' inside one observation. The paired comparisons taken from the three relative performances are not corollary independently. Within a triplet, (A-B-C), the three pairwise comparisons cannot be considered independent. For example, if C was by coincidence favoured by some local circumstances, that will impact two pairwise comparisons (A-C and B-C) rather than just one.

Firth and Turner (2012) provide an implementation of the Bradley-Terry model in R, which is exemplified using a triadic crop variety trial.

2.3 Plackett-Luce model

An extension of the Bradley-Terry model, the Plackett-Luce model (PL-model), is able to deal with more objects in a comparison. Like the Bradley-Terry model, the Plackett-Luce model is also based on maximum likelihood estimation (*Guiver, 2009*). The difference between the BT-model and the PL-model, is that each incomplete ranking per observation can be seen as a whole, rather than the objects in the ranking have to be compared pairwise (*Table 2*). The BT-model used a contest between two objects, resulting in a winner and a loser. The PL-model uses an (incomplete) ranking of all objects available within a single comparison (*Caron, 2012*). For the triadic crop variety trial, this means the ranking of a triplet (A-B-C) could be included as one ranking in the PL-model compared to the three paired comparisons (A-B), (A-C) and (B-C) in the BT-model. The paired comparisons between the improved varieties (A, B or C) to the local variety (local) are by themselves equal, but are treated differently in the BT-model and the PL-model.

Table 2. Partial rankings for Bradley-Terry model compared with the Plackett-Luce model for the Bioversity (2016) dataset

Bradley-Terry model	Plackett-Luce model
1. (A-B)	1. (A-B-C)
2. (A-C)	2. (A-local)
3. (B-C)	3. (B-local)
4. (A-local)	4. (C-local)
5. (B-local)	
6. (C-local)	

The Bradley-Terry model considers the six different rankings as independent observations, while the Plackett-Luce model treats the partial rankings as parts of a single observation (*Caron, 2012*). One observation in the Plackett-Luce model thus consists of four partial rankings, while one observation in the Bradley-Terry model consists of one paired comparison. It was decided to completely shift to the Plackett-Luce model, as this model is capable of dealing with more varieties simultaneously per observation and an incomplete ranking dataset.

2.3.1 Data inconsistency in observations

However, the partial rankings used in the Plackett-Luce model are not necessarily correct. It would be expected that as soon as a certain ranking exists for a location, based on the three varieties, the local variety fits in one location to create a ranking of four. However, it was found in the data the partial rankings per observation were inconsistent, for example:

(A > B > C) Meaning variety A is preferred most and variety C is preferred least

When the local variety was included, it was expected all the sub-comparisons would lead to one particular order of the varieties, such as:

(A > B > C > local) Meaning variety A is preferred most and the local variety is preferred least

The expected sub-comparisons would be:

- (A > local) A is preferred over the local variety
- (B > local) B is preferred over the local variety
- (C > local) C is preferred over the local variety

However, the sub-comparisons were not necessarily in line with the ranking as given in (A > B > C > local). Therefore, the ranking (A > B > C > local) was not used, but a combination of the four rankings: (A > B > C), (A > local), (B > local) and (C > local). The Plackett-Luce model is able to deal with the comparisons, even when they were not consistent (*Turner, 2018*) by listing the four rankings in a matrix. The columns in the matrix represent the 10 different bean varieties and a local variety. The rows represent the partial rankings, resulting in a length of four times the observations. In order to rank different objects from a dataset with incomplete rankings, each object in the ranking is given a value representing the place in the relative ranking per observation.

For the ranking of the three varieties A, B and C, the best performing variety receives value 1, the second best value 2, and the worst performing variety value 3. Variety A, B and C get their assigned values, and the values not evaluated in that observation, are given the value 0. The varieties which were represented with a value 0 were not seen as best or worst, but it was simply a way of making clear they were not observed in that observation. In addition were the three varieties also all compared to the local variety. For the comparison between one of the tree varieties with the local variety, the values for the local variety and the tested variety were either 1 or 2 (1 for best and 2 for second best) and the other varieties were represented by a 0.

As mentioned before is the Plackett-Luce model able to deal with multiple rankings per observation, using the 'grouped_rankings' from the pltree package in R (*Turner, 2018*). The four different partial rankings are combined in 1 column and 1 row as follows:

(A > B > C , A > local , B > local , C > local) 1 observation of four partial rankings

Observations with the format as described above, are the input for the Plackett-Luce model. *Hunter (2004)* elaborates further on algorithms for the generalized Bradley-Terry model from which the Plackett-Luce model is one. Warwick University was requested by Bioversity to develop the Plackett-Luce model in R (*Hunter, 2004; Turner 2017*). The Plackett-Luce model in R is still under development and the package was introduced in R during the time this thesis was formed.

2.4 Model-based recursive partitioning with PL-model

Just like in the example of the soccer pool, there can be advantageous or disadvantageous factors influencing the probability of a certain outcome (in a ranking). In the earlier example, the 'home-field advantage' was referred to. In the case of the chocolate example, a distinction could be made between preferring one sample over another sample after drinking either coffee or tea. The actual chocolate types themselves do not change over the different observers, but the observers themselves can be divided into two groups: coffee drinkers and tea drinkers. The preference of one chocolate type over another type is no longer only influenced by the taste of the observer, but also by the drinking coffee or tea. If a stimulant like drinking coffee or tea might influence the preference of the chocolate types, the dataset can be partitioned into two datasets: a set with observers who were drinking coffee and a set with observers who were drinking tea. The advantages or disadvantages possibly influencing the preference of the samples (such as drinking coffee or tea) is called a covariate. In this way it is possible to find out whether there is a difference in outcomes when the covariates change.

In model-based partitioning is the dataset not partitioned beforehand in for example observers drinking coffee or tea, but is a model-based partitioning algorithm used as explained by *Strobl (2011)*. The dataset will only be partitioned in coffee-drinkers or tea-drinkers if the datasets show a significant difference in preference of the chocolate types. The partitions in the dataset can be seen as and visualised as a tree. *Strobl (2011)* described model-based partitioning as part of the BT-model, but it is also implemented as part of the PL-model using the package *psychotree* in R (*Turner, 2018*). In *Van Etten (2016)* the CART algorithm is used (*Crawford, 1989*), which is based on using one tree with the optimal partition threshold for each node using the whole dataset.

A dataset can contain different groups of observers identified with similar covariates, but with differences in their preference on the outcomes (*Garge, 2013*). The principle of model-based recursive partitioning is similar as for classification and regression trees (*Breiman, 2017*), but model-based partitioning differs in its ability to find differences in values in the covariates in order to create significantly different models in order to fit the total model, rather than finding differences in the preferences of the outcomes (*Strobl, 2011*). In the PL-model, model-based partitioning aims to find significantly different models by stratifying the covariates in order to partition the dataset in such a way, the goodness-of-fit is optimised.

All covariates are tested whether partitioning of this covariate leads to an improvement in the fit of the overall Plackett-Luce model and the covariate with the smallest p-value is chosen to partition the dataset. The lowest p-value corresponds to the highest significance that the use of the partitioning is improving the overall Plackett-Luce model for each node in the Plackett-Luce tree. All the values in the covariates are ordered and subsequent splits are tried. The partition is made where the likelihood of the PL-model is at its maximum (*Strobl, 2011*).

Upon a binary split, each subset is used in order to fit the Plackett-Luce model again. Both subsets are partitioned again, as long as there is significant change in improving the overall fit of the Plackett-Luce model, or as long as no stop criteria are reached (*Strobl, 2011; Turner, 2018*). Stop criteria are for example: reaching a significance level (per node), a minimal subsample size per leaf, or a maximum depth of partitioning (= maxdepth) (*Turner, 2018*). The significance level and minimal subsample size should be chosen beforehand. A significance level of 5% is deemed appropriate, but it should be lower when the sample size is large, otherwise the tree gets too complex for the used dataset. When the model is too complex, it risks induced overfitting: a result might only be valid under a very specific set of covariates, which is hard to interpret. The minimal subsample size should provide a sufficient basis to be able to draw conclusions and should be increased when the number of variables used to get to that subsample size is too high (*Strobl, 2011*).

3. Explorative Analysis

The dataset with the results from the Seeds for Needs experiment contains relative evaluations of performances of different types of the common bean in Nicaragua (*Bioversity, 2016*). Since the location of the trials was recorded, it is possible to link the trial points to environmental variables. An explorative analysis was performed to obtain information about the interaction between crops and environmental conditions and the common bean in general.

3.1 Environmental Conditions

The performance of a crop variety is highly dependent on the interaction with the environment (*Atlin, 2001*) and therefore it cannot be concluded straight away that one crop variety is better than all other crop varieties at every place and time (*Patterson, 1995*). The influence and impact of the presence of certain environmental conditions change over place and over time and interact with crop factors that express the performance of crop varieties. Since the experiment was done under real field-circumstances instead of designed conditions, the environmental conditions should be included in the analysis. The statistical analysis can be adapted to the real field-circumstances by including the circumstances important to the performance of the crop.

The characteristics of a crop influence the interaction between the environmental conditions and the crop performance (*Allen, 1998*). The crop characteristics consist of the growth stages, crop factor and the yield response factor. The growth stages can be divided in the initial, crop development, mid-season and late season stage. For beans in general, is a growing period between 95 – 110 days (*Brouwer, 1986*). The growing period as used in this thesis for the common bean in Nicaragua is 110 days, which is close to the range as mentioned by *Brouwer (1986)* and *Gómez and Blair (2004)*. This growing period can be divided into four different growth stages: the initial growth stage (20 days), the crop development stage (30 days), the mid-season stage (40 days) and the final stage (20 days) (*Brouwer, 1986*).

The next crop characteristic is the crop factor; this crop factor is used to calculate the crop water requirements of a bean. It is of interest to calculate the crop water requirements, this is a measure of the compensation of evapotranspiration of a crop (*Todorovic, 2005*). It states exactly how much water a crop needs from day to day (or over any chosen time span) over the growing period. In order to calculate the actual evapotranspiration of the bean, the evapotranspiration from a reference crop (green grass, 8 – 15 cm tall) is multiplied by the crop factor for beans. This crop factor resembles the relationship of the reference crop to the crop which is calculated. The crop factor changes per growth stage, because per growth stage, the canopy of the crop changes and thus also the evapotranspiration (*Brouwer, 1986*). In *Table 3* are the length of the growth stages and their corresponding crop factor given. The last crop characteristic is the yield response factor, which represents the sensitivity to a water deficit for the crop, resulting in a stress for the crop and a yield reduction. As soon as the yield response factor is higher than 1, the crop is very sensitive to water deficit, especially during the initial and crop development stage. The yield response factor for beans is *1.15*, which means the crop is sensitive, especially during the initial and crop development stage (*Steduto, 2012*). When the crop is older (in the mid-season and late season stage) it is more capable to recover from stress compared to the initial and crop development stage. Therefore, a distinction will be made in the different growth stages to calculate whether the crop has enough, an excess or a deficit of water. Both, the crop factor and yield response factor are considered to be uniform over Nicaragua.

Table 3. Length growth stage of bean and the corresponding crop factor (Kc) (source: Brouwer, 1986)

Growth stage	Length growth stage (days)	Crop factor bean, dry: kc (-)
1. Initial stage	20 days	0.35
2. Crop development stage	30 days	0.70
3. Mid-season stage	40 days	1.10
4. Late season stage	20 days	0.30

Environmental conditions occur spatially clustered rather than randomly spread over space and there are data products describing these conditions. The observations in the experiment have geographical coordinates. Hence, they can be mapped to environmental conditions. The environmental conditions influencing the performance of a crop can be divided into three different conditions: management, weather, and environmental variables (*Allen, 1998*).

Management conditions consist of the managing from the farmer when it comes to irrigation, fertilization and mulching. Irrigation management, the use of fertilizers, mulching and the presence of pest and diseases were not taken into account in this thesis, due to the absence of data for the study area. It was assumed each farmer treated the tested crop varieties equally. Furthermore, the evaluation of the variety performance was relative, and therefore management conditions can be excluded from this research.

Weather variables concern insolation, (air) temperature, humidity and wind speed. These factors all influence the evapotranspiration of a crop, and thus the amount of water the crop needs (*Allen, 1998*). Temperature in general influences the growth of crops, and for each crop variety optimal performance occurs at different temperatures (*Hatfield, 2015*). Additionally, the maximum night-time temperature is considered to be important concerning the performance of the beans (*Bioversity, 2016*).

Based on crop characteristics and weather variables, the crop water requirements of a crop can be calculated. Comparing the crop water requirements to the precipitation on the specific locations, results in either an excess or deficiency of water, leading to a different performance of the crop. Beans need a moderate amount of water (300-600 mm per vegetative cycle), especially early in initial, crop development stage and in the mid-season stage. Late rains are bad for the beans, since their colour changes and the beans have a lower market value (*Gómez, 2004*).

