# Quantifying the effect of ecological restoration on soil development

Constructing a SOC chronosequence in the Baviaanskloof, South Africa

*MSc thesis*

**Jip Koster. Registration number: 951009470060**

**Supervisors: Titia Mulder (Soil Geography and Landscape, Wageningen University and Research), Louise Willemen (Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente)**

**March 2019**

# Abstract

Ecological restoration has the potential to tackle major global problems, but knowledge to predict its effects are currently limited due to a lack of available multidisciplinary monitoring methods, such as mapping soil organic carbon. Therefore, this research aims at constructing a method to quantify soil development as a result of ecological restoration. This is done in the semi-arid Baviaanskloof catchment, South Africa. This area offered the possibility to derive a 13-year chronosequence using 50 soil samples, which were analyzed on SOC. Further analyses of these samples, local legacy data and remote sensing derivatives were used to 1) investigate the main drivers of soil development in this study area by applying a principal component analyses and investigating the effects of the single soil forming factors, 2) assess the suitability of a mechanistic model approach to predict the spatial distribution of the observed SOC fractions and thereby the current soil development state, 3) assess the potential to predict the temporal development of the soil state as a result of ecological restoration, and 4) assess the suitability of established carbon models for quantifying soil development (RothC and Carbon Benefit Project) by comparing them with the chronosequence findings.

This study shows that SOC was an appropriate indicator for quantifying soil development, in contrast to the NDVI. The most important drivers were parent material and livestock, while the contribution of erosion and climate to soil development were hard to quantify. Moreover, the mechanistic model was successful in the prediction of the measured distribution of SOC ($R^2$ = 0.67 and thereby improved the original model ($R^2$ = 0.41). Given these models limitations are addressed, the mechanistic approach has the potential to be applied for the monitoring of restoration practices. The use of a chronosequence approach for predicting temporal development resulted in an identification of divergence, an alternative state and the soil development state. Although these identifications involved large uncertainties, the contribution to the currently scarce knowledge on the behavior of soil development to ecological restoration is profound. Moreover, the existing carbon models turned out to overestimate and misjudge this temporal behavior of SOC, which means their limitations are critical for the prediction of SOC development. This study concludes that despite its complex behavior, the importance of soil forming processes in driving the ecological system, teaches us that a multidisciplinary method, such as applied in this research, are crucial in the application of restoration practices, monitoring and modelling scenarios

# Table of Contents

# 1. Introduction

Landscape restoration is increasingly used for reversing the global pattern of anthropogenically and climate change induced land degradation. By also including the local communities in the process of introducing vegetation, landscape restoration projects have the aim to return social, financial and natural capital, but also inspiration to the people (Ferwerda, 2015). These returns, at the end, are driven by the enhancement of ecosystem services including food supply, flood risk and carbon sequestration (Keesstra, et al., 2018)., which are described in the recent Sustainable Development Goals (SDG's)

The effectiveness of returning these ecosystem services as a result of the restoration projects is currently actively discussed and monitored (Chen and Duan, 2009). Especially in semi-arid areas, where ecological systems are highly fragile to small environmental disturbances. To successfully describe the system regeneration, one requires detailed knowledge on the change of system state over time, which should be acquired by active monitoring of restoration sites.

This active monitoring of the advances is most often done using Normalised Difference Vegetation Index (NDVI) as an indicator of vegetation cover (Smit, 2014, Zucca et al., 2015), but this might not be suitable for monitoring the soil state as a result of ecological restoration. In fact, the semi-arid ecological system this study focusses on, is very complex in terms of processes interacting. Soil development is thereby known to be driven by properties related to climate, parent material, topography, time and organisms (Jenny, 1941). Since vegetation cover is not included in all of the properties it is most likely not simultaneously behaving with soil functioning (Bullock et al., 2011; Chen et al., 2007). Therefore, an enhancement of vegetation cover can only be considered a mitigation measure for improving soil functioning instead of an indicator of the soil development. An approach is needed which is multidisciplinary to consider all single soil forming factors affecting the eventual development.

For this, the infiltration capacity is known to be an indicator for soil development in arid regions (Van Luijk et al., 2013). Because it relates to the process of erosion (Constantini, 2016), it is an important indicator of soil state in degraded semi-arid systems. At the same time, infiltration capacity is only an indicator of soil structure (Wang and Shao, 2013) and therefore does not tell us everything about soil development. Alternatively, soil organic matter (SOM) is known to influence water retention (Haynes and Naidu, 1998), soil fertility (Tiessen et al., 1994) and considered a product of mainly biological activity. Due to the fact it covers multiple disciplines, SOM is a better option as an indicator of the physical, biological and chemical state of a soil (Post and Kwon, 2000). Consequently, SOM is deemed useful for monitoring regeneration of degraded areas (Constantini, 2016). To strengthen this assumption, multiple examples of increases in SOM as a result of vegetation change can be found (Post and Kwon, 2000; Shourzan et al.,2005).

SOM is hard to measure and therefore usually approximated by using Soil Organic Carbon (SOC) (Read and Ridgell,1922). One of the methods for describing SOC is using soil sample measurements in combination with geostatistical techniques (Heuvelink and Webster, 2001). This approach is not only costly, but these statistical models also tend to exaggerate the interpolation error in comparison with the mechanistic method (Robinson and Metternicht, 2006). An alternative method involves the construction of a SOC chronosequence, which means dynamics can be set up based on the spatial difference in age. However, the chronosequence needs a large range of ages (i.e. long-term experiments) to be able to accurately describe the temporal development (Insam and Domsch, 1988). This data is most often not available and therefore the chronosequence method has been applied very scarcely in semi-arid regions. Other methods for quantifying SOC development are focusing on quantifying large-scale carbon sequestration as a result of land use change. They most often use the empirical relations with land use properties (GEF, 2010). However, these carbon models do not include important soil forming processes in their calculations. These processes are essential for monitoring restoration ecology (Viaud et al., 2010). Therefore, by lacking any implementation of these processes the established carbon models might not be suitable for the case of ecological restoration. Concluding on the above, most of the methods focusing on quantifying SOC, have serious limitations in their current application in the case of monitoring ecological restoration.

In contrast, remote Sensing (RS) is widely available and can be a time and cost-efficient tool for obtaining environmental data. Nowadays, it is also applied to find relationships between soil processes and environmental parameters, in time and space (Conant et al., 2011; Mulder et al., 2011). For example, SOM stocks are mapped by relating soil processes to environmental legacy data (Minasny and McBratney, 2006; Hendriks, 2018). Moreover, RS-sensing parameters can be used as input in dynamic models, such as Century (Brady and Weil, 2010), RothC (Coleman and Jenkinson, 1995) and Ecosys (Grant, 1995). These models are able to predict vegetation dynamics and ultimately quantifying SOC on timescales of decades up to centuries.

However, applying remote sensing has been only scarcely used in the past to assess the effectiveness of restoration measurements. This is caused by a lack of available knowledge of remote sensing techniques in the domain of ecology (Aplin, 2005). Moreover, to relate remote sensing images to measurements of SOC or soil development one needs a large amount of measurements, both in the spatial as well as temporal domain. Mainly due to financial constraints regarding these measurements, scientists tend to focus on monitoring vegetation activity (Chen et al., 2007; Chen and Duan, 2009; Liu et al., 2012), which is only partly indicating the development of soil. Eventually this results in a mismatch of expected restoration results with the observed state caused by this lack of systemic knowledge on soil development behavior (Bestelmeyer et al., 2013).

In this perspective, this research aims to construct a method for monitoring the effect of ecological restoration on soil development, mainly focusing on semi-arid areas. To achieve this, the following questions need to be answered:

- What are the main environmental drivers of soil development as a result of ecological restoration?
- Can the current soil development state of the system be found?
- Can we predict the temporal development of the soil state a result of ecological restoration by constructing a restoration curve?
- How do the methods in this research compare to other methods for quantifying soil development in the research area?

In the following sections the materials and methods used to answer the questions will be described (Chapter 2) and eventually its results presented (Chapter 3) and thoroughly discussed (Chapter 4 and 5).

# 2. Materials and Methods

## 2.1 Study Area

To solve the stated research questions, the study was conducted in a semi-arid area where the vegetation has been restored over a timescale where changes in ecosystem services could be detected. The Baviaanskloof, South Africa, is known to be such an area and is therefore a popular study site. It is even assigned as UNESCO world heritage site, due to its unique ecosystem diversity (Mills and Cowling, 2010). There, mainly the native *Spekboom* vegetation (*Portulacaria afra*) has been severely declined in the late 1980's due to grazing of sheep and goats (Jansen, 2008). This has led to frequent erosional events and a severe decline in overall soil development (Draajer, 2010).