The environmental conditions influencing the adaptation of beans are altitude, slope, soil type, soil depth, drainage, pH and aluminium content (*Gómez and Blair, 2004*).

The altitude influences the available water for the crops and the temperature. Generally, the temperatures are lower at higher altitudes, compared to lower altitudes (*Haverkort, 1990*). The altitude and slopes in an area can result in a difference on the performance of crops. Altitude has a significant influence on the rainfall in an area and subsequently influences the performance of crops (*Trapnell, 1960*). Furthermore, slopes can influence the suitability of land for crop cultivation. Steeper slopes cause more runoff compared to areas with (almost) no slope (*Audsley, 2006*) and additional measures to the land need to be taken in order to obtain the same yield as for flat or gently sloping areas. Additionally, it is harder for farmers to cultivate the land and to harvest the crop, compared to flat areas. The altitude and slope are in this sense, factors influencing the performance of the crop varieties.

The soil type in combination with the soil depth also affect the water potential in the soil. Next to precipitation, the soil type and soil depth influence the ability of the crops to actually use the water. A soil with a higher water holding capacity is more beneficial compared to a soil with no pores, as the plant has more access to water. Furthermore, a plant should also be able to penetrate its roots into the soil in order to actually reach the available water. A deeper soil depth also allows the soil to store precipitation. A high-water content in the soil is good for a plant but good drainage is equally important. In the absence of drainage the roots of a plant are always under water, resulting in oxygen stress. *Letey (1958)* analysed the crop production as a result of different soil physical properties in detail.

When it comes to the chemical composition of the area, the acidity of the soil and the related aluminium content are influencing the performance of crops. A soil is considered too acidic when the pH is lower than 4.0, which reduces the growth of roots (*Foy, 1992*). For beans a pH higher than 5.5 is the minimum requirement, but optimal yields are obtained with a pH around 6.5 (*Gómez and Blair, 2004*). Aluminium in a soil is toxic for a plant and hinder growth; aluminium contents which are too high, reduce the performance of a crop (*Foy, 1992*). Furthermore have soils in Latin America very often a deficit in phosphorus and in combination with a high aluminium content in the soil, beans are affected (*Gómez and Blair, 2004*).

The altitude, slope, soil type, soil depth, drainage, pH and aluminium influence the performance of crop varieties in their own way. Additionally, the different environmental conditions over space, possibly influence the other environmental conditions due to reciprocal correlations. When the varying influences on the varieties and the spatial variation of the environmental conditions itself are combined, the performance of the crop varieties are also assumed to be spatially varying.

3.2 Covariate Selection – Nicaragua

The ecology desirable for the cultivation of the common bean in Nicaragua is described by *Gómez, (2004)*. The environmental conditions as mentioned in the *Table 4* are important for the adaptation of the common bean to different zones in Nicaragua and vary over space. These environmental conditions can give an insight in the adaptation of varieties to the important environmental conditions for the bean in Nicaragua.

For the environmental conditions, the altitude, the slope and the soil type were analysed, as suggested by *Gómez and Blair (2004)*, see *Table 4*. Other environmental conditions, such as the soil depth, drainage, pH and aluminium content are characteristics influencing the adaptation of beans in Nicaragua. The research was limited to altitude, slope and soil type only, simply due to a lack of datasets regarding the other characteristics.

Table 4. Adaptation of the common bean to different zones in Nicaragua based on their climatic and edaphic characteristics (source: Gómez & Blair, 2004)

Adaptation	Altitude (m.a.s.l)	Tempe-rature (°C)	Precipitation		Soil				
			Accumulated (mm)	Period (month)	Depth (cm)	Slope (%)	Drainage	pH	Al (%)
Optimal	450-800	17-24 17-20	200-450	6	>60	<15	Good	6.5	20
Intermediate	200-450	23-27	450-700	4	40-60	15-30	Moderate	6.0	50
Marginal	100	<17 >27	>700 <200	<4 >6	<40	>30	Imperfect	5.5	>50

A literature study resulted in a selection of datasets as mentioned before. Not all above-mentioned variables could be used in this thesis, because for some variables no data were available. Additionally, some variables are more or less homogeneous over the study area so that they are not useful for stratifying crop performance. From the available datasets were the raster-values extracted to the locations of the observations in ArcGIS using the tool; 'Extract Values to Points' or 'Sample'. This resulted in a csv-file, which was used as input for the Plackett-Luce and the tree to fit. The covariates used in the model and their origin are clarified below: altitude, slope, season, soil types, maximum night-temperature (TNmax) and the water balance.

<u>Altitude</u> - The altitude as provided in the Seeds for Needs experiment (*Bioversity, 2016*) was replaced by data extracted from the Digital Elevation Model (*Table 4*) because of lacking information on the source of the altitude as provided by Bioversity.

<u>Slope</u> - The Digital Elevation Model was used to create a slope map, using ArcGIS and the 'slope' tool (percent rise). For each observation in the Bioversity dataset, the slope (in percentage) was extracted from this slope map.

<u>Season</u> - When it comes to seasons, there are three different seasons distinguished in Nicaragua in which farmers cultivate their crops: Primera, Prostrera and Apante. Primera is the first season from the end of May until the beginning June and is the start of the raining season. The second season is called Postrera and the bean cultivation is during this season from September until December. The last season is Apante, which is from November until March and most important in Nicaragua (*Gómez and Blair, 2004*).

<u>Maximum night-temperature (TNmax)</u> - Each observation in the Bioversity dataset has a corresponding planting date. This date is the start of the 110 days growing period. The

temperature data is available in composites of 8 days, therefore both day and night temperatures were linear interpolated, starting from the planting date on. The maximum night-time temperature was important during the implementation of the Plackett-Luce model for the beans dataset in Nicaragua (*Turner, 2018*). The night-temperature is assumed to be the lowest temperature on a day and the maximum values of these minimum temperatures over a growth stage is chosen as maximum night-temperature. Beans are mostly growing during dusk and dawn, which results in a high influence on the performance of the beans considering the maximum night-temperature. The different growth stages were also distinguished, and the maximum night-time temperature for each growth stage was added to the dataset as a covariate. 'TNmax' represents the temperature during the complete 110 days vegetative cycle, the maximum night-time temperatures per initial, crop development, mid-season and late season growth stage are represented respectively: 'TNmax_ini', 'TNmax_cd', 'TNmax_ms', 'TNmax_ls'.

<u>Water Balance (WB)</u> - In order to calculate the water balance, the crop water requirements and precipitation were needed. For the crop water requirements, daytime temperature data (*Table 4*) was needed. For each location in the Bioversity dataset is the temperature per day extracted to the corresponding point.

The crop water requirements were calculated using the formula for determining the evapotranspiration for the reference crop, multiplied by the crop factor for beans for the different growth stages. There are at least four different methods to calculate the evapotranspiration of crops: FAO Blaney-Criddle, FAO Radiation, FAO Penman and Pan Evaporation (*Smith, 1998; Kra, 2013; Subedi, 2015*). Based on the available data the FAO Blaney Criddle method was used, because it only requires measured temperature data, while values for humidity, wind speed and sunshine can be estimated. This method was suitable, since data on humidity, wind speed, sunshine and evaporation for the study area in Nicaragua were not available. The steps to calculate the crop water requirements, using Blaney-Criddle are described by *Brouwer (1986)*.

Step 1: Determination of mean daily temperature: Tmean (°C)

$$Tmax = \frac{\text{sum of all Tmax values during the month}}{\text{number of days of the month}} \tag{1}$$

$$Tmin = \frac{\text{sum of all Tmin values during the month}}{\text{number of days of the month}} \tag{2}$$

$$Tmean = \frac{\text{Tmax} + \text{Tmin}}{2} \tag{3}$$

The values for the temperature available, were the mean daily land surface temperatures. To get the temperature in degrees Celsius, the temperature values as given in the dataset were multiplied with the value 0.02, followed by subtracting 273.15. For each observation, 110 days with temperatures were selected, starting from the given planting date for that observation. Rather than using the mean minimum and maximum temperatures per month, the interpolated daily temperatures were used in the calculation. Using the mean minimum and maximum temperatures is done in order to predict the evapotranspiration from crops for future use. For these calculations is the daily temperature used to calculate the evapotranspiration, because later on the evapotranspiration is compared to the actual daily precipitation for the specific growing period after the planting date.

Step 2: Determination of the mean daily percentage of annual daytime hours: p-factor (-)

The mean daily percentage of annual daytime hours is the percentage of hours in a day (24 hours) in which the sun is shining and influencing the evapotranspiration. The daily percentage of annual daytime was based on the latitude as proposed by *Brouwer (1986)*. The daily percentages were divided by 100 and used as p-factor. The latitudes as given in the Bioversity dataset range from 12.65776 to 13.46576. These values are between the 10 and 15 degrees and therefore the corresponding values from *Brouwer (1986)* were used.

Step 3: Calculate evapotranspiration reference crop ETo (mm/day) with the Blaney-Criddle Method

$$ETo = p\ (0.46 * Tmean + 8\ ) \tag{4}$$

ETo = Reference crop evapotranspiration (mm/day) as an average for a period of 1 month

Tmean = Mean daily temperature (°C)

p = Mean daily percentage of annual daytime hours

Step 4: Calculate crop evapotranspiration ETcrop (mm/day) (bean)

$$ET\ crop = ETo * Kc \tag{5}$$

ET crop = Crop evapotranspiration or crop water need (mm/day)

Kc = Crop factor

The crop factor is dependent on the type of the crop, the growth stage but also the climate. Crops grow faster in warmer climates, compared to colder climates and strong winds ask for higher Kc factors compared to climates with little wind. The difference in crop factors could be 0.05 higher for high wind speeds and 0.05 lower for low wind speeds (*Brouwer, 1986*).

For every location the actual evapotranspiration per day in millimetres was calculated with a total length of 110 days. The sum was taken of the total actual evapotranspiration per growth stage and compared to the precipitation (*Table 4*) of that same time span. This resulted in a water balance (an excess or deficit of water availability) per location per growth stage. The values of this water balance for each location per growth stage starting from the planting date on were added to the Bioversity dataset under the names: 'Initial', CropDev', 'MidSeason', and 'LateSeason'.

All precipitation was assumed to be effective precipitation, meaning all the water becomes available for the crops. In other words, deep percolation and run-off which are part of the total water balance were ignored. This was done because of lack of data about the soil depth. Furthermore, it should be noted that irrigation was not taken into account during any of the calculations, because the varieties are rainfed (*Jacob van Etten (Bioversity), personal communication, April 13 2018*).

Soil types - The soil types were used as provided by Bioversity. The geostatistical wizard in ArcGIS was used to perform indicator kriging in order to interpolate the soil types between the known soil points. The soil types were divided into two categories, combining Eutric Regosols with Haplix Nitisols and Haplic Phaeozems with Humic Nitisols. The result was a raster with the probability where which category was expected. The selection of the soil types per category was chosen and made sense when the results of the Plackett-Luce model had to be visualised over space [*see: Chapter 5.1.3*].

Altitude, temperature and precipitation were obtained from additional datasets to complete the dataset, see *Appendix I*. The environmental variables used as covariates in the Plackett-Luce model were season, altitude, slope, soil type and crop water requirements (*Table 4*):

Table 5. Environmental variables used as covariates in the Plackett-Luce model

Covariates	Measurement scale / unit
Season	Categorical: {Primera (May – August), Postrera (September – October), Apante (November – January)}
Altitude	Meter above sea level
Slope	Percentage slope
Soil type	Categorical: {Eutric Regosols, Haplic Nitisols, Haplic Phaeozems, Humic Nitisols}
Night-time temperature	The maximum night-time temperature (°C/growth stage)
Water Balance[1]	mm/growth stage

3.3 Correlations between covariates

The Plackett-Luce tree will partition the data on thresholds values in a covariate. It is possible some environmental conditions are closely correlated, and the partitioning is in that case not only representing one environmental condition, but two or even more when correlations are high. In order to investigate this, the linear correlations between the different environmental conditions were calculated. The season and soil type are not included in the correlation calculations, because they are both nominal data. In *Figure 1* the correlations between the covariates are shown. Since

[1] A negative value indicates a water shortage, while a positive value indicates a water surplus

the Plackett-Luce tree uses model-based partitioning, the resulting tree will be optimal. Therefore all covariates are given as input for the model. The correlations to the covariates as chosen and used to partition in the Plackett-Luce model are used in order to discuss the results.

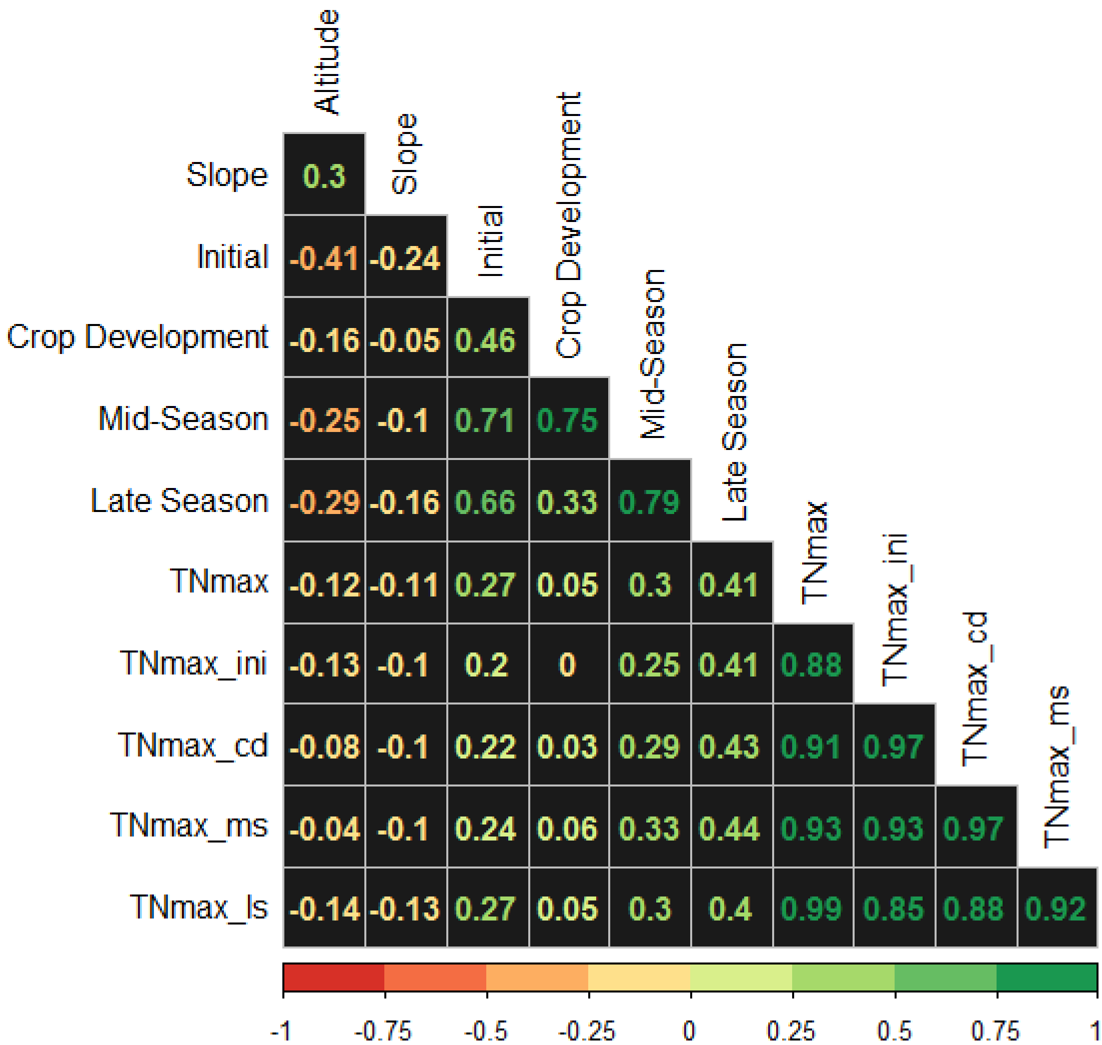


Figure 1. Correlations between covariates used in Plackett-Luce model (Altitude, slope, water balance (initial, crop development, mid-season and late season stage), TNmax maximum night-temperature (whole vegetative cycle, initial, crop development, mid-season and late season stage)

4. Materials and methods

4.1 Dataset Seeds for Needs

The dataset provided by Bioversity comprises results of an on-farm trials distributed over four provinces (Atlantico Norte, Jinotega, Matagalpa and Boaco) in Nicaragua (*Figure 2*).

Figure 2. Tricot Locations per province in Nicaragua (source: Bioversity, 2016 and National Geographic, Esri, Garmin)

The dataset consists of 842 different locations in Nicaragua, where bean varieties were assessed based on their relative quality in performance. The ten bean varieties used in the experiment are: ALS 0532-6, BRT 103-182, INTA Centro Sur, INTA Ferroso, INTA Matagalpa, INTA Precoz, INTA Rojo, INTA Sequia, PM2 Don Rey and SJC 730-79. Each observation includes the relative performance of three randomly chosen crop varieties on a certain location (longitude, latitude). The relative performance of the bean varieties are incomplete categorical ranked data. The data is categorically ranked, with incomplete rankings, because a crop variety is ranked based on the classes: 'best', 'medium', and 'worst'. For every location the relative performance of three out of the ten bean varieties was evaluated. The three bean varieties in each observation were also relatively evaluated to the local bean variety the farmer grew, resulting in classes: 'better' and 'worse' performing varieties compared to the local variety. Each farmer is cultivating his/her own local bean variety, but the local varieties are not necessarily equal for all farmers.

Each observation also includes the soil type, the planting date, the season and year on/in which the beans were cultivated. To augment the existing dataset, different environmental conditions were added, based on their importance to the adaptation of the beans to the environment and the availability of the data.

4.2 Model-based partitioning with Plackett-Luce model

In order to fit a tree in the Plackett-Luce model, the different partial rankings were used per observation. In the usage of the PlackettLuce model in R were the observations connected to the covariates. In summary the following steps were taken in order to fit the Plackett-Luce model, combined with the model-based partitioning for different bean varieties in Nicaragua:

1. A Plackett-Luce model with all observations was fitted, which means the crop varieties were ordered per observation and their worth parameter to perform best are the result;
2. The stability of the worth parameters (relative performance) is assessed with respect to each available environmental conditions (covariate);
3. "If there is significant instability, split the full data by the covariate (environmental condition) with the strongest instability and use the cut-point with the highest improvement in model fit;
4. Repeat steps 1-3 until there are no more significant instabilities, or a split produces a sub-group below a given threshold" (*Turner, 2018*)

The following model parameters were chosen:

- Alpha : 0.01
- Minsize : 100
- Maxdepth : 5

Alpha is the significance level of the instability of the covariates in the Plackett-Luce tree. The minimal amount of samples in a leaf of the tree was set at 100, which is more than 10% of the total sample size. A minsize of at least 100 observations was assumed to avoid over-fitted models. The maximum depth of the model was set on 5 in order to avoid the model to get too complicated to interpret.

4.3 Scoring varieties

All covariates were used as inputs in order to score all the varieties using model-based recursive partitioning in the Plackett-Luce model (*Table 5*).

The model resulted in a visualisation of a tree in which the rankings were visualised per leaf based on worth parameters. Every node in the tree is a partitioning in the dataset, based on the covariate with the strongest instability, resulting in a different outcome of the Plackett-Luce model fit. The covariates in the model are environmental conditions used as explanatory variables for the preference of the different crop varieties (*Turner, 2018*). The worth parameters for the crop variety dataset (*Bioversity, 2016*) in the Plackett-Luce model represent the preference of one variety over another variety compared to a reference variety, based on their relative performance.

The reference variety, the local variety, got the value 0.0 and the other varieties perform better (>0) or worse (<0) compared to the reference variety. The crop varieties were ordered per observation and the probability of each variety to perform as best was calculated based on the probability of other varieties and their performance in other observations (*Turner, 2018*). Varieties with a value higher than 0, perform better than the first tested variety and the variety with the highest worth parameter is the 'winner'. Comparably, every variety with a worth parameter lower than 0 represents varieties performing worse than the reference variety, and the variety with the lowest worth parameter performs worst. The worth parameters were displayed, as default plot, on a log scale. The worth parameters were rescaled until they summoned to 1, giving probabilities which were used in order to visualise the relative preferences of varieties per given set of environmental conditions.

4.5 Validation Plackett-Luce model

The validation of the Plackett-Luce model was done with a variation on the k-fold cross validation (*Kohavi, 1995*). The leave-one-out method was not considered to be of additional value, because the model is fit to all data except for one observation. The ranking for the observation left out can be predicted, but because the rankings are relative and incomplete per observation, the leave-one-out method does not give the goodness-of-fit for the model. The choice was made to validate the model using a blocked-cross-validation, in which not one observation, but a set of observations is left out.

In this case, the data is partitioned, based on the different growing seasons (Primera, Postrera and Apante) per year resulting in five subsets of seasons. In each cross-validation, a Plackett-Luce tree is created based on four out of five subsets of the data, using only the covariates used in the tree. The deviance is calculated over the dataset not used for training. This is done five times in order to get a deviance for each season. The deviances per season were divided by the square root of observations in that season, reducing the bias resulting after dividing the dataset in folds based on seasons. The deviances were aggregated by summation per model. This method was discussed with and explained by Kaue de Sousa (*Consultant Bioversity, personal communication, April 25, 2018*) and Jacob van Etten (*Bioversity, personal communication, April 13 2018*).

To investigate whether the Plackett-Luce model with model-based partitioning was sensitive to including the rainfall and temperature dependent water balance, the model was firstly run without the water balance per growth period. The sensitivity of the model to using the maximum night-time temperatures (per growing period) was also investigated this way. A weighted negative log likelihood was calculated and used in order to compare the influences of the covariates by leaving them out of the model fitting.

To check whether the results were in line with the results in previous research, the outcomes of the model in this thesis was compared with the outcomes of the Bradley-Terry model (*Bioversity, 2016*), the Plackett-Luce model (*Turner, 2018*) and the environmental conditions influencing the adaptation of the bean in Nicaragua (*Gómez and Blair, 2004*). The comparisons between the results as found in this thesis and the previously found conclusions are done by comparing the top three in the rankings resulting from the different partitioned dataset rankings.

4.6 Presentation of outcomes

The presentation of the outcomes consist of roughly three parts:

- Recursive model-based partitioned tree from the Plackett-Luce model.
    - Using all covariates [Chapter 5.1.1]
    - Using spatial covariates, season and water balance [Chapter 5.1.2]
    - Using spatial covariates [Chapter 5.1.3]
    - Validation of models [Chapter 5.2]
- Visualizations of the preferred common bean varieties under certain environmental conditions in Nicaragua and the probabilities a variety is preferred per region and/or scenario [Chapter 5.3].

The result of a Plackett-Luce tree is a tree with nodes, representing partitioning of the dataset, and leaves representing graphs showing the worth parameters of the varieties in comparison to a reference variety. These worth parameters can be converted into probabilities a variety is preferred under a certain set of environmental conditions. The spatial variation can be distinguished and visualized in graphs (*Bioversity, 2016* and *Turner, 2018*), but it is either possible to visualize the outcomes of the model in a spatial way by creating maps with the specified environmental conditions and the corresponding crop variety preference. Some environmental conditions are spatially explicit, such as the soil type, the altitude and the slope. These conditions can be mapped, because their spatial variation might explain the locations of the preferences for different varieties. Some covariates are not spatially explicit, because they change over time rather than over space, such as temperature and precipitation. The covariates for the TNmax and water balance per growth stage might be dependent on the planting date of the observations, because they cover a different growing period, influencing the performance of the varieties.

When several spatial conditions were used as covariates in the Plackett-Luce model, the different spatial conditions were classified in ArcGIS. This resulted in areas with the same spatial environmental conditions as represented in a node in the result of the Plackett-Luce model. These areas were called regions and connected to a probability pie of preferred varieties. Probability pies adjacent to a map make for an easier user read and allow for a better visualization of more data, compared to superimposed maps (*Luzzi, 2016*). The same was done for environmental conditions, which could not be represented over space. A region is considered to be spatially explicit, while a scenario is considered to be uncertain over time and space and therefore not visualised in a 2D map. The spatially explicit regions can be visualised with maps and to each region in the map corresponds a probability pie. There are also probability pies for scenario's in which the conditions are met as given by a not meeting or exceeding the threshold value in a node in the Plackett-Luce tree. Scenario's do not necessarily have to be corresponding to a certain region in a map, but can be based on covariates changing over space and time only.

5. Results

5.1 Plackett-Luce Model

The visualization of the model-based partitioning in the Plackett-Luce model resulted in a partitioning tree. The worth parameters of different varieties under the environmental conditions pertinent to the leaves of the tree are displayed in plots. In the previous chapter, three different combinations of environmental conditions were described, which each resulted in a tree. The first tree could choose from all covariates, i.e., spatial components (altitude, slope, soil type), season, the water balance (per growth stage) and the TNmax (maximum night temperatures). The second tree could only use the spatial components, the season and the water balance. The third tree was built using only the spatial components.