Since the year 2001, the Baviaanskloof UNESCO world heritage site has been included in the Spekboom restoration program under the flag of the PRESENCE learning network. As a result of this around 65% of the area has been put under *Spekboom* restoration management. This means soil is restored by replanting of vegetation and partly excluding livestock (Blanksma, 2011). For this restoration purpose, the native original *Spekboom* vegetation is an effective species to use for restoration due to its strong resistance to drought and slope processes. Moreover, its well-known ability to sequestrate substantial amounts of organic carbon into the soil (Mills and Cowling, 2010), contributes to its suitability for restoration. Also, the fact the *Spekboom* does not reproduce via seeds, makes this species easy to monitor due its passive behavior. Another consequence of this lack of seed source is the requirement of *Spekboom* to be planted as cuttings (Mills and Cowling, 2010) of about 40 cm in size. This size dependents on the type of *Spekboom* planted which can differ between two subspecies (Mills and Cowling, 2010).

The *Spekboom* restoration management area (446 $km^2$) is part of the 1234 $km^2$ Baviaanskloof catchment. Despite focusing on a local scale, we can expect significant spatial differences in environmental variables. In fact, the area consists of more than 5 different parent materials strong alluvial formations (Bobbins, 2011) and pronounced differences in land use, such as farmland, livestock and natural area. Moreover, there is both a large interannual (100-700 mm) as well as monthly temporal variability (10-70 mm) in the amount of rainfall in the Baviaanskloof catchment. The latter causes the area to be sensitive to both drought and erosion (Jansen, 2008). This underlines the importance for managing soil water infiltration and retention.

## 2.2 Data collection and preprocessing

*Table 1 Overview of environmental variables used as input in this research, including source, area coverage and scale/resolution*

| Soil forming factor | Description | Variable | Source | Area | Vector/Raster |
|---|---|---|---|---|---|
| **Climate** | Mean and standard deviation (2000-2018) | Yearly average precipitation<br>Heavy rainfall<br>Cumulative dry days<br>Min. temperature<br>Mean temperature<br>Max. temperature | South African Environment Observation Network (SAEON) | Study area | Vector |
| | | Mean solar radiation<br>Land Surface Temperature (LST)<br>Soil Moisture Index<br>Albedo | DEM (ASTER)<br>Landsat 8, 09-2018, NASA<br>Sentinel 09-2018, ESA<br>Sentinel 09-2018, ESA | Study area<br>Study area<br>Study area | Raster 30x30 m<br>Raster 30x30 m<br>Raster 10x10 m |
| **Organisms** | NDVI (2000-2018) | NDVI timeseries (2000-2018)<br>Average NDVI<br>NDVI September | Landsat 7 (2000-2016), 8 (2017-2018) (NASA)<br>Sentinel 09-2018, ESA | Study area | Raster 30x30 m<br>Raster 10x10m |
| | Vegetation Height | Vegetation height (cm) | Fieldwork campaign | Plot | Vector |
| **Relief** | Digital elevation model derivatives | Elevation<br>Slope<br>Aspect<br>Flow accumulation (MUFF)<br>Topographic Wetness Index (TWI) | DEM (ASTER) | Study area | Raster 30x30 m |
| **Parent material** | Geological map | Geomorphological unit | SAEON | Study area | Vector |
| | Field measurements | Stone Fraction<br>Texture | Lab analyses | Plot | Raster 30x30 m |
| **Time** | Date of restoration | Restoration date | Livinglands | Plot/area | Vector |
| **Other** | Field measurements | pH, EC, nutrients, SOC. | Lab analyses | Plot | Raster 30x30 m |
| | Field description | Surface conditions<br>Descriptive soil horizons | Field data | Plot | Raster 30x30 m |

### 2.2.1 Environmental variables

In order to quantify soil development, information about the environment was needed to teach us more about soil functioning and to use for establishing a prediction of the soil development state.

#### *2.2.1.1 Local data*

The majority of the required environmental data was derived from existing maps and databases. More specifically, climatic and geological data as well as the date of restoration (Table 1) were received from a local observation network (Table 1) and the local policy makers in the study area, respectively.

Thereby, the geological data did not need any adjustments since it is supplied as a spatial map with each polygon representing a different parent material type and age.

Precipitation data was obtained from a local observation network (Table 1). This data covers 16 weather stations across the Baviaanskloof. From this data, the measurement period of 2015 to 2016 is used to predict the yearly average rainfall, cumulative heavy rainfall (97.5% percentile) and the number of dry days (2.5% percentile), since in this period available data was most complete for all stations. The goal was to use the rainfall variables in assessing differences between sites, but also to predict the spatial variability in SOC. To achieve this the point data of the rainfall variables were converted to a spatial map over the study area. This was done by performing regression kriging with the elevation as explanatory variable (Teng et al., 2014). This has eventually resulted in a spatial map of precipitation, at 30x30m resolution. For this study area, the temperature data did not offer the possibility to construct derivations other than the mean, since it was only available as a map of yearly maxima and minima.

The date of restoration was available as a polygon across the full study area. Some polygons included multiple restoration dates because they were repeatedly planted with vegetation due to a failure of initial vegetation. It is assumed that the first restoration date represents all *Spekboom* that is planted for each polygon, because it was unknown which planting practice has the largest contribution to the total planted *Spekboom* within a polygon. Subsequently, the derived planting dates were converted to the number of days after restoration by calculating the time difference with the 21th of September 2018, which is the fieldwork date. The uncertainty associated with the assumption of a single representative restoration date was quantified by calculating the standard deviation of the different days of restoration, associated with the multiple planting dates stated.

#### *2.2.1.2 Remote Sensing*

The resulting environmental variables from the data collection were derived from freely accessible RS-images (Table 1). Preprocessing of satellite data was done according to standardized processing methods if needed. Three sources for the acquisition of the remote sensing images were used and will be more thoroughly described in the following sections.

*Landsat images*

The NVDI dataset, was derived from 30x30m resolution Landsat daily imagery from the year 2000 to 2018 (https://earthexplorer.usgs.gov/) according to the following Equation 1 by Tucker (1979):

$$NDVI = \frac{(NIR-VIS)}{(NIR+VIS)} \qquad \textit{Equation 1}$$

where the NIR (near infrared) and VIS (visible red) were obtained from Landsat band 5 and 4 respectively. Images where reported cloud cover exceeded 10% were rejected. Moreover, the Landsat Control Report (QR) attached per image was used to identify flaws in the cell values of the timeseries images and set them to NA in case of the presence of any type of possible flaw. This means designated fill, dropped frame, terrain occlusion, water, snow, ice, cirrus and clouds are removed. Also values below zero and above one, were set to NA, since they do not represent realistic conditions.

Based on the previously obtained Landsat NDVI at the month of fieldwork (September 2018), the land surface temperature was derived according to the following Equation 2:

$$LST = \frac{BT}{(1+\frac{0.00115*BT}{1.4388}*Ln(\varepsilon)} \qquad \textit{Equation 2}$$

where the land surface temperature (LST) was defined based on the brightness temperature conversion (BT) and emissivity ($\varepsilon$) as a product of NDVI, whose equations are described in Avdan and Jovanovska, (2016).

Subsequently the Soil Moisture Index (W) was calculated using the following Equation 3:

$W = \frac{i_d + s_d NDVI - LST}{i_d - i_w + (s_d - s_w) NDVI}$ *Equation 3*

where $i_d$ and $i_w$ equal the dry and wet intercept, while $s_d$ and $s_w$represent the dry and wet slope of the edges of the trapezoid approach described by Sadeghi et al. (2017).

*Sentinel images*

In order to derive a more accurate prediction of the actual vegetation status, the NDVI was calculated with a smaller spatial resolution (Table 1: 10x10m). Thereby, Equation 1 is applied to band 8 (NIR) and band 4 (VIS) of Sentinel 2 images to derive the NDVI for the month of fieldwork (September 2018).

Subsequently, using the derived NDVI the Albedo ($\alpha$) was calculated with the following Equation 4:

$\alpha = \sum_{b_i} \left| \rho_{b_i} * w_{b_i} \right|$ *Equation 4*

where $\rho$ is the reflectance for band number $(b_i)$ and $w$ equals the weight for the associated band number ($b_i$) as described by D'Urso and Calera (2006).

The derived Sentinel albedo and NDVI images, were resampled to the target plot resolution of 30x30m using the aggregated mean value.

*DEM derivatives*

The DEM is derived from ASTER 30x30m images . Following that, the DEM derivatives (slope, aspect and flow accumulation, mean solar radiation) were generated using the designated ArcMap tools. Moreover, the Topographic Wetness Index (TWI) has been calculated using Equation 6.

$TWI = \ln \frac{\alpha}{\tan(\beta)}$ *Equation 5*

where $\alpha$ is the ratio of catchment area and contour length and $\tan(\beta)$ representing the slope (Beven and Kirkby, 1979).

All the environmental variables are ordered based on the main soil forming factor they are related to (Table 1). Eventually parameters as shown in Table 1 were available over the study area. While for SOC, pH and fine fractions this was available for the chosen plots, the RS-derived covariates (including NVDI) were known for the whole study area on various scales (Table 1).

## 2.2.2 Fieldwork

### *2.2.2.1 Site selection*

In order to obtain enough information to quantify the development of soils in the restoration area, soil samples were taken in the Baviaanskloof catchment. The sampling sites were chosen based on a combination of purposive sampling (Patton, 1990) and stratified sampling (Neyman, 1934). This sampling strategy starts with the creation of strata which were based on each unique combination of parent material and the Year Of Restoration (YOR) and livestock presence. These were hypothesized to be the most important spatial factors contributing to the SOC distribution. YOR and livestock exclusion were expected to determine the tree height distribution while the geology was expected to be the driver of soil physical conditions. Therefore, both were deemed important for soil restoration in semi-arid areas (Van Luijk et al., 2013). Each stratum was at least once represented in the selection of sampling sites, except for the sites with livestock inclusion. These sites were expected to all show reduced tree growth conditions and therefore strata with livestock exclusion were reduced to a coverage of 13 samples. Secondly, in each strata combination, spatial coverage and accessibility were taken into account as factors for the selection of the plots (Fig. 1). Finally, the site-specific conditions regarding the success rate of *Spekboom* restoration were accounted for by only sampling at sites which had a success rate ranging from 75% to 100%. This success rate was based on surviving *Spekboom* observed. In case these 75-100% desired success conditions were not present in the strata, the highest success rate found was used. This was done to make sure sites were comparable in terms of land cover, but also to be able to detect the biggest possible change in soil development allowing to set up significant relationships.

#### *2.2.2.2 Site description*

There was need for a method to support the results of the analysis of quantitative variables. More specifically, in case the measured quantitative variables would be limited in their coverage, contain large errors or were just absent, qualitative variables can be used for a better understanding of environmental conditions. Also, a first understanding of site-specific conditions, providing an important contextual background for further analysis in this research was required. In order to achieve this, a qualitative analysis was performed. This means, at the designated plot sites, the soil forming factors were qualitatively assessed at 3 different categories according to FAO (2017). These include general site information, locations, soil formation factors and soil description.

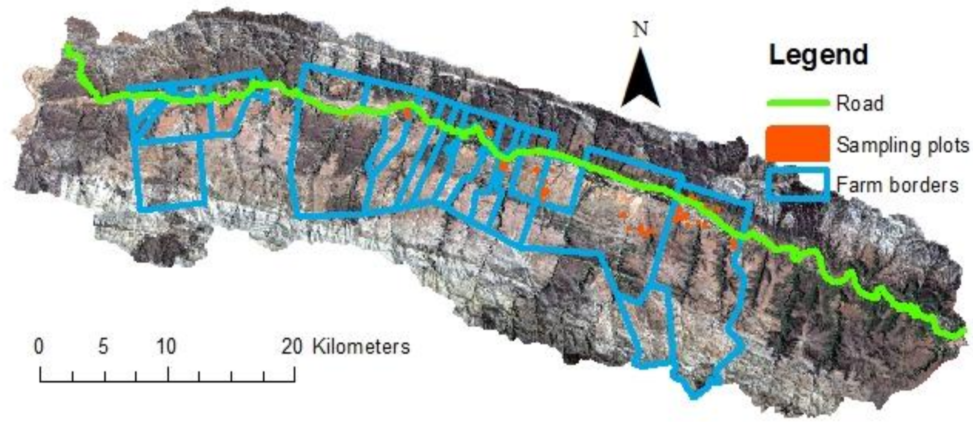


*Fig. 1 Map of sampling locations along the Baviaanskloof catchment including farm borders and the road.*

#### *2.2.2.3 Soil samples*

At every 30x30 meter sampling plot in Fig. 1, Valeri plot sampling (Rossello and Baret, 2007) was used to account for spatial variability. This means sampling was done at 10-15 locations per plot until the 2-liter bag was full, including the corners and center of the plots using a 100 ml iron sampling ring. Only the top 10 cm of the soil was sampled, because soils were barely exceeding the 20 cm depth due to their eroded state. Also, the SOC accumulation is expected to be highest in the top part of the soil. In case the soil was too dense or stony to sample with the ring, an iron shovel was used to carefully imitate the ring sampling procedure. During sampling, locations were chosen within a 50 cm range of any *Spekboom* shrub. This purposive sampling method (Neyman, 1934), was done based on the assumption that the relatively young and small *Spekboom* vegetation has not yet been able to affect more than 50 cm of its environment (Mills and Cowling, 2010). Sampling outside this range would have caused uniform values over the sampling sites, which would be hard to compare. Afterwards the collected soil was mixed into a 2-liter bag.

The collected soil samples were physically and chemically analyzed on nutrients, texture, pH, EC and SOC (Walkley and Black, 1934) (Table 1). These analysis (*appendix,* Table A9) were done at a lab.

Subsequently, the mean and standard deviation of the collected variables were computed and their relationship with YOR was analyzed.

## 2.3 Relation between variables

In order to obtain knowledge on the most important variable determining soil development, but also to gain an overview of the behavior of variables over space and time, the relation between variables was assessed. Eventually this was the main source of information in the derivation of the main driver of soil development in the area. Firstly, this was done by performing a Principal Component Analysis (PCA; Hotelling, 1933). Subsequently the Pearson correlation coefficient (PCC) (Pearson, 1896) between all variables was used as a first identification of temporal variability. Finally, a general overview of variables was made by making use of summary statistics. Moreover, the variables were individually assessed based on the main soil forming factors (Jenny, 1941), since each of them is expected to behave differently in relation to the restoration practices applied.

### 2.3.1 Assessing the relation between variables using PCA

As a first step in determining the main drivers of soil development, a PCA (Hotelling, 1933) was performed to analyze the different domains the variables are covering in explaining the total variability in dataset of environmental variables. The PCA could teach us which variables were similar in describing the variability in SOC following the principal components of the dataset variables. Not only this but also the distribution of pristine, degraded and restoring samples was valuable for interpreting the effect of the environmental variables in perspective of degradation reversal.

### 2.3.2 Temporal restoration curve

There was need to gain insight in the dynamics of SOC over time, related to the desired goal of predicting the temporal development of SOC. Eventually, this forms the basis for the comparison with other methods for quantifying soil development, but also offers a possibility to get a perspective view on the derived SOC values and maps by assessing the values' temporal proximity to pristine conditions. In order to achieve these goals, there was need for a temporal representation of SOC data in the research area. Unfortunately, past SOC measurements were not available in any restored part of the Baviaanskloof catchment. Therefore, we made use of a soil development chronosequence, which was constructed by selecting plots differing in the year and month they have been restored, from 2006 up to 2015. The unrestored degraded plots in the study area were assumed to show the initial baseline conditions (restoration of 0 days). In addition, the pristine conditions were assumed to have a YOR of 7000 days, which corresponds with the moment the SOC is expected to reach pristine conditions as is found by Jia et al. (2005).

Using a chronosequence for such purpose is not necessarily appropriate (Walker et al., 2010). In case initial conditions are different, mainly in terms of primary and secondary ecological succession, the site-specific development of these ecological successors may differ. This could potentially cause a multiple pathway response in these successors, where the diverging pathways are least appropriate for establishing a temporal relationship useful for understanding the trend in SOC development. In the current situation, this would mean that e.g. differences in initial microbiological communities (type and magnitude) are likely to exist due to different initial land cover state (Sparling, 1992) or aggregate stability (Hattori, 1988) may cause a different response of the system to the *Spekboom* establishment. The system response might be seen in the SOC chronosequence and is expected to follow multiple distinguishable trends, i.e. different pathways. This effect for SOC has been seen in earlier studies done at an old mining reclamation site (Insam and Domsch, 1988), where the microbial development follows two distinguished paths with a similar weakening effect.

To analyze how appropriate the application of a chronosequence was to the current area, the potential occurrence of pathways should be identified. This has formerly been done by using repeated measurements over time (Lebrija et al., 2010), but only by monitoring variables of primary and secondary ecological succession. For the Baviaankloof catchment, these measurements nor timeseries of SOC were available. However, a wide range of environmental variables were measured which could most likely heavily influence the dynamics of organic carbon and therefore its potential pathways. This means for the purpose of distinguishing pathways, the chronosequence might not be fully suitable. In that case, the similarity of important SOC drivers would have distinguished the pathways, rather than its temporal development.