5.1.1 Covariates: spatial components, season, water balance and TNmax

The tree based on all covariates shown in *Figure 4* has splits on the water balance of the initial growth stage, altitude and the TNmax (max night temperature) of the initial growth stage. The full dataset is first partitioned with a threshold of an excess of 1.64 mm on the water balance in the initial growth stage. When this threshold is not exceeded, the dataset is partitioned, based on an altitude of 826 meters. Worth parameters for the varieties resulted after the second partitioning on the left side in the preference of the local variety (11) over all other varieties. However, at an altitude below 826 m.a.s.l., the second-best variety was INTA Matagalpa, closely followed by SJC 730-79 (10) and INTA Sequia (8). At altitudes above 826 meters, the differences between the varieties were more pronounced. The SJC 730-79 (10) performed second worst, while the INTA Sequia (8) performed the second best after the local variety (11).

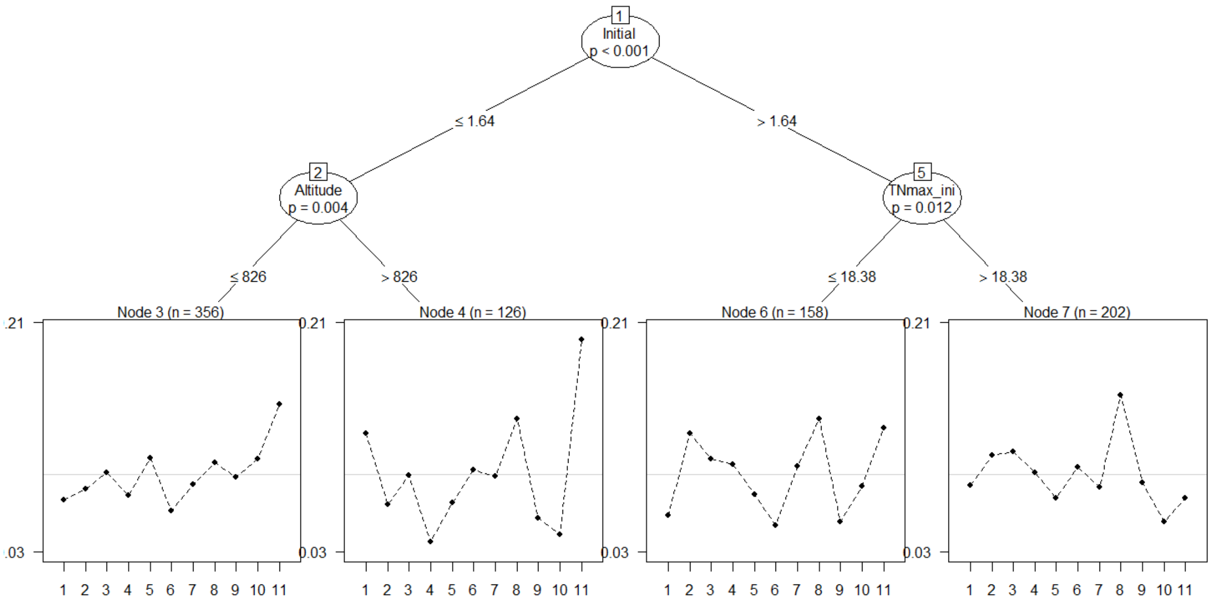


Figure 3. Worth parameters of overall performance of the ten trial varieties and the local variety of common beans in Nicaragua for each node in the Plackett-Luce tree built using all ccovariates.
(Covariates: Initial (water balance during initial stage in mm), Altitude and TNmax_ini (maximum night temperature during initial growth stage in °C), p refers to significance level of the instability of the covariate) (Varieties are: 1: ALS 0532-6, 2: BRT 103-182, 3: INTA Centro Sur, 4: INTA Ferroso, 5: INTA Matagalpa, 6: INTA Precoz, 7: INTA Rojo, 8: INTA Sequia, 9: OM2 Don Rey, 10: SJC 730-79, 11: Local Variety)

On the right side, for observations with more than 1.64 mm on the water balance during the initial growth stage, a partitioning was made based on the TNmax during the initial growth stage. When the maximum night-temperature dropped beneath 18.38 °C, the INTA Sequia (8) performed best and it was very closely followed by the local variety (11) and the BRT 103-182 (2). When the maximum night-temperature exceeded 18.38 °C, the INTA Sequia (8) again performed best while the local variety (11) performed second worst.

5.1.2 Covariates: spatial components, season and water balance

The covariates chosen for the tree in *Figure 4* excluded the TNmax for all growth stages. The left side of the tree is comparable to the tree in *Figure 3*, as the first partitioning is again based on the water balance during the initial growth stage. Also the second split is again on altitude. The worth parameters for the first two nodes in both trees (*Figures 3 and 4*) are therefore equal.

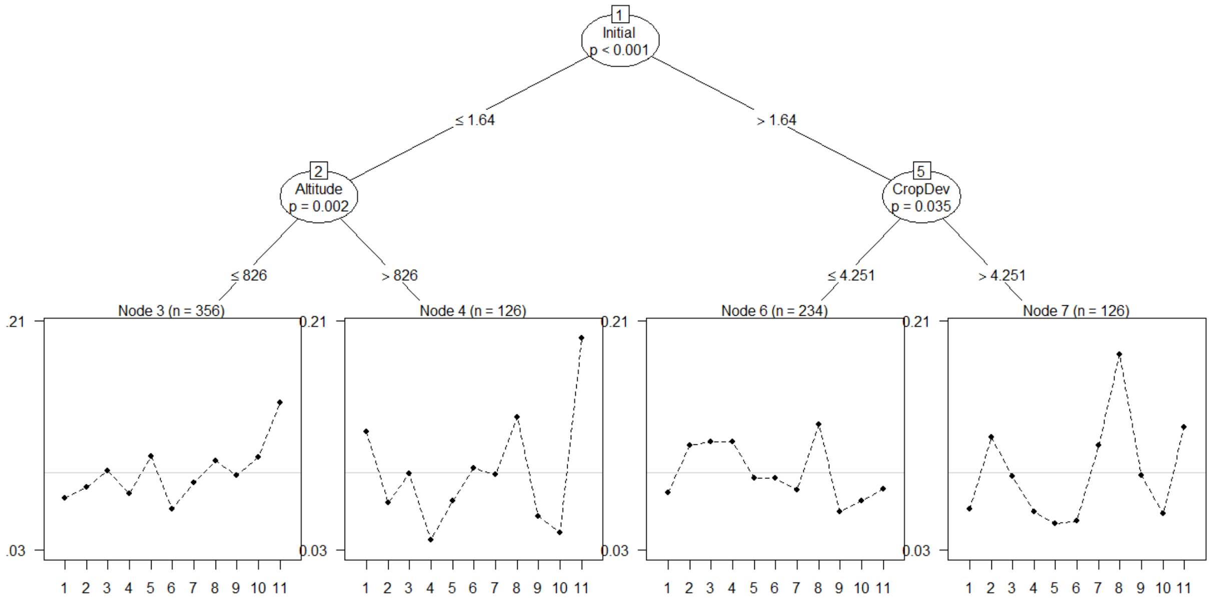


Figure 4. Worth parameters of overall performance of the ten trial varieties and the local variety of common beans in Nicaragua for each node in the Plackett-Luce tree built using spatial components, season and water balance.

(Covariates: Initial (water balance during initial stage in mm), Altitude and CropDev (water balance during the crop development stage in mm), p refers to significance level of the instability of the covariate)
(Varieties are: 1: ALS 0532-6, 2: BRT 103-182, 3: INTA Centro Sur, 4: INTA Ferroso, 5: INTA Matagalpa, 6: INTA Precoz, 7: INTA Rojo, 8: INTA Sequia, 9: OM2 Don Rey, 10: SJC 730-79, 11: Local Variety)

On the right side of the tree, the next split is based on the water balance during the crop development stage (the second growth stage). In case of less than 4.25 mm on the water balance during that stage, the INTA Sequia (8) performed best, but was closely followed by the INTA Centro Sur (3), INTA Ferroso (4) and BRT 103-182 (2). The local variety (11) performed worse than most other varieties.

On the other hand, if the water balance exceeded 4.25 mm during the crop development stage, INTA Sequia (8) performed best followed by the local variety (11) and BRT 103-182 (2).

5.1.3 Covariates: spatial components

The covariates of the tree shown in *Figure 5* were chosen from the spatial components only. Altitude, with a threshold at 526 m.a.s.l., is used as the first split. Cases below the threshold were not subdivided: INTA Sequia (8) was most preferred, closely followed by the INTA Centro Sur (3) and the local variety of a farmer (11).

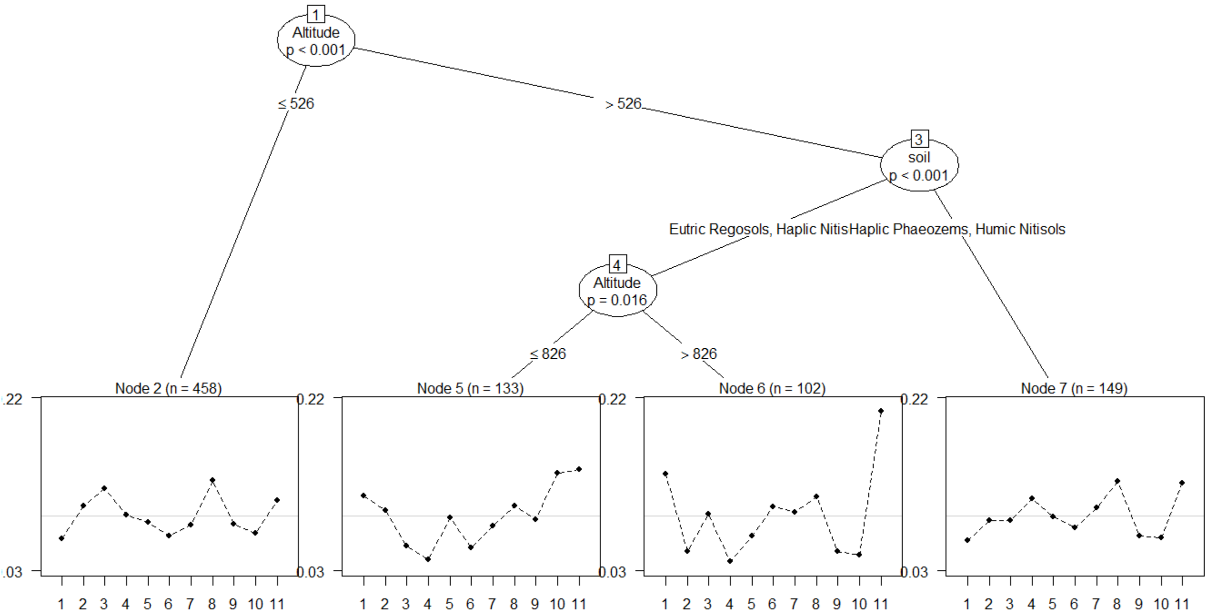


Figure 5. Worth parameters of overall performance of the ten trial varieties and the local variety of common beans in Nicaragua for each node in the Plackett-Luce tree built using only spatial components.

(Covariates: Altitude and soil type, p refers to significance level of the instability of the covariate)
(Varieties are: 1: ALS 0532-6, 2: BRT 103-182, 3: INTA Centro Sur, 4: INTA Ferroso, 5: INTA Matagalpa, 6: INTA Precoz, 7: INTA Rojo, 8: INTA Sequia, 9: OM2 Don Rey, 10: SJC 730-79, 11: Local Variety)

At altitudes above 526 m.a.s.l., another split based on the soil type was made. On Haplic Phaeozems or Humic Nitisols, the INTA Sequia (8) and the local variety of the farmer (11) were most preferred. On Eutric Regosols or Haplic Nitisols below 826 meters the local variety (11) performed best, very closely followed by the SJC 730-79 (10) while above 826 meters, the local variety (11) outperformed all other varieties and the second best performing variety was ALS 0532-6 (1).

5.2 Validation Plackett-Luce models

The validation of the Plackett-Luce models was done by a blocked-cross-validation in which a weighted negative log likelihood is calculated for each of the different covariate selections (*Table 6*).

Table 6. Negative log likelihood from a blocked-cross-validation for the different inputs in the Plackett-Luce models

Covariates in Plackett-Luce tree model	Negative Log Likelihood [blocked-cross-validation]
Water balance (initial), altitude, TNmax (initial) [Chapter 5.1.1]	1009
Water balance (initial, crop development), altitude [Chapter 5.1.2]	1029
Altitude and soil type [Chapter 5.1.3]	1026

The model with the lowest negative log likelihood is the Plackett-Luce tree in which all covariates were used and the water balance during the initial stage, the altitude and the maximum night-temperature during the initial stage were selected in the tree for partitioning. The negative log likelihood after excluding the maximum night-temperatures is higher, compared to including them. Excluding the water balance, leaving only the spatial covariates altitude and soil type, leads to a negative log likelihood which is lower than the model including the water balance, but higher than the model including the water balance and the maximum night-temperatures.