Instead of distinguishing pathways to quantify the potential behavior of soil organic carbon, different development scenarios were established for this purpose. These scenarios would tell us more about the potential paths SOC can follow, without the need to take complexity of environmental variables into account. The maximum and minimum scenario were calculated based on the 95% confidence interval of the SOC measurements of subsequent years. Such that every range included at least 4 measurements which were similar in days of restoration. This method was applied with the assumption that the standard deviation between the set of points of similar restoration date would change following these days. This would happen due to the expected divergence of the scenarios over time caused by a different development pathway. Eventually this would mean the points are not normally distributed and the overall standard deviation would not be valid. To validate the significance of the scenarios, three trendlines were fitted through the scenario points and the original chronosequence points, by using a linear exponential model. The exponential model was expected to represent realistic conditions since it is in line with the theoretical framework describing a positive feedback mechanism (Suding et al., 2004) as a result of a SOC increase. Following the derivation of the trendlines, the quality of the fitted trends was evaluated using the R-squared value of the applied model and its P-value of significance.

### 2.3.3 Organisms

In order to understand the role of vegetation in soil development, the relationship between tree height and SOC was analyzed. Tree height was estimated at plot level and related with the chronosequence time. This temporal trend could teach us more about the succession of vegetation for the chosen plots. Thereby, the vegetation was expected to follow a logistic growth curve (Hunt, 2012) since this logistic behavior had been formerly associated with *Spekboom* development (Lombard et al., 2001). This logistic behavior means that proceeding with time a maximum tree height will be reached. This height is assumed to be represented by the observed vegetation height at the pristine plots.

The tree height measurements, however, did not have the desired full spatial coverage and therefore were not suitable to be used to monitor SOC development in the Baviaanskloof catchment. Consequently, the relationship with a proxy for vegetation height should be investigated. In this case the NDVI was chosen, which had already been used intensively for monitoring restoration practices activity in semi-arid areas (Chen et al., 2007; Chen and Duan, 2009; Liu et al., 2012,). The overall mean NDVI

(Table 1) was extracted at plot level by using the mean of all pixels. Following this they were related to YOR and tree height observations.

Insight in the suitability of the NDVI for observing vegetation activity in the research area was important since the NDVI might be useful for monitoring purposes and would act as input for the prediction of SOC. In order to analyze this suitability, the change of NDVI as a result of the applied restoration practices was assessed. This was done by creating the yearly timeseries of the NDVI-index for a wide time range of observations (2000-2018). Additionally, timeseries of yearly rainfall measurements was added since it is previously reported by Davenport and Nicholson (1993) that the NDVI is sensitive to changes in rainfall. Both the variables observations were reduced to plot level and averaged among plots in order to get an overview of the response of NDVI to the restoration efforts. Finally, the timeseries trends were identified by performing loess regression and smoothing in R.

### 2.3.4 Relief

In the perspective of soil erosion, relief would be important in decreasing SOC stocks. Therefore, the effect of slope, TWI, aspect and flow accumulation on SOC was assessed using the PCC and P-value of significance. In addition, the PCA scores were compared. These scores are important for understanding the similarity of the DEM-derived variables, which can be very much similar due to their similar origin which is the digital elevation model.

### 2.3.5 Climate

Due to the potential impact of climatic variables on multiple processes determining SOC stocks, it was important to look at them separately. In fact, temperature and rainfall both potentially would have an effect on tree growth, mineralization and erosion (Jenny, 1941). Therefore, in order to understand the potential behavior of the climatic variables, the scores of the PCA were analyzed for each climatic variable. Moreover, the PCC offered more insight in the direct relationship with SOC.

### 2.3.6 Parent material

The relation of parent material with SOC can be found in the texture class, stoniness and location of the geographical units. This section investigates the effect of parent material on the potential differences in pathways followed by SOC. This was achieved by constructing chronosequence curves using an exponential linear model for all different parent materials and assessing the goodness of fit with the R-squared value and P-value of significance obtained. A statistical T-test was performed on categorical variables, like Parent material to explain differences in environmental conditions

### 2.3.7 Other

Livestock grazing is an important factor in determining the growth of *Spekboom*, because it was the main cause of soil degradation in the Baviaanskloof catchment (Jansen, 2008). Since livestock is not fully excluded in the study area, its presence could have inhibited the growth of trees and therefore the SOC development after restoration. In order to investigate the potential effect of livestock presence, the development of the tree height and SOC stock was assessed for two separate datasets of livestock exclusion and inclusion. Curves of tree height and SOC stocks over the chronosequence were fitted according to the expected temporal relation of SOC (exponential) and tree height (logistic). These curves were again assessed on the goodness of fit, which could offer more insight in the potential effect of livestock.

## 2.4 Mechanistic model

### 2.4.1 Model setup

In order to understand and predict the spatial variability in SOC, a mechanistic model was set up (Fig. 2). This was done according to a method documented in recently published data (Hendriks, 2018). This method used well known relationships of 1) litter production (LP), 2) turnover rate (TR), 3) mineralization (MR) and 4) soil erosion (ER) with the SOM stock. Since the SOC stock is a good indicator of the SOM dynamics during the land use change process (Leifeld and Kögel, 2005), it was expected the same approach could be used. This approach relies on the assumption of an equilibrium state between the afore mentioned 4 processes influencing SOC stocks according to Equation 7:

$$\frac{LP*TR}{MR+ER} = 0 \qquad \textit{Equation 7}$$

eventually coming up with the following Equation 8:

$$SOC_{stock} = \frac{LP*TR}{MR+\frac{ER}{BD*SD*100{,}000}} \qquad \textit{Equation 8}$$

Where BD and SD equal bulk density ($g\ cm^{-3}$) and soil depth ($cm$), respectively. As mentioned above, the 4 processes determining SOC are LP, TR, MR and ER. These processes were predicted using known physical relationships with environmental variables (Hendriks, 2018) represented in equations 9-12:

$$LP = k_1 + (k_2 * Th) \qquad \textit{Equation 9}$$

$$TR = \frac{1}{3.09+2.7e^{Cl}} \qquad \textit{Equation 10}$$

$$MR = k_3 e^{(k_4*Sm+k_4*Te)} \qquad \textit{Equation 11}$$

$$ER = k_6 + (k_7 * S) + (k_8 * Wa) \qquad \textit{Equation 12}$$

Where $Th$, $Cl$, $Sm$, $Te$, $S$ and $Wa$ equal tree height, clay content, soil moisture, temperature, slope processes and water accumulation respectively.

The k values ($k_1$-$k_8$) are constants that were optimized in the calibration process explained in section 2.4.2.

The adjustments to the original mechanistic model of Hendriks (2018) can be found in the variables used for describing the processes. Previously no differences in clay content were found and a combination of

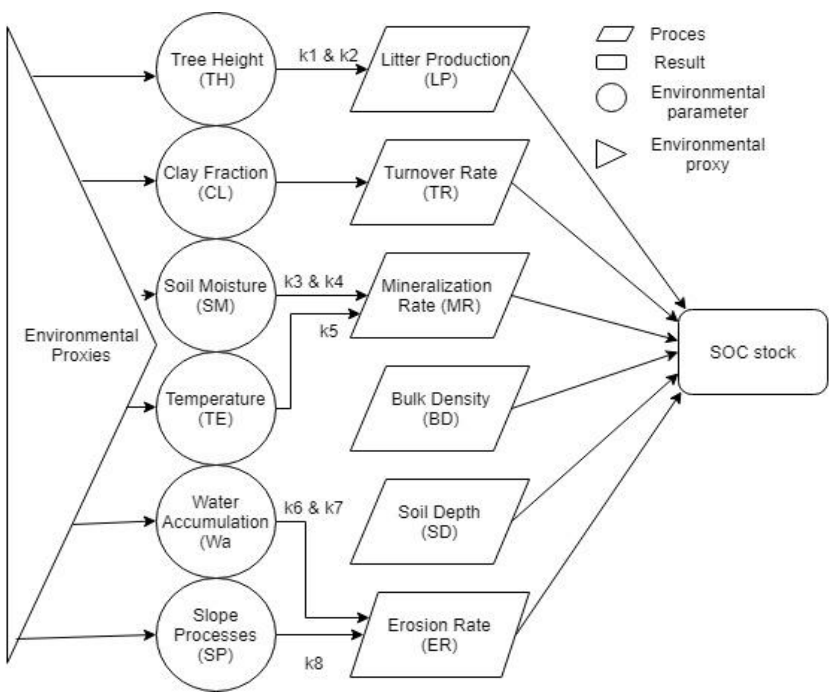


*Fig. 2 the structure of the mechanistic model approach for predicting SOC carbon* temperature and soil moisture in the calculations for the mineralization rate were absent. Moreover, previously only the slope had been taken into account for predicting the erosion rate, while it is known that more variables are influencing this rate (Ziadat and Taimeh, 2013). It was expected to find better prediction results by increasing complexity and therefore it was deemed important to describe processes realistically.

### 2.4.2 Model calibration

To determine which environmental proxy would be suitable to predict the variables in the equations described above, expected physical relationships formed the first basis. This means tree height was expected to follow the same trend as NDVI, clay content would be determined by parent material and water accumulation determined by relief. Additionally, the combined results of the PCA and correlation coefficients were used as a confirmation. The variables which did have a similar direction in the PCA were used interchangeably, whereas the largest correlation with SOC was also a determining factor for selecting the proxies to use to predict SOC.

The R optim. base function was used to optimize the k-values in the model equations (Eq. 9:12) for the root mean square deviation (RMSD), which equals:

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(SOC_{obs,i} - SOC_{pred,i})^2} \qquad \textit{Equation 13}$$

Where $n$ is the number of observations and $SOC_{obs,i}$ and $SOC_{pred,i}$ are both the observed (Equation 14) and predicted SOC stock (Equation 8) in sample number $i$, respectively. The observed SOC fraction ($fSOC$) first needed to be converted to SOC stocks. Since the stone fraction could not be accurately determined with the current sampling methods and only the first 10 cm was sampled, the calculation of $SOC_{obs}$ was done using a constant value for stone fraction ($SF = 33\%$) and bulk density ($BD = 1.53\ g/cm^3$) according to Equation 14:

$$SOC_{obs} = fSOC * (SD * BD * (1 - SF)) \qquad \textit{Equation 14}$$

Where, soil depth (SD) in cm was applied up to 10 cm.

The k values ($k_1 : k_8$) in equation 9:12 were calibrated by optimizing for the k values which generated the least RMSD when applied to the formulas together with the chosen environmental proxies.

The calibration process resulted in a RMSD and a correlation coefficient between the predicted and the measured SOC values. These values were compared with a calibration run with the original set of proxies and the original results when setting up the model by Hendriks (2018). Moreover, the residuals or unexplained variance of the data, represented by the observed SOC stock, were subtracted from the difference of observed and modelled SOC stock. Following that the residuals were analyzed for its potential causes. These causes were expected to relate to assumptions made in the model regarding spatial uniformity. To test this hypothesis the relations with environmental variables that are known to not be accurately represented in the model, were assessed.

### 2.4.3 Spatial mapping

In order to extrapolate the SOC measurements from plot to catchment level the previously calibrated model needed to be applied to catchment scale. This spatial prediction was performed by applying equation 8 to 12 to the two calibration runs: Hendriks original model and the current application. The input of the equations was the calibrated k-values and the set of environmental proxies (on catchment level instead of plot level). To generate a final SOC map for the Baviaanskloof catchment, the residuals of the calibrated model were interpolated using ordinary kriging (Elbasiouny et al., 2014) and added to the spatial model equation prediction. Hereby, negative SOC stock values were physically impossible and therefore set to 0. The standard error of the kriging prediction was visualized and used to assess the quality of the derived SOC maps.

## 2.5 Comparison other maps/models

### 2.5.1 Spatial

The goal of the next applied method was to get more insight in the quality of the prediction of the SOC map and therefore the suitability of this mechanistic mapping method for monitoring current soil development as a result of ecological restoration. Moreover, the following methods did aim at getting more insight in the impact of choosing different environmental proxies. To achieve those goals, a comparison of the mechanistic model and chronosequence approach with other methods was required. Therefore, the generated SOC map was compared with a SOC map for the Baviaanskloof catchment which was derived with the same environmental proxies as originally designed by Hendriks (2018).

### 2.5.2 Temporal run, Carbon Benefit Project and RothC

In order to assess the suitability of the current chonosequence method for predicting the temporal development of SOC, but also to get more insight in its relationship with the established carbon models, a comparison between methods was required. Therefore, the temporal run of the mechanistic model was compared with two former methods: Carbon benefit project (Victoria et al., 2012) – detailed assessment and RothC (Coleman and Jenkinson, 1995) for assessing carbon sequestration in soils as a result of land use change (Table 2). Since both methods grant an annual SOC sequestration value, they were compared with the slope of both the temporal mechanistic model run and the chronosequence restoration curve derived in section 2.3.2. This was done by calculating CO2e.year$^{-1}$ into C.year$^{-1}$ using the moll fraction CO2:C.

The RothC and CBP calculations were performed on catchment scale, while the chronosequence was only derived on plot scale. Comparison required a setup of the carbon models such that the derived average catchment conditions would be similar to the ones observed at plot level. Both the CBP and RothC model provided the opportunity to put the chronosequence scenarios in a qualitative perspective. This was done by differing the magnitude of certain variables, such that SOC behavior can be mimicked and understood.

The following sections describe the setup of the Carbon benefit project – detailed assessment (Table 2) and the RothC tool for ecological restoration case in the study area.

*<u>Carbon benefit project</u>*

The carbon benefit project detailed assessment compares the baseline scenario (degraded) with the actual project scenario (restored) in terms of carbon sequestration according to the conceptual framework described by Victoria et al. (2012). These scenarios differ in the fractions of the different types of landuse (Table 2). For the Baviaanskloof area this would mean an increase of 'Forestland' and a decrease of "Severely degraded grassland". Where the latter represents degraded area. The magnitude of these estimations was done according to field observations documented in section 3.1.2, but also according to local information as well as personal communication with restoration project leaders. Using the CBP detailed assessment tool a maximum and minimum project scenario was implemented. The maximum scenario would mimic what was measured on plot level by sampling close to trees which would be a success rate and tree cover of near 100%. The minimum scenario represents what was expected to happen on plot/catchment level, when assessing the average tree cover observed on plot level, which would be a 100% success rate.

| Method | Mechanistic model | Carbon benefit project | RothC | Chronosequence |
|---|---|---|---|---|
| Approach | Semi-emperical | Conceptual | Conceptual | Observations |
| Purpose | Predicting measurements | Land use change | Crop management | Temporal behavior |
| Climate input | P, Temp | None | P,Temp | None |
| Texture input | Clay fraction | None | Clay fraction | None |
| Grazing pressure input | None | Vegetation and Manure | Manure | None |
| Vegetation parameter | Empirical fitting | Crop selection | Decomposition parameter | None |
| Spatial component | Spatial input | Land use cover fractions | None | None |
| Temporal component | None | Linear yearly | Linear yearly | Exponential daily |

*Table 2 Overview of different methods applied in the research. These are compared using their purpose, input parameters and temporal/spatial components*

The carbon benefit project tool also required a selection of the vegetation type and status, which was selected to be mature dry tropical forest plantation. This vegetation is well representative for *Spekboom*, because similar to *Spekboom,* the tropical vegetation, has a strong potential to sequestrate carbon (Eleanor Milne, personal communication, Januari, 2019; Mills and Cowling, 2010). The fractions of grassland types (moderately degraded, heavily degraded and natural) were chosen according to average measured land use fraction. Grass, bare soil and shrubs represent these grassland types, respectively. The cover of annual crops, perennial crops, agroforestry, wetlands and settlements were set to 0 ha since the existence of a 'natural' system is assumed. The livestock parameter could be changed to investigate the effect of livestock, where the implementation of the number and type of animals were based on personal communication with farmers. However, the model assumes livestock grazing is not affecting the implemented vegetation, since dry tropical forest species are too tall for being grazed by livestock. Consequently, this parameter was excluded from analysis.

*RothC*

Another method that was used for comparison is the RothC carbon model (Coleman and Jenkinson, 1995). This model requires climatic data (maximum, minimum, annual mean precipitation and temperature), crop cover, crop sequester potential and clay content as input (Table 2). The former was provided using the available precipitation and temperature datasets. Crop cover was set to equal to the coverage of *Spekboom* at the comparison scale. In case of plot level, this meant crop cover was set close to 100%, since sampling was solely done in a range of 50 cm from the vegetation. In contrast, catchment/plotwide comparison needs a cover commonly found at the plots. Changing crop cover could be a powerful tool for mimicking the effect of livestock grazing, but its effect is absent in the RothC model. Crop sequester potential was set to equal tropical species, since its potential equals *Spekboom* potential (Mills and Cowling, 2010). Required clay fraction input was available from lab measurements as described in section 2.2.2. Other factors that could be changed in RothC such as residue and yield management, were not applicable to the management practices in the Baviaanskloof or any information about them was not available (manure input).