5.3 Probability preference of varieties

The worth parameters of the varieties resulting from the Plackett-Luce trees were converted into probabilities of each variety to be preferred over the other varieties. For each leaf of the trees shown in *Figures 3-5*, the probability distribution is shown in a pie-chart, in which the probabilities are given as percentages.

5.3.1 Covariates: spatial components, season, water balance and TNmax

The spatial variation in the split of node 2 in *Figure 4* is visualised in regions in a map (*Figure 6*) corresponding to two pie charts (*Figure 8 and 9*) and the split of node 5 is visualised in two different scenarios using pie charts (*Figure 10 and 11*).

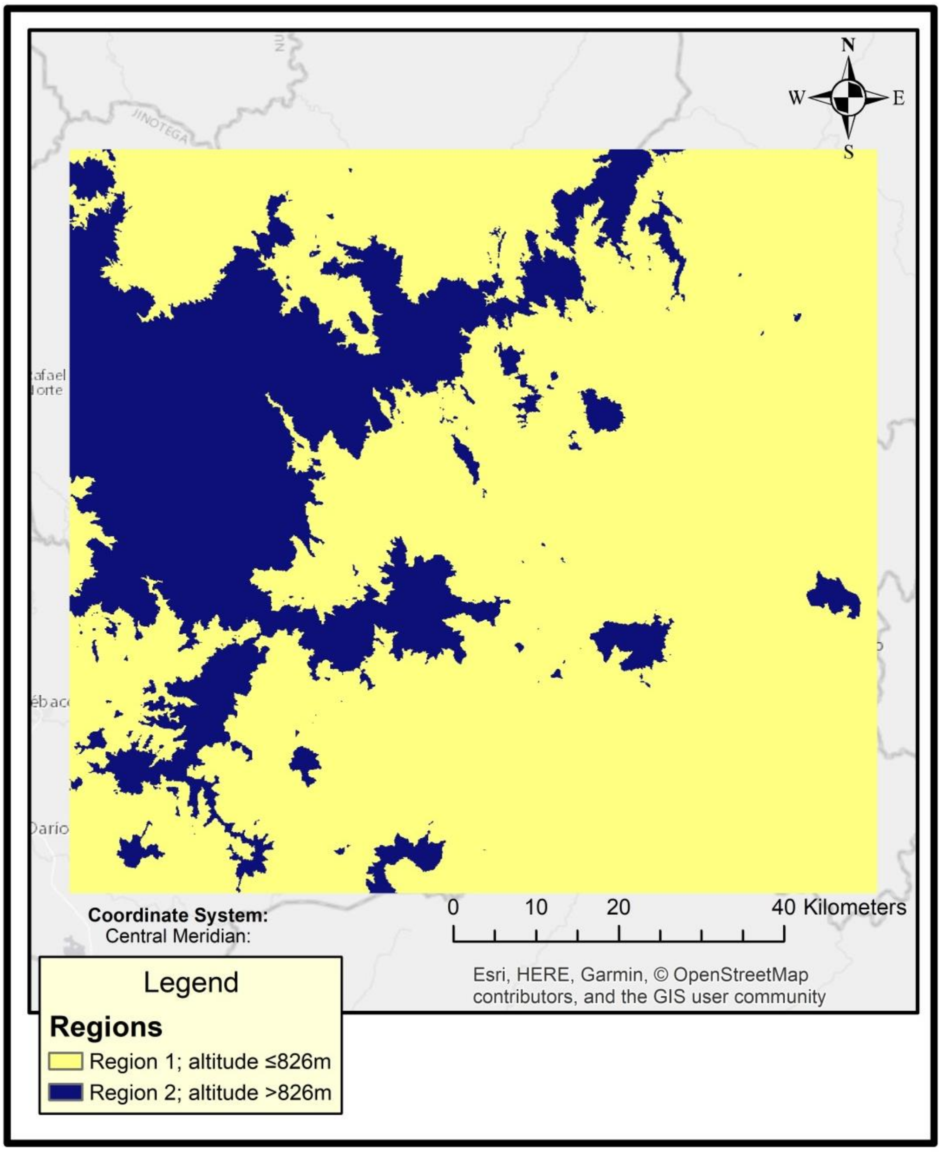


Figure 6. Regions based on altitude relevant to the PLtree of Figure 4.

When the result on the water balance during the initial growth stage is below 1.64 mm, the local varieties were most probable to be preferred over the other varieties (*Figure 8 and 9*).

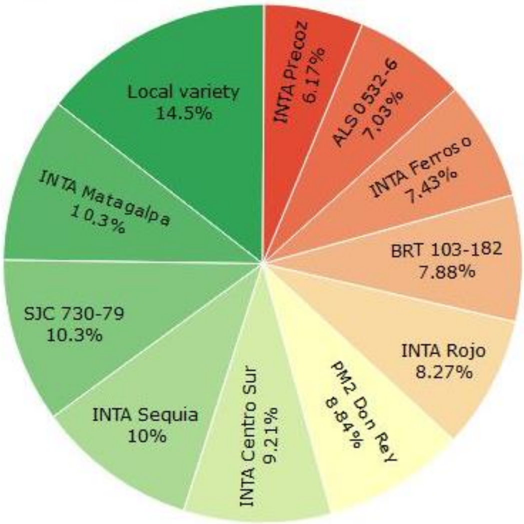


Figure 7. Probability pie
WB initial stage ≤ 1.64 mm
Altitude ≤ 826 m

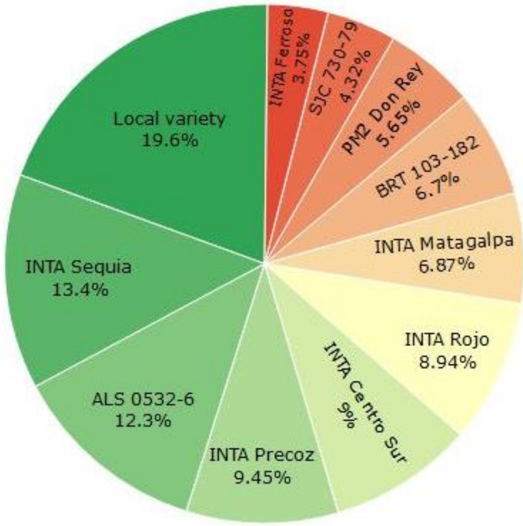


Figure 8. Probability pie
WB initial stage ≤ 1.64 mm
Altitude > 826m

When the water balance is above 1.64 mm during the initial growth stage, the TNmax during the initial growth stage becomes important. This results in the preference of the INTA Sequia (8) in two scenarios in which the temperature is either exceeding or not exceeding the TNmax of 18.38 °C (*Figure 10 and 11*).

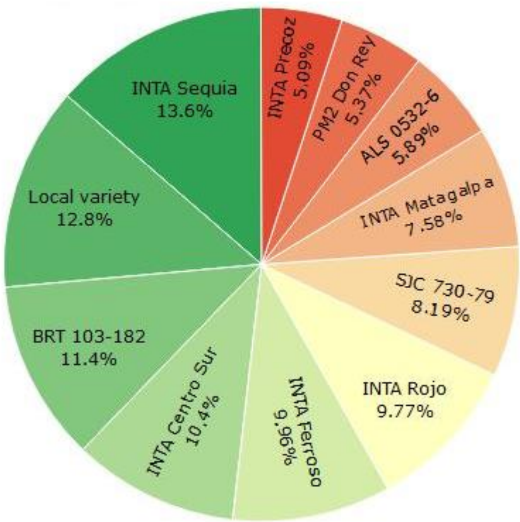


Figure 9. Probability pie
WB initial stage > 1.64 mm
TNmax initial stage ≤ 18.38 °C

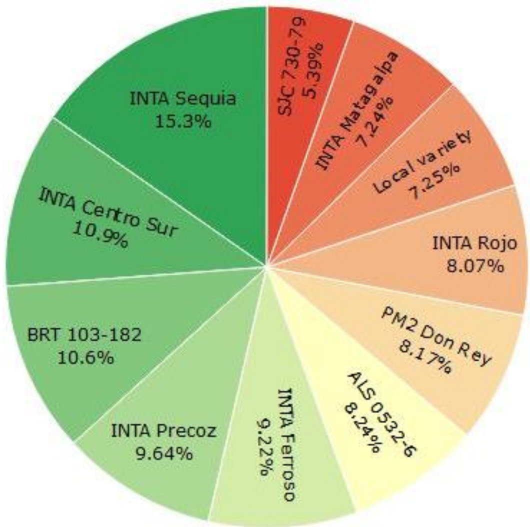


Figure 10. Probability pie
WB initial stage > 1.64 mm
TNmax initial stage > 18.38 °C

5.3.2 Covariates: spatial components, season and water balance

Excluding the TNmax in the Plackett-Luce resulted in a difference on the right side of the Plackett-Luce tree (*Figure 4* & *Figure 5*). Therefore, the probability pies resulting from *Figure 10* and *Figure 11* apply for *Figure 5* as well. In this case is the amount on the water balance during the crop development stage used to partition the data. This resulted in two scenarios, either not exceeding or exceeding 4.25 mm on the WB during the crop development stage (*Figure 12 and 13*). In both scenarios is the INTA Sequia (8) most preferred and are the probabilities of *Figure 12* comparable to *Figure 11* and *Figure 13* is comparable to *Figure 10*.

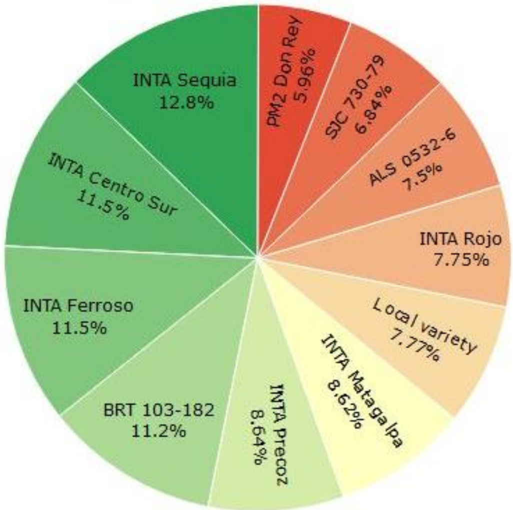


Figure 11. Probability pie
WB initial stage > 1.64 mm
WC crop development stage ≤ 4.25 mm

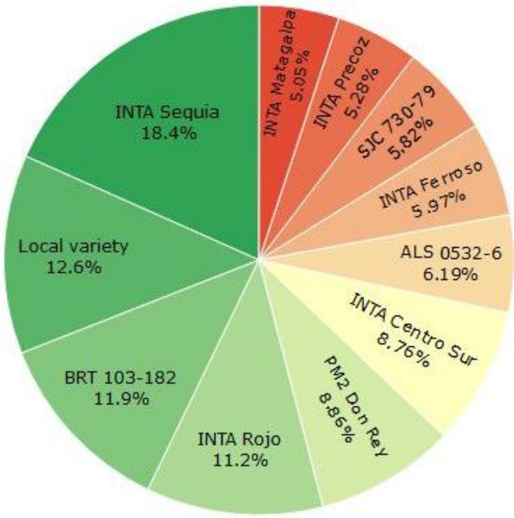


Figure 12. Probability pie
WB initial stage > 1.64 mm
WC crop development stage > 4.25 mm

5.3.3 Covariates: spatial components

Combinations of the environmental conditions pertinent to the tree of *Figure 6* divide the study area in four regions as shown in *Figure 13* based on altitude and soil types. Region 1 corresponds to node 2, region 2 to node 5, region 3 to node 6 and region 4 to node 7 from the Plackett-Luce tree *(Figure 5)*.

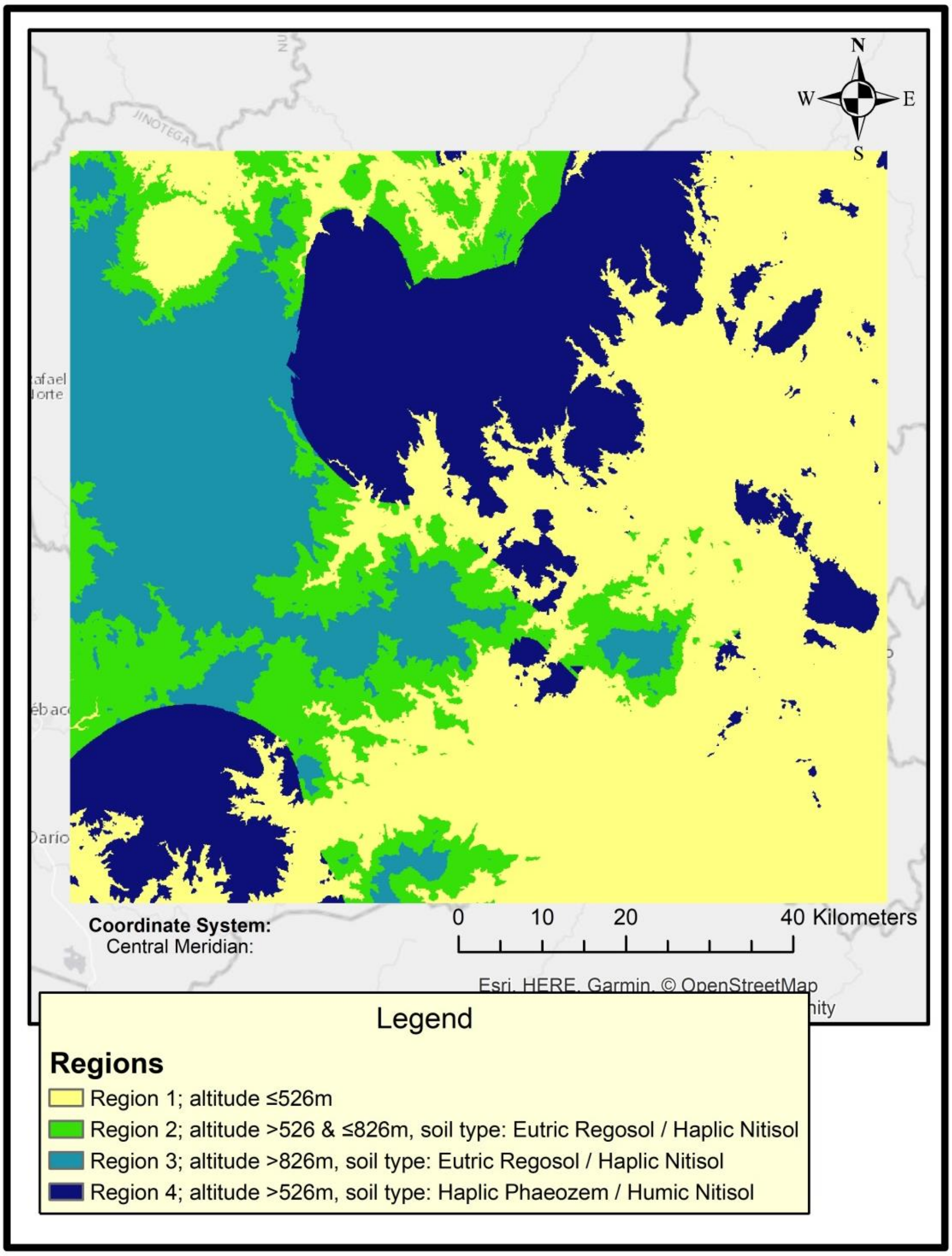


Figure 13. Physiographic regions relevant to the PLtree of Figure 6.

Figures 14 - 17 represent the probabilities of the preferred varieties for the different regions as distinguished in *Figure 14*. The local variety (11) is most preferred with the soil types Eutric Regosol or Haplic Nitisol, while the INTA Sequia (8) is most preferred below an altitude of 526 meter or for the combination altitude above 526 and a Haplic Phaeozem or Humic Nitisol.

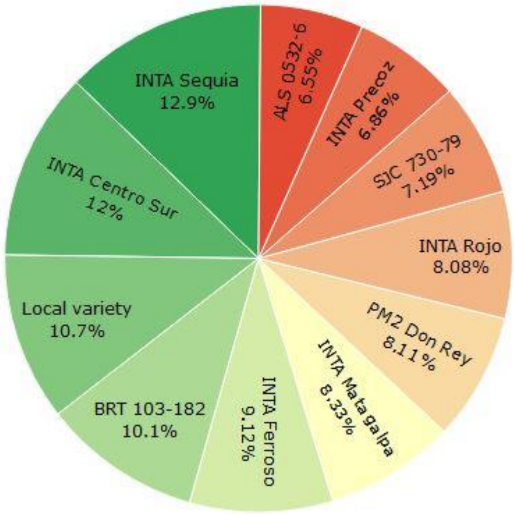


Figure 14. Probability pie
Altitude ≤ 526m

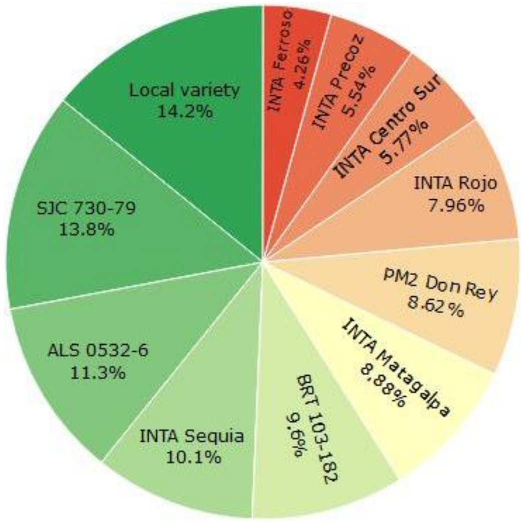


Figure 15. Probability pie
Altitude > 526 & ≤826 m
Soil type: Eutric Regosol / Haplic Nitisol

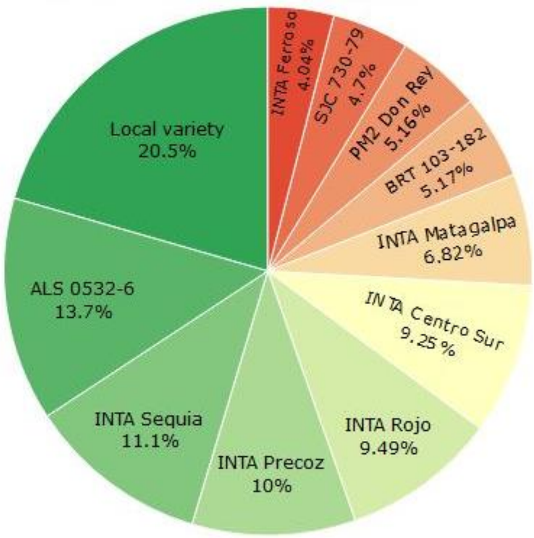


Figure 17. Probability pie
Altitude > 826 m
Soil type: Eutric Regosol / Haplic Nitisol

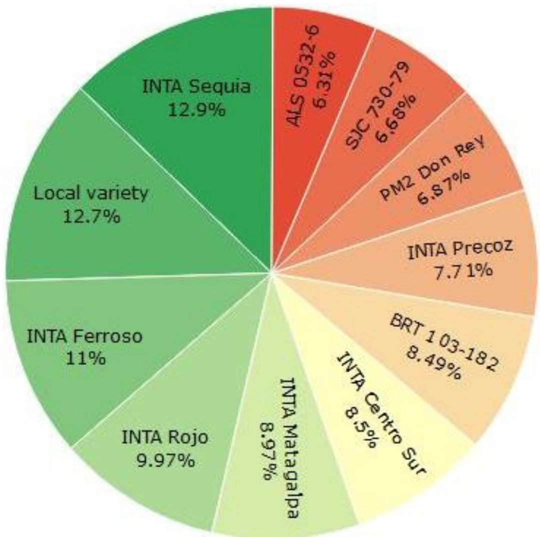


Figure 16. Probability pie
Altitude > 526 m
Soil type: Haplic Phaeozem / Humic Nitisol

5.4 Summary results

Table 7 shows the summarized top three preferred varieties per given set of environmental conditions per model.

Table 7. Overview top three preferred varieties per given set of environmental conditions per model

Covariates chosen to partition data			Top 3 probability a variety is preferred over others		
1	**Covariate 1**	**Covariate 2**	**1st**	**2nd**	**3rd**
Chapter 5.1.1 (Figure 3)	Initial (mm) ≤ 1,64	Altitude (m) ≤ 826	Local variety 11	INTA Matagalpa 5	SJC 730-79 10
	Initial (mm) ≤ 1,64	Altitude (m) > 826	Local variety 11	INTA Sequia 8	ALS 0532-6 1
	Initial (mm) > 1,64	TNmax_ini (°C) ≤ 18.38	INTA Sequia 8	Local variety 11	BRT 103-182 2
	Initial (mm) > 1.64	TNmax_ini (°C) > 18.38	INTA Sequia 8	INTA Centro Sur 3	BRT 103-182 2
2	**Covariate 1**	**Covariate 2**	**1st**	**2nd**	**3rd**
Chapter 5.1.2 (Figure 4)	Initial (mm) ≤ 1,64	Altitude (m) ≤ 826	Local variety 11	INTA Matagalpa 5	SJC 730-79 10
	Initial (mm) ≤ 1,64	Altitude (m) > 826	Local variety 11	INTA Sequia 8	ALS 0532-6 1
	Initial (mm) > 1,64	CropDev (mm) ≤ 4,25	INTA Sequia 8	INTA Centro Sur 3	INTA Ferroso 4
	Initial (mm) > 1.64	CropDev (mm) > 4,25	INTA Sequia 8	Local variety 11	BRT 103-182 2
3	**Covariate 1**	**Covariate 2**	**1st**	**2nd**	**3rd**
Chapter 5.1.3 (Figure 5)	Altitude (m) ≤ 525		INTA Sequia 8	INTA Centro Sur 3	Local variety 11
	Altitude (m) > 526 & ≤ 826	Soil Type Eutric Regosol / Haplic Nitisol	Local variety 11	SJC 730-79 10	ALS 0532-6 1
	Altitude (m) > 526	Soil Type Eutric Regosol / Haplic Nitisol	Local variety 11	ALS 0532-6 1	INTA Sequia 8
	Altitude (m) > 526	Soil Type Haplic Phaeozem / Humic Nitisol	INTA Sequia 8	Local variety 11	INTA Ferroso 4

6. Discussion

6.1 Analysis of Partial Ranked Variety Scores

The data used in this thesis were results from the Seeds for Needs experiment in Nicaragua and consisted of relative evaluations about the performances of varieties by farmers. Triplets were chosen at random, out of a total of ten varieties and then compared to each other and also an available local variety. This evaluation was based on the overall performance of the variety, with the overall performance being a sum of the yield, consumption quality and marketability (*Bioversity, 2016*). It is unknown how the scores of the partial performances are combined in an overall performance keeping in mind the subjectivity of the farmers. The partial relative performances can be of additional value when it comes to explaining inexplicable variety performances under a given set of environmental conditions.

The Bradley-Terry (BT) model and Plackett-Luce (PL) model are both suitable for doing statistical analysis with incomplete rankings. The PL model is chosen, because the model is designed to deal with observations independently, using more partial rankings per observation. Furthermore, the PL model is insensitive to inconsistencies the partial ranking inside an observation might have. Model-based recursive partitioning is used on all covariates corresponding to the locations of the observations in the experiment. The parameters used in the model were an alpha of 0.01 and a max depth of 5, which are in line with *Turner (2018)*. The minimal amount of observations per leaf in the tree is set at 100 (10% of the total amount of observations) but is stated to be optimal at a value of 200 in *Turner (2018)*. The influence of using a min size of 200 in the trees of this thesis, resulted in a model only using the covariates of the water balance during the initial stage and the altitude. The TNmax during the initial growth stage and the water balance of the crop development stage, as used to partition the data in *Figure 3 and 4,* are not explanatory when the minimum subsample size is increased to 200. When considering the spatial components only, as in *Figure 5*, the soil type is no longer considered to be an explanatory variable.

The PL model is still relatively new in R, lacking in sufficient documentation, especially in the starting period of this thesis. Therefore, the validation is a variation on the k-fold cross validation method (*Kohavi, 1995*) and based on conversations with Jacob van Etten (*personal communication, April 13, 2018*) and Kaue de Sousa (*personal communication, April 25, 2018*). Instead of the leave-one-out method, a season is left out in the cross-validation, resulting in a blocked-cross-validation in order to analyse the goodness-of-fit of the model. The dataset had three different seasons over different years, resulting in five different seasons and thus five subsets of data. The deviances can be calculated over the season left out and divided by the square root of the number of observations in that leaf, in order to reduce the bias, resulting from the folding in seasons (*Kaue de Sousa, personal communication, April 25, 2018*). Furthermore, a forward selection is done during the calculations of the deviances, meaning the covariates as chosen in the PL model, are input to fit a new PlL model to the four seasons. The results of this validation are given in negative log-likelihoods and the lower this value, the better the model is when it comes to the goodness-of-fit. The PL model in which all covariates were used as input, has a negative log-likelihood of 1009, which is lower than the value for including only the spatial components. It was expected that the use of more covariates than only the spatial, but less than all covariates would lead to a negative log-likelihood between 1009 and 1029. The negative log-likelihood for using the water balance and the spatial components is however 1029, which is close to, but higher than if using spatial components only. The reason for this might be that using the spatial components only results in two explanatory values (soil type and altitude) (*Figure 5*), compared to three explanatory values (water balance in the first two growth stages and altitude) (*Figure 4*). Another reason can be the response of the varieties during the different seasons, since seasonality can also be seen as explanatory variable (*Kaue de Sousa, personal communication, April 25, 2018*).