# 3. Results and Discussion

## 3.1 Data Collection

### 3.1.1 Environmental Variables

In total, 50 soil samples were taken from 50 plots covering the study area. From these 50 plots 4 had pristine conditions and 6 were non-intervened. The remainder of the plots had been restored between the year of 2006 and 2015, which resulted in an average value of years of restoration (YOR) for the samples of 7.86 years (Table 3).

*Table 3 Results of remote sensing and legacy data, including mean, standard deviation (st.dev) and the pearson correlation with year of restoration (YOR)*

| Variable | Mean (st. dev) | Pearson R2 with YOR |
|---|---|---|
| Years of restoration (yr.) | 7.86 (4.7) | 1 |
| Normalised Difference Vegetation Index (-) | 0.29 (0.1) | 0.45 |
| Topographic Wetness Index (-) | 6.13 (1.3) | -0.37 |
| Elevation (m) | 497.44 (86) | 0.1 |
| Slope (deg.) | 13.55 (7.8) | 0.48 |
| Albedo (-) | 0.12 (0) | -0.07 |
| Solar radiation (W/m2) | 369.41 (39) | 0.1 |
| Land Surface Temperature (°C) | 38.97 (2.2) | -0.52 |
| Max. temperature (°C) | 19.43 (0.5) | 0.29 |
| Mean. temperature (°C) | 12.50 (0.4) | 0.18 |
| Min. temperature (°C) | 5.56 (0.4) | 0.1 |
| Mean Yearly Rainfall (mm) | 279.41 (23) | 0.07 |
| Mean Yearly Heavy Rainfall (mm) | 174.37 (15) | -0.52 |

A large range of topographic conditions was sampled. Slopes were ranging from 0 to 35 degrees and elevation from 350 to 720 meters.

The climatic conditions were variable with time, having a mean, maximum and minimum temperature of 12.5, 19.4 and 5.6 degrees over the year, respectively. Spatially, however, these temperatures did not show too much variation between plots (st.dev = 0.5 degrees). This invariability was caused by the coarse resolution of the temperature raster but is also related to the relatively small surface area of the study area compared of the climatic variations. The lack of spatial variability holds true as well ?? for mean yearly rainfall which only had a standard deviation of 23 mm derived from an average of 279.41 year$^{-1}$ (Table 3) for the study area.

### 3.1.2 Fieldwork

The sampling plots varied in land cover (Table 4). On average, 50% of the surface area of the plots contained bare soil while only 25% was covered by trees. This tree cover value is positively correlating with YOR (cor.= 0.82), which also holds for the tree height at these plots (cor.= 0.89).

*Table 4 Results of the fieldwork campaign site description including mean, standard deviation (st.dev) and the pearson correlation with year of restoration (YOR)*

| Variable | Mean (st. dev) | Pearson cor. with YOR |
|---|---|---|
| Surface Stoniness (%) | 42.88 (22) | 0.38 |
| Bare Soil (%) | 49.82 (25) | -0.55 |
| Shrubs (%) | 14.7 (18) | -0.36 |
| Grass (%) | 11.5 (22) | -0.03 |
| Tree (%) | 23.58 (26) | 0.82 |
| Tree Height (cm) | 72.7 (66) | 0.89 |
| Soil Depth (cm) | 24.42 (13) | 0.24 |
| Soil Stoniness (%) | 47.0 (25) | 0.19 |

The total surface stoniness, which consist of a sum of stone, boulder and rock fraction, covered a substantial fraction of the plots (mean = 42.88), which might be a possible indication of eroded conditions.

In general, the soils were relatively stony (mean = 47%) as well. This high fraction of stones in the soil increased forced use of a sampling shovel instead of the regular sampling ring and therefore introduced errors in sampling. More importantly the larger stones made it hard to sample the first 10 cm accurately. The stone fraction might be overestimated due to these phenomena. Generally, due to the degraded past of the study sites, average soil depth was rather shallow (24.42 cm). Nevertheless, there was still a substantial fraction of clay left (mean=14.08).

### 3.1.3 Lab Analysis

The analysis of 50 soil samples resulted in an average SOC per plot of 1.49% (Table 5) which ranged from 0.5% for degraded to 5% for pristine conditions. These SOC values showed a strong correlation with YOR (cor. = 0.75), but also other soil quality indicators, like bulk density, pH, permeability are correlated with YOR. These correlation values with YOR, respectively -0.13, -0.35 and 0.63 were smaller than the value for SOC, which might indicate SOC being a better indicator for the successfulness of restoration monitoring. Total nutrients did also increase with time (cor.=0.45), most likely due to the increase of organic matter content. Nevertheless, some nutrients, such as phosphorus (cor.=-0.38) did decrease with time, which might be an indication of continuously ongoing erosion. The same might hold for clay fractions which were not higher for older plots (cor.=-0.18), while you might expect a stabilized

*Table 5 Results of the fieldwork campaign lab analyses including mean, standard deviation (st. dev) and the pearson correlation with year of restoration (YOR)*

| Variable | Mean (st. dev) | Pearson cor. with YOR |
|---|---|---|
| Clay (%) | 14.08 (3.7) | -0.18 |
| Silt (%) | 13.92 (4.7) | 0.06 |
| Stone Fraction (%) | 0.33 (0.2) | 0.21 |
| Bulk Density (kg/m3) | 1.53 (0.3) | -0.13 |
| Permeability | 1.72 (1) | 0.63 |
| pH | 6.47 (1.2) | -0.35 |
| P (mg/kg) | 99.42 (76) | -0.38 |
| C (%) | 1.49 (0.9) | 0.75 |
| Total Nutrients (mg/kg) | 820 (403) | 0.45 |

concentration due to increased soil stabilization by vegetation.

## 3.2 Relations between variables

### 3.2.1 Assessing the relation between variables using PCA

By extracting the principal component of the different variables described in section 2.1, relations between the collected environmental variables were found. Soil organic carbon was situated mainly in the first dimension (Dim1) of the Principal Component Analysis (PCA), which can be seen in Fig. 3. SOC was explaining most variation in this dimension, which was in its own case the most important dimension (31.1%) of the dataset. That is why SOC (*appendix*, Table A11 cor.= 0.81) can be seen as the most important variable explaining the variation in the total dataset, together with restoration days (cor.=0.86), tree height (cor.= 0.89), and permeability (cor.= 0.86). These three last variables, on their turn, were all not accurately determined and therefore their scale of contributions to the restoration process becomes doubtable.

Dimension 1 also highly determines the transition from degraded to restoring to pristine in the graph. In this transition, most of the samples taken were more similar to being at the degraded site of the graph rather than the pristine corner, which tells us something about the current state of the system already. More specifically the magnitude of the environmental variables for the restoring plots showed more similarities to degraded plots rather than pristine plots, which indicates a tendency for the majority of