6.2 Environmental data

A selection of environmental data used as covariates available and accessible with relevance to the ecology of beans for the study area were used: altitude, slope, season, soil type, night-temperature and water balance. The last two covariates were calculated per growth stage. The different growth stages were the initial stage (20 days), crop development stage (30 days), the mid-season stage (40 days) and the late season stage (20 days) (*Brouwer, 1986*). It was expected that the conditions during the initial growth stage would be influencing the preference of the bean varieties, based on the findings for the common bean in Nicaragua by *Gómez (2004)*. Based on the PL model (*Figure 3 and 4*), it can be stated the water balance during the initial growth stage is the most important.

Since this dataset has been used for statistical analysis before, there was some insight towards important environmental conditions regarding the Seeds for Needs dataset in Nicaragua. The precipitation was considered to be important, because this resulted from research in which soil information and extreme precipitation indices were used as covariates (*Bioversity, 2016*). This thesis did not use the number of dry days or the maximum precipitation but rather a water balance per growth stage. This water balance was computed from the comparison between the evapotranspiration a plant would use during a certain growth stage and the water which was available through precipitation in the same growth stage. The evapotranspiration was calculated with the relatively simple Blaney-Criddle (*Smith, 1998*) method. It was expected, the water balance in the initial stage would be an important covariate because the yield of beans is sensitive to water deficit during the initial growth stage and the crop development stage (*Steduto, 2012*). The water balance in the initial growth stage is positively correlated to the water balance in the crop development (0.46), mid-season (0.71) and late season (0.66) stage (*Figure 1*). For the water balance in the crop development stage the strongest correlation can be found with the mid-season stage (0.75) and positive correlations between the late season stage (0.33) (*Figure 1*). Dividing the water balance in different growth stages is advisable because the correlations between the different stages are varying. Each water balance of a growth stage might have a different impact on the performances of different varieties. Since the water balance is such an important factor in the partitioning of the dataset, especially for the initial stage, it might be of additional value to calculate the evapotranspiration with a more accurate method such as Hargreaves (*Kra, 2013*), since the Blaney-Criddle method underestimates the evapotranspiration in general (*Subedi, 2015*).

Furthermore, the maximum night-temperature was assumed to be an important environmental condition for comparing the preferences of the different common bean varieties *Turner (2018)*. In *Turner (2018)* the TNmax was only used as an input for the whole vegetative cycle, but in this model the TNmax was additionally calculated per growth stage. In *Figure 3* the second partitioning is done using the TNmax during the initial growth stage. The additional value of using the TNmax per growth stage over using the TNmax for the whole vegetative cycle is not considered to be important, since the different TNmax values are closely correlated to each other (0.88 - 0.99) (*Figure 1*). The TNmax from the initial growth stage is correlates strongly to the maximum night-time temperature of the crop development stage (0.97), followed by the mid-season stage (0.93) and the late season stage (0.85). The TNmax, which was chosen in the PL tree, was from the initial stage and it can be concluded the initial stage has most impact on the further development and thus the overall performance of the varieties. The maximum night-time temperatures are positively correlated with the water balance in the late season, varying from 0.4 – 0.44. The correlations between the night-time temperatures and the water balance covariates during the other growth stages are varying from 0 – 0.33.

The covariate altitude is not strongly correlated to any other covariate. The highest correlation for altitude is a negative correlation of -0.41, with the water balance in the initial growth stage. There is a difference in height from East to West (high to low) (*Figure 6*). *Trapnell (1960)* mentioned the significant influence of altitude on rainfall in an area and based on the correlations in *Figure 1* higher altitudes might correspond to lower precipitation amounts during the initial growth stage. Additionally, the difference in height was also highlighted by Kaue de Sousa (*personal communication, April 25, 2018*), because there was a climatic gradient, which made it of additional use to include the longitude and latitude as covariates as well (*Haverkort, 1990*). The gradient is now only captured when the altitude is an explanatory variable, but the influence of including the longitude and latitude on the model outcomes is unknown.

When it comes to the interaction of the crops with the soil there are shortcomings in this thesis, that might influence the preferences of the varieties. The precipitation is assumed to be effective precipitation, meaning all water becomes available for the crops, but deep percolation and run-off were ignored. The soil type is known, but knowledge about the soil depth and the ability of the soil to infiltrate water is not included. Especially in combination with unexpected (short) tropical storms with large amounts of precipitation (*Gómez and Blair, 2004*). The amount of precipitation might be very high, resulting in a positive water balance, but is not necessarily distributed over the growth period. It is not known whether the soil is able to infiltrate some of the water or that all water is lost due to surface-runoff in combination with steep slopes. The ability of the soil to keep water is important, because for example during the *Postrera* the precipitation is low and humidity very high, yet the soil contains enough moisture to provide for the crops (*personal communication, April 25, 2018*).

6.3 Variety Scores

Based on the results, it can be concluded the INTA Sequia (8) and the local variety (11) of the farmer are the best performing varieties, because either one of them is the most preferred variety per set of given environmental conditions (*Table 7*). One should keep in mind the local variety is not one and the same variety for each location, but differs from farmer to farmer. The local variety is best seen as reference variety and not as a free-standing variety to prevent biased comparisons. For 50% of the given set of variables is the local variety preferred over other varieties or in other words: is none of the other varieties outperforming the local variety of the farmer. For the other 50% of the given set of variables is the INTA Sequia (8) preferred over the other varieties. The INTA Precoz (6), INTA Rojo (7) and PM2 Don Rey (9) are the only varieties that do not occur in the top three preferred varieties under any given set of varieties in this thesis.

Bioversity (2016) concluded the following, based on a Bradley-Terry model with soil information and extreme precipitation indices as covariates: The INTA Sequia (8) had the best overall performance when there was low precipitation (less than 57 mm) during the growing period. The local variety (11) was not clearly outperformed, even when the number of consecutive dry days was greater than 15 days. When there was more than 57 mm of precipitation during the growing period, the INTA Sequia (8) had a similar performance to the other varieties. The INTA Matagalpa (5) had the best performance with a precipitation higher than 57 mm per growing period.

This thesis also concludes the best overall performance is of the INTA Sequia (8) and the local variety (11) under all the given sets of environmental data. It is interesting to notice the INTA Sequia (8), being the drought-resistant variety, is outperforming the other varieties in *Figure 9, 10, 11 and 12*, because there is more than 1.64 mm surplus on the WB in the initial growth stage and in *Figure 12* also more than a surplus of 4.25 mm on the WB during the second growth stage.

The performance of the INTA Matagalpa (5) was worst when the surplus on the water balance exceeded the 4.25 mm during the crop development stage (*Figure 12*). As resulting from this thesis is the INTA Matagalpa (5) able to deal with a surplus of water during the initial stage, but during the crop development stage, this surplus should not exceed 4.25 mm. The INTA Matagalpa was preferred as second best variety as long as the WB during the initial stage is not exceeding the 1.64 mm in combination with an altitude below 826 meters (*Figure 7*).

The results of the Plackett-Luce model (*Turner, 2018*) were based on the use of the season, year and maximum night-temperature as covariates to partition the data. The findings of *Turner (2018)* mentioned the preference of the INTA Rojo (7) in the early season (Primera) and where night-time temperatures were not exceeding 18.7 °C, closely followed by the local variety (11). During the other growing seasons (Postrera and Apante) and night-time temperatures lower than 18.7 °C the local variety (11) was most preferred, followed by the INTA Sequia (8). When the night-time temperature was exceeding the 18.7 °C, the INTA Sequia (8) was still preferred, closely followed by the BRT 103-182 (2) and INTA Centro Sur (3).

Most findings in this thesis are in line with the conclusions from *Turner (2018)*. The findings concerning the INTA Rojo (7) are not in line with this research, because the INTA Rojo (7) is not mentioned once in the top three preferences of the varieties. Regardless the night-time temperatures in either the findings of *Turner (2018)* or in this thesis is the INTA Sequia (8) most preferred. In *Figure 9 and 10* is the top three almost similar to the findings of *Turner (2018)*.

The common bean adapts differently to environmental characteristics in Nicaragua (*Gómez and Blair, 2004*). The characteristics are for the common bean in general in Nicaragua and were divided into different zones: optimal, intermediate and marginal. The altitude, temperature, precipitation, soil-depth, the slope, the drainage, pH and aluminium content in the soil were the characteristics important concerning the adaptation of the bean to its environment. This thesis only covers the altitude, temperature, precipitation and the slope as covariates for the input of the Plackett-Luce model. In all the Plackett-Luce trees, the altitude was an important covariate. The point in partitioning this covariate was most of the times 826 meters, and for the spatial variation tree (*Figure 5*) there was a second threshold around 526 meters. The optimal zone is between 450-800 meters (*Gómez and Blair, 2004*) , which is in line with the findings of this thesis. The temperature is considered to be optimal between 17 and 24°C. It is harder to check whether these temperatures are in line with this research, while the maximum night-time temperature is used. The threshold is 18.38°C, so the night-time temperatures are going below the 17°C. The daytime temperatures are generally higher than the night-time temperatures and included in the water balance, but not separately analysed. *Gómez (2004)* mentioned a lower market value (and thus lower overall performance) could be the result of late rains, but this could not be found based on the Plackett-Luce models. The slope was not chosen as covariate, and not considered to be instable enough, compared to the other covariates.

6.4 (Geo-)Graphical presentation

The default visualization of the PL model is a tree with line-graphs representing the leaves. The line-graphs represent the worth parameters of the different varieties and are connected with a line. This line has no meaning, because the ordering of the varieties on the x-axis is alphabetically fixed and a different order would result in a complete different line. Furthermore, a base-line is visible in the different graphs, but this line has no meaning, since it is not corresponding to the reference variety. To present the data in a rational way, the worth parameters were rescaled until the summation was 1 and are now probabilities the varieties are preferred over other varieties under the given set of environmental conditions. Visualising the probabilities can be done with either a bar graph or a pie chart (*Friendly, 2000*). A pie chart is considered to be most sufficient because the probabilities are adding to one.

7. Conclusion

The goal of this research was to contribute to enhanced statistical analysis of on-farm trials of the Seeds for Needs experiment in Nicaragua. A spatial analysis was part of the result after combining existing methods with environmental data to highlight spatial variation.

7.1 Suitable statistical method – incomplete ranking

The methods suitable for the statistical analysis of ranked sets of variety trials in combination with explanatory environmental variables, are the Bradley-Terry (BT) model and the Plackett-Luce (PL) model. The Bradley-Terry model uses paired comparisons, while the Plackett-Luce model uses partial rankings of multiple items per observation. The Plackett-Luce model was preferred over the Bradley-Terry model because the Seeds for Needs data has scored three varieties along with a local variety in each experiment. BT would rely on an invalid independence assumption, since one paired comparison is seen as one observation, while having six paired comparisons per observation. PL is able to group several rankings per observation and is insensitive to inconsistency between these rankings.

7.2 Selection environmental conditions

The environmental conditions available for the study area and which were considered to be influential to crop performance were: altitude, slope, soil type, season, planting date, water balance (WB) per growth stage and maximum night-temperature per growth stage (TNmax). The environmental conditions were assigned to the locations of the observations and used as possible explanatory variables in the PL model in R. The variables most explanatory for different performances of the varieties, were the water balance during the initial growth stage, the altitude and TNmax during the initial growth stage. The selection of these three covariates resulted in the lowest negative log likelihood and therefore in the model with the highest goodness-of-fit. When only including spatial varying environmental conditions, the altitude and soil type were considered to be most explanatory.

7.3 Relative scores varieties

Based on the explanatory variables resulting from the PL-tree, the study area and observations could be divided in either regions or scenarios with different rankings of preferences for varieties. The probabilities of varieties being preferred over other varieties per region and/or scenario are visualized in pie charts.

Based on the three most important explanatory variables (WB initial, altitude and TNmax initial), the study area could be divided into two different regions based on altitude for a WB below a threshold of 1.64mm. In both regions the local variety (11) of the farmer was the most preferred over all other varieties. A WB exceeding the 1.64mm during the initial growth stage resulted in two scenarios in which the INTA Sequia (8) was preferred in both scenarios.

The study area could be divided into four regions, when using the spatial explanatory variables only. For altitudes below 526 meters, the INTA Sequia (8) was most preferred, regardless the soil type. The INTA Sequia (8) was also preferred with an altitude above 526 meters and a soil type of either Haplic Phaeozem or Humic Nitisol. In the regions with soil types Eutric Regosol or a Haplic Nitisol and an altitude above 526 meters was the local variety most preferred.

Except for the INTA Precoz (6), INTA Rojo (7) and PM2 Don Rey (9) all varieties occur at least once in a top three of preferred varieties under a given set of environmental conditions. The first places in the top three under a given set of environmental conditions in this thesis is always either the INTA Sequia (8) or the local variety (11) of the farmer.