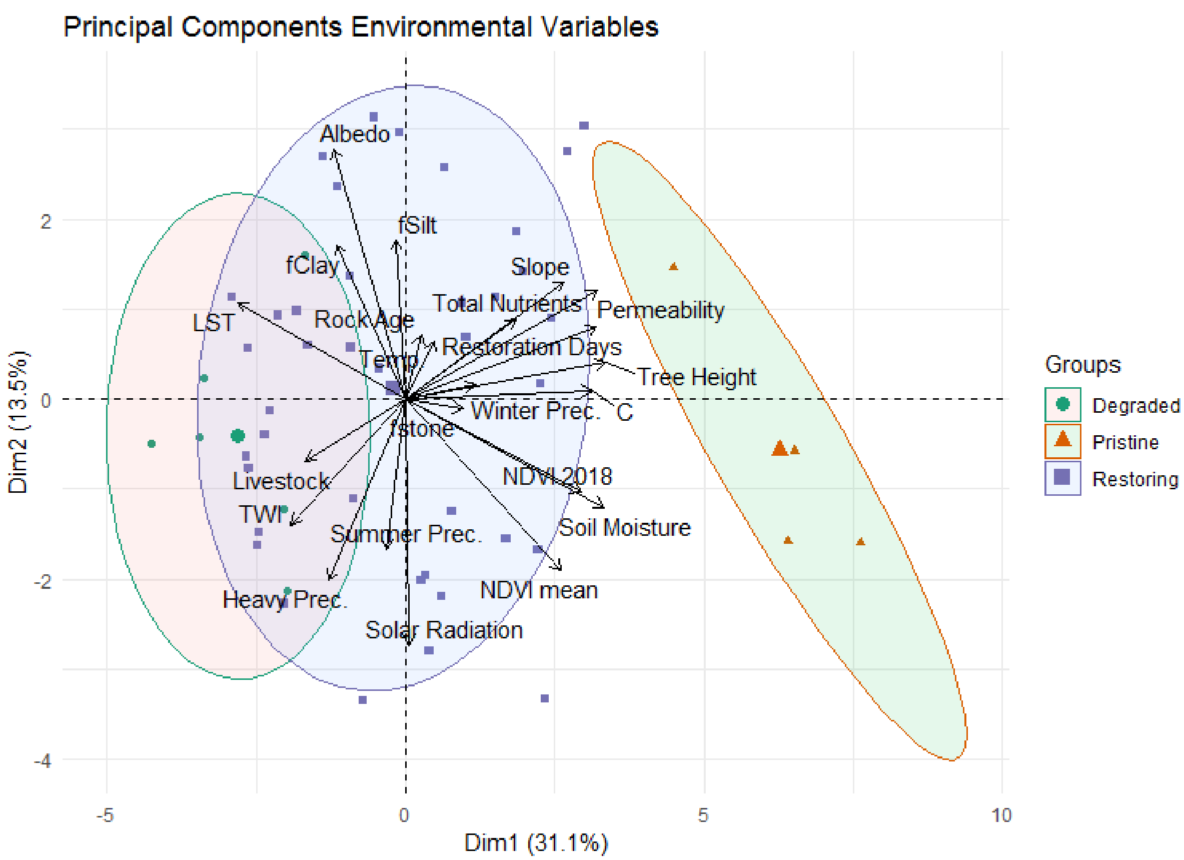


*Fig. 3 The Principal Component Analysis (PCA) of the most important environmental variables collected in describing the variability in the dataset.*

the variables to be more degraded instead of pristine. In fact, at least 5 samples can be qualified for being still completely degraded, based on the PCA.

What also became clear of the principal component analysis (Fig. 3) was the division of the soil forming factors in their scores on the different dimensions. Climatic variables (albedo, summer precipitation and solar radiation) were closely situated to each other in the second dimension. Whereas, for instance the variables related to parent material (rock age, clay and silt) were positioned lower in the second dimension. Because of the positions of the different soil forming factors on the PCA, also their contribution for explaining variation in soil organic carbon, and thus the total restoration process, were separately distinguished.

For the purpose of predicting SOC, the most important variables were assessed based on their scores on Dim1. Restoration days, slope and NDVI-related variables were the most important variables in Dim1, whereas the temperature, silt fraction and solar radiation seemed to be less important in influencing SOC. In case the absolute PCA scores overlapped, the variables could be interchangeably used for predicting SOC. In case of the variables derived from the NDVI (LST, soil moisture and NDVI 2018) this overlap was to be predicted. In contrast, the tree height, permeability and restoration days were from other origin, but explained similar variability in the dataset. Nevertheless, the contribution of predictors in explaining SOC by using Dim1 of the PCA, should be interpret with caution. In fact, SOC is also represented in Dim3 (cor.= 0.19), Dim4 (cor. = -0.16) and Dim5 (cor.=0.19). Therefore, the PCA cannot offer a definitive answer on the most important drivers of SOC. As a consequence, the drivers should be individually related to SOC and carefully interpret.

### 3.2.2 Temporal restoration curve

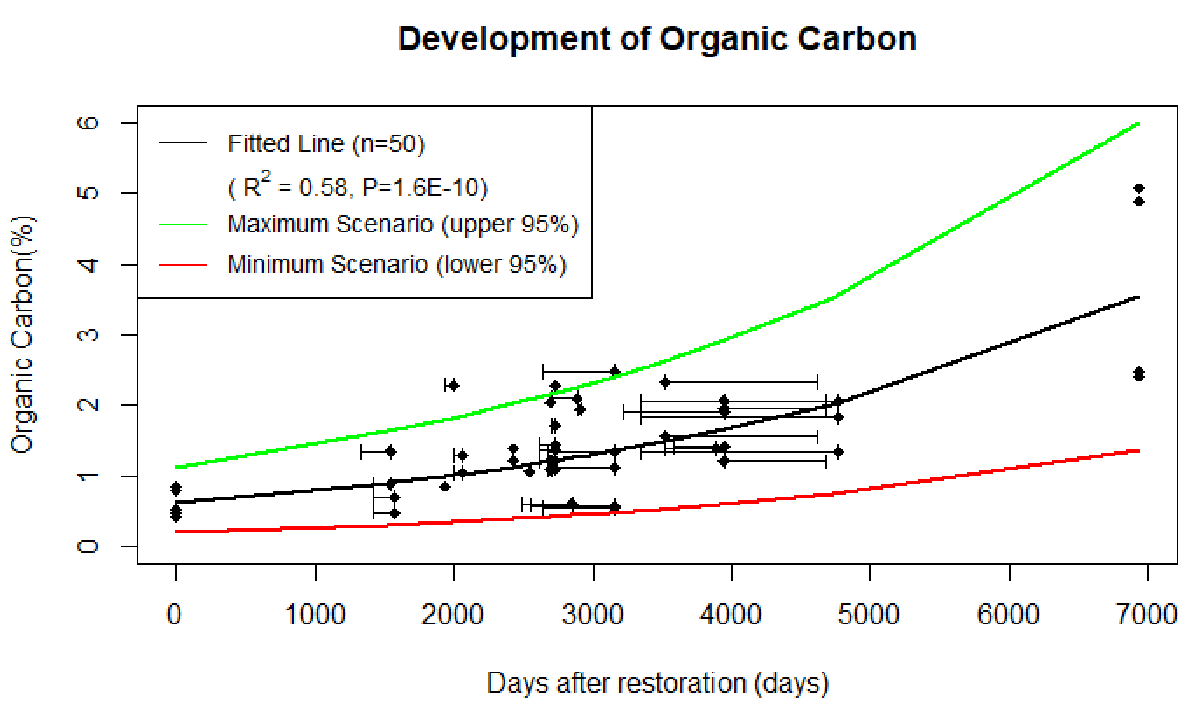


*Fig. 4 Development of organic carbon as a result of the chronosequence approach including horizontal uncertainty errors and multiple development scenarios*

Years after restoration (YOR), which is from now on visualized as days after restoration, is considered one of the most important factors in the restoration process (Fig. 4). Time was heavily correlated (*appendix,* Table A11; cor.=0.75) to organic carbon development which is understandable concerning the restoration measurements. Also, the found positive exponential trend with time agreed with the understanding of a theoretically stronger organic carbon buildup due to SOC's positive effect on soil functioning.

As a result of the chronosequence approach applied here, multiple scenarios of organic carbon development have been distinguished (Fig. 4). Both scenarios showed a divergence from the mean with increasing restoration days, which is in line with the hypothesis of the possible occurrence of pathways. Both the exponential models connecting the points representing the scenarios contained a R-squared value of 0.74. This high predictability of the scenarios strengthens the observations of significant divergence. The divergence means the standard deviation of the points did increase with time, such that there was potentially a driver which caused a difference in slope which was not linear. The points can be considered very unpredictable in time, because the possible development of SOC in time varies substantially. Where at day 0 the SOC could range from 0 to 1% it could differ at day 7000 from 1%-7% using the established scenarios in Fig. 4.

The creation of the chronosequence restoration curve could have some important uncertainties. Firstly, the placement of some points involved uncertainty on the x-axis, due to inaccurate monitoring of the date of restoration (YOR). This uncertainty is represented by the arrows in Fig. 4. Time of restoration was also important considering the initial climatic planting conditions, which potentially could be a factor influencing tree height success rate and finally development of SOC. The most important limitation regarding the approach of identifying the divergence is related to sample size. The limited measurements available created an uncertainty in determining the divergence or increased the standard deviation in sample points. The determination of divergence in the range from initial to restoring state was not affected by sampling size. In this range a lot of measurements defended the observations. In contrast, the measurements approaching the end of the chronosequence timescale, in particular pristine conditions, cannot be fully justified with only four measurements. Additionally, farms associated with heavy degradation due to livestock abundance are not sampled for a year of restoration before 2013, which causes an absence of points associated with livestock presence. Another potential uncertainty involving the interpretation of the derived restoration curve is related to variables spatially affecting SOC. These can be significant in determining divergence and therefore the effect of single environmental variables on the development of SOC will be more thoroughly discussed in the following sections.

### 3.2.3 Organisms

The role of organisms or vegetation (tree height) in the development of soil functioning can be found in both dimension 1 (*appendix*, Table A10; cor. = 0.88) as well as dimension 2 (cor. = 0.13) of the PCA (Fig. 3). The tree height is one of the most important variables in explaining SOC (*appendix*, Table A10; cor.=0.79). Tree height follows a logistic temporal trend (Fig. 5; $R^2$.= 0.88) which is similar to the exponential trend SOC follows in time (Fig. 4).

This determination of tree height as a driver of SOC might be partially affected by the high variability (Table 4; st.dev = 66) in the tree height of plots. More specifically ,this variability exceeds other environmental variables which leads to a less pronounced effect of uncertainty in its derivation and a more significant relation with SOC. The results of tree height for the plots can be deemed less accurate due to the judgmental measurements of heights higher then 2m or inaccessible stems.