7.4 (Geo-)graphical representation

Visualising the results of the PL tree with the attached plot method in R might be a sufficient way of explaining results, but this lacks the visual representation of the spatial variation. Each leaf at the end of the PL tree consists of a line-graph in which the varieties are alphabetically ordered on the x-axis and plotted against the worth parameters of the varieties. The worth parameter of each variety is connected to the next variety with a line, but due to the alphabetically ordered varieties this does not explain anything. Another way of visualising this spatial variation, is by converting the worth parameters of the varieties compared to a reference crops, into the probability a variety is preferred over other varieties. These probabilities are basically rescaled worth parameters summoned to 1 and can be visualised in a pie-chart, corresponding to a certain region or scenario with environmental conditions.

7.5 Recommendation

The main question of this research was, 'Which variety performs best under given environmental conditions?'. Through the statistical analysis shown throughout this research, it was shown that a varieties performance depends largely on the discussed environmental conditions and thus an overall 'best' performing variety was not found. Instead, a 'best' performing variety resulted from each set of environmental variables. However, the INTA Sequia and local variety of the farmer are performing very well under many chosen sets of environmental conditions. The environmental conditions which are important explanatory variables regarding the preference of the varieties, are considered to be the water balance and the maximum night-temperature during the initial stage and the altitude.

The question now posed is 'what and how can these findings be applied and implemented into real-life scenarios?'. One suggestion could be providing in-depth feedback for the farmers with the aim to improve their knowledge concerning other varieties than their local variety. When other varieties besides their own variety are preferred under given environmental conditions, the performance of the variety is proven to the farmer. The results of the Seeds for Needs experiment though, cannot only be feedback for the farmers participating in the experiment, but also for their neighbours. Farmers tend to adopt other agricultural technologies when the technology is proven in an area with similar (environmental) conditions (*Muthoni, 2017*). This thesis contributes to a potential extrapolation of the findings, by finding the regions spatially similar when it comes to the environmental conditions. The spatial environmental conditions as found can be a starting point for trying to explain the spatial variation of the preferences for different varieties. The methodology to include model-based partitioning on covariates variable over time and space as done in this thesis, resulted in significantly different rankings of preferences of varieties. The methodology can be expanded to other datasets resulting from the Seeds for Needs experiments of Bioversity (*Bioversity, 2016*).

Furthermore, the findings of this thesis can contribute in agro-climate risk management, from which varietal diversification is a way to decrease yield reduction as result of climate change (*Sukcharoen and Leatham, 2016*). Next to the spatially explicit regions different scenarios are created, focussing on the preference of varieties above or below a certain precipitation and temperature threshold. Precipitation and temperatures are highly variable over space and time, resulting in uncertainty when it comes to the performance of crops. Some crops are more resilient to certain circumstances compared to others but are not necessarily the best for all circumstances. In this case varietal diversification is a possibility, in which a set of varieties is chosen to reduce the climatic risks for a farmer (*Sukcharoen and Leatham, 2016*). The uncertainty of using only one variety or a certain set of varieties can be compared to different scenarios and analysed. Variety Release Committees and/or governments can advise a certain variety or set of varieties to reduce the risks on the existing yield because of climate change effects (*Lemke, 2007*) or to reduce yield failures due to the different interaction of varieties with the environment (*Sukcharoen and Leatham, 2016*).

Finally, this thesis can be seen as valuable for breeders and plant pathologists, since a unique combination of field data and the performance of the designed breeds is analysed. The analysis is based on the performance as evaluated by the farmers themselves and give an idea of the performance of the different varieties when cultivated by farmers rather than under designed, controlled, laboratory-like conditions.

References

Adams, R. M., Rosenzweig, C., Peart, R. M., Ritchie, J. T., McCarl, B. A., Glyer, J. D., ... & Allen, L. H. (1990). Global climate change and US agriculture. Nature, 345(6272), 219-224.

Agresti, A., & Kateri, M. (2011). Categorical data analysis. In International encyclopedia of statistical science (pp. 206-208). Springer Berlin Heidelberg.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300(9), D05109.

Atlin, G. N., Cooper, M., & Bjørnstad, Å. (2001). A comparison of formal and participatory breeding approaches using selection theory. *Euphytica*, *122*(3), 463-475.

Audsley, E., Pearn, K. R., Simota, C., Cojocaru, G., Koutsidou, E., Rounsevell, M. D. A., ... & Alexandrov, V. (2006). What can scenario modelling tell us about future European scale agricultural land use, and what not?. *Environmental science & policy*, *9*(2), 148-162.

Bentley, J. W. (1989). What farmers don't know can't help them: The strengths and weaknesses of indigenous technical knowledge in Honduras. Agriculture and Human Values, 6(3), 25-31.

Beza, E., Steinke, J., Van Etten, J., Reidsma, P., Fadda, C., Mittra, S., ... & Kooistra, L. (2017). What are the prospects for citizen science in agriculture? Evidence from three continents on motivation and mobile telephone use of resource-poor farmers. PloS one, 12(5), e0175700.

Bioversity (2016) Seeds for Needs – Emerging evidence on triadic comparisons of technologies (tricot). Internal report Bioversity international.

Breiman, L. (2017). *Classification and regression trees*. Routledge.

Brouwer, C., & Heibloem, M. (1986). Irrigation water management: irrigation water needs. *Training manual*, *3*.

Caron, F., Teh, Y. W., & Murphy, T. B. (2014). Bayesian nonparametric Plackett–Luce models for the analysis of preferences for college degree programmes. *The Annals of Applied Statistics*, 1145-1181.

Crawford, S. L. (1989). Extensions to the CART algorithm. International Journal of Man-Machine Studies, 31(2), 197-217.

Fielding, W. J., & Riley, J. (1997). How big should on-farm trials be and how many plots should be measured. PLA Notes, 29, 19-22.

Firth, D., & Turner, H. L. (2012). Bradley-Terry models in R: the BradleyTerry2 package. Journal of Statistical Software, 48(9).

Fischer, G., Shah, M., Tubiello, F. N., & Van Velhuizen, H. (2005). Socio-economic and climate change impacts on agriculture: an integrated assessment, 1990–2080. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 360(1463), 2067-2083.

Foy, C. D. (1992). Soil chemical factors limiting plant root growth. In *Limitations to plant root growth* (pp. 97-149). Springer, New York, NY.

Friendly, M., & SAS Institute. (2000). *Visualizing categorical data*. Cary, NC: Sas Institute.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data*, *2*, 150066.

Garge, N. R., Bobashev, G., & Eggleston, B. (2013). Random forest methodology for model-based recursive partitioning: the mobForest package for R. *BMC bioinformatics*, *14*(1), 125.

Gómez, O. (2004). *Evaluation of Nicaraguan common bean (Phaseolus vulgaris L.) landraces* (Vol. 476).

Gomez, O. J., Blair, M. W., Frankow-Lindberg, B. E., & Gullberg, U. (2004). Molecular and phenotypic diversity of common bean landraces from Nicaragua. *Crop Science*, *44*(4), 1412-1418.

Guiver, J., & Snelson, E. (2009, June). Bayesian inference for Plackett-Luce ranking models. In proceedings of the 26th annual international conference on machine learning (pp. 377-384). ACM.

Hatfield, J. L., & Prueger, J. H. (2015). Temperature extremes: effect on plant growth and development. *Weather and climate extremes*, *10*, 4-10.

Haverkort, A. J. (1990). Ecology of potato cropping systems in relation to latitude and altitude. *Agricultural Systems*, *32*(3), 251-272.

Hunter, D. R. (2004). MM algorithms for generalized Bradley-Terry models. Annals of Statistics, 384-406.

Kohavi, R. (1995, August). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*(Vol. 14, No. 2, pp. 1137-1145).

Kra, E. Y. (2013). Hargreaves equation as an all-season simulator of daily FAO-56 Penman-Monteith ETo. *Agricultural Science*, *1*(2), 43-52.

Lemke, P., Ren, R., & Alley, I. (2007). The physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Clim. Change, 2007, 337-383.

Letey, J. O. H. N. (1958). Relationship between soil physical properties and crop production. In *Advances in soil science* (pp. 277-294). Springer, New York, NY.

Lobell, D. B., Burke, M. B., Tebaldi, C., Mastrandrea, M. D., Falcon, W. P., & Naylor, R. L. (2008). Prioritizing climate change adaptation needs for food security in 2030. Science, 319(5863), 607-610.

Luzzi, D. (2016). Communicating spatial uncertainty to non-experts using R. Thesis Report for Laboratory of Geo-Information Science and Remote Sensing, Wageningen University and Research Centre.

Muthoni, F. K., Guo, Z., Bekunda, M., Sseguya, H., Kizito, F., Baijukya, F., & Hoeschle-Zeledon, I. (2017). Sustainable recommendation domains for scaling agricultural technologies in Tanzania. *Land Use Policy*, *66*, 34-48.

Mwongera, C., Boyard-Micheau, J., Baron, C., & Leclerc, C. (2014). Social process of adaptation to environmental changes: How eastern African societies intervene between crops and climate. Weather, Climate, and Society, 6(3), 341-353.

Patterson, D. T. (1995). Effects of environmental stress on weed/crop interactions. Weed Science, 43, 483-483.

Silva, E. L. D. G. S., Oliveira, C. M. M., Mendes, A. B., & Guerra, H. M. G. F. O. (2017). Technology and Innovation in Agriculture: The Azores Case Study. *International Journal of Interactive Mobile Technologies (iJIM)*, *11*(5), 56-66.

Silva, E. L. D. G. S., Oliveira, C. M. M., Mendes, A. B., & Guerra, H. M. G. F. O. (2017). Technology and Innovation in Agriculture: The Azores Case Study. International Journal of Interactive Mobile Technologies (iJIM), 11(5), 56-66.

Smith, M., Allen, R., & Pereira, L. (1998). REVISED FAO METHODOLOGY XA9848174 FOR CROP-WATER REQUIREMENTS. *Management of nutrients and water in rainfed arid and semi-arid areas*, 51.

Steduto, P., Hsiao, T. C., Fereres, E., & Raes, D. (2012). *Crop yield response to water* (Vol. 1028). Rome: FAO.

Steinke, J. (2015). Citizen science with resource-poor farmers as a new approach to climate adaption and food security: Evidence from Honduras.

Steinke, J., van Etten, J., & Zelan, P. M. (2017). The accuracy of farmer-generated data in an agricultural citizen science methodology. Agronomy for Sustainable Development, 37(4), 32.

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychological methods, 14(4), 323.

Subedi, A., & Chávez, J. L. (2015). Crop evapotranspiration (ET) estimation models: a review and discussion of the applicability and limitations of ET methods. *Journal of Agricultural Science*, *7*(6), 50.

Sukcharoen, K., & Leatham, D. (2016). Mean-Variance Versus Mean–Expected Shortfall Models: An Application To Wheat Variety Selection. *Journal of Agricultural and Applied Economics*, *48*(2), 148-172.

Todorovic, M. (2005). Crop water requirements. *Water Encyclopedia*.

Trapnell, C. G., & Griffiths, J. F. (1960). The rainfall-altitude relation and its ecological significance in Kenya. *The East African Agricultural Journal*, *25*(4), 207-213.

Turner. H., (2018) PlackettLuce, R Package for Plackett Luce models under development, GitHub repository, <https://github.com/hturner/PlackettLuce>

van de Gevel, J., Gellaw, A., & Fadda, C. (2013). Seeds for needs: adaptation to climate change-Innovative tools to match seeds to the needs of women farmers in Ethiopia.

van Etten, J. (2011). Crowdsourcing crop improvement in sub-saharan africa: a proposal for a scalable and inclusive approach to food security. IDS bulletin, 42(4), 102-110.

Van Etten, J., Beza, E., Calderer, L., Van Duijvendijk, K., Fadda, C., Fantahun, B., ... & Kiambi, D. (2016). First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). Experimental Agriculture, 1-22.

Appendix I : Additional data sources used

Variable : Altitude	
Dataset	SRTM Digital Elevation Data produced by NASA (SRTM_19_10)
Acquired	http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1
Resolution	90 meter
Variable : Temperature	
Dataset	Land Surface Temperature (MYD11A2)
Acquired	https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table
Resolution	1000 meter
Temporal granularity	Composites – 8 days 110 days starting from the planting date Note: The needed data is from 10/09/2015 – 29/04/2017, but due to the composites of 8 days, the downloaded data is from 06/09/2015 - 30/04/2017. The dataset was provided by Kaue de Sousa (consultant of Bioversity).
Variable : Precipitation	
Dataset	CHIRPS (Climate Hazards Infrared Precipitation with Stations (*Funk, 2015*)
Acquired	http://chg.geog.ucsb.edu/data/index.html [CHIRPS > Data > global_daily > tifs > p05 > year]
Resolution	0.05°
Temporal granularity	Daily (10/09/2015 – 29/04/2017)