Another restriction in the analysis of the tree height contribution to soil development is related to the overlap of tree height with YOR in the PCA, which means their variability over the plots is similar. Keeping in mind tree height generally increases with time as a result of tree growth, YOR is correlated with tree height. In case SOC and time also positively correlate as a result of an annual positive carbon balance, the tree growth correlation is to be higher due to this phenomenon then it would be in a condition with uniform YOR. Therefore, in general the contribution of tree height might be potentially overestimated. The same overestimation would be caused by SOC increasing fertility and water retention, such that it stimulates tree growth. The correlation coefficient includes the multiple effects and thus overestimates the single physical effect of tree height on SOC.

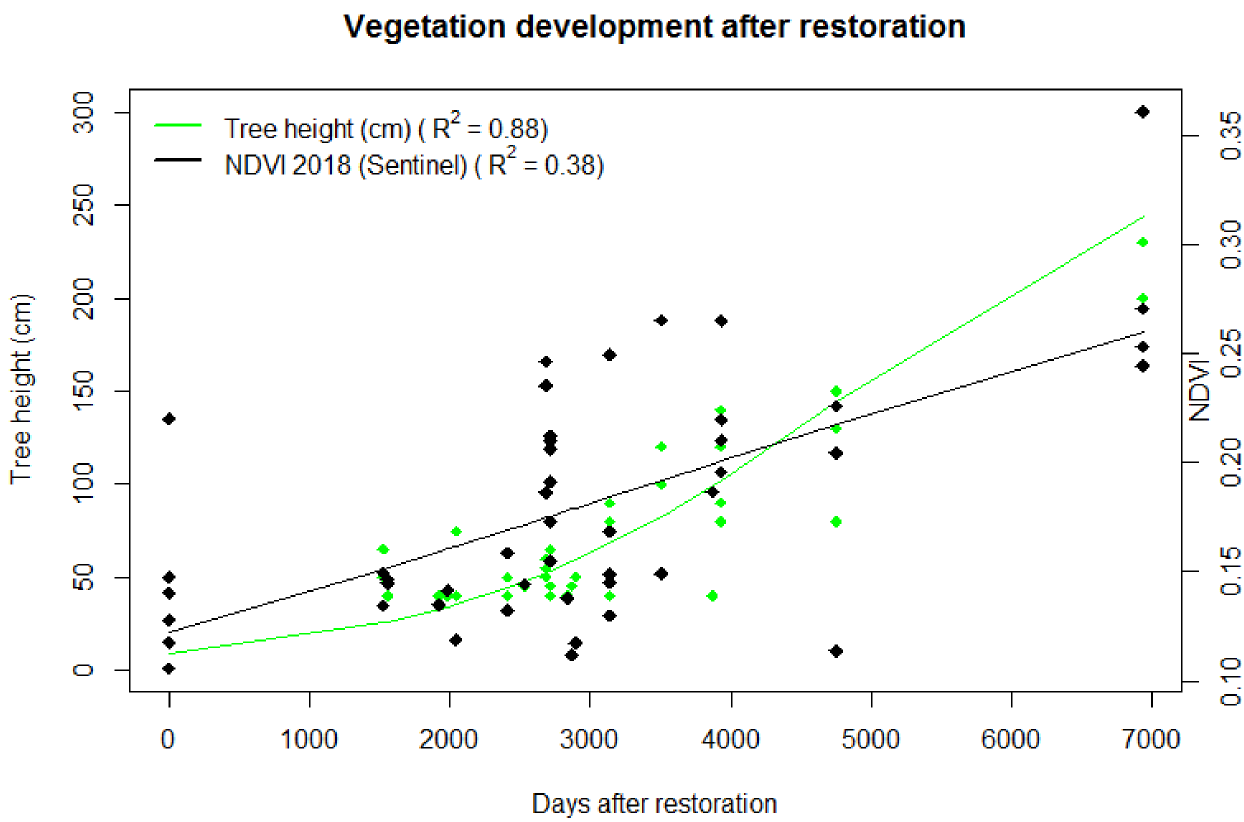


*Fig. 5 Vegetation development after restoration, including the measured Tree height in cm and the NDVI-index derived from September 2018 for 0 to 7000 days after restoration*

The results of the chronosequence applied the NDVI values show an increase of NDVI with time (Fig. 5; $R^2$.=0.49). Therefore, the NDVI is correlated with tree height (*appendix*, Table A10; cor. = 0.62) and SOC (cor. = 0.51). However, it follows a different temporal trend when fitting a logistic model to the chronosequence timescale (Fig. 5). Therefore, an adjustment should be made when predicting tree height/litter input with NDVI.

The NDVI timeseries derived from the time period of 2000 to 2018 for the 50 plots barely shows an increase in vegetation activity as a result of the restoration practices (Fig. 6). Instead, the yearly NDVI shows the same, slightly delayed timing of peaks, as the yearly average precipitation. The relation with NDVI and precipitation can also be seen back in the PCA (Fig. 3), where winter precipitation is explaining similar variance in the dataset as the NDVI parameters does. This implies the effect of precipitation on the NDVI is dominating the increase in NDVI due to tree growth, at least for most of the plots. Other possible causes of the weakly observed NDVI trend can be explained by a situation where the observed greenness of the vegetation does not change as a result of its growth. This situation could have been caused by livestock grazing decreasing the present observable leave area while unobservable roots and stems are unaffected. The lack of a clear trend in the NDVI as a result of ecological restoration makes it almost impossible to derive a prediction of the future NDVI and therefore tree height further in time.

### 3.2.4 Relief

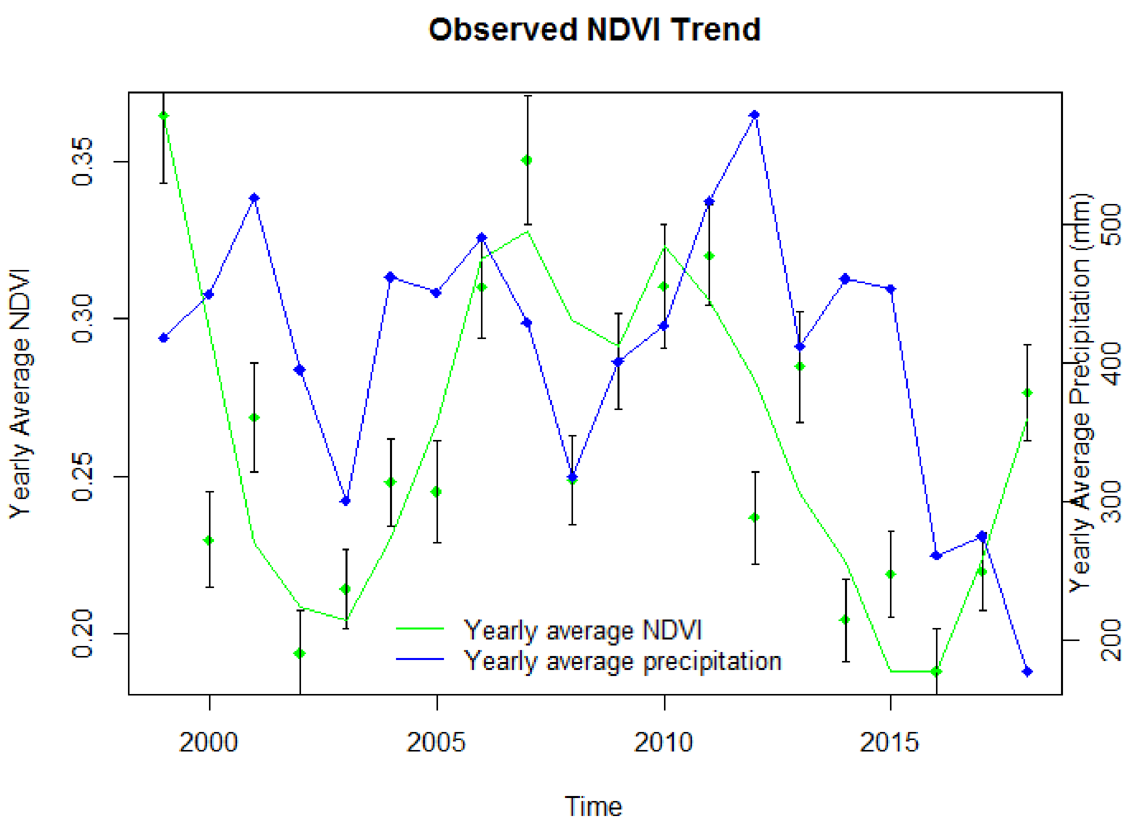


*Fig. 6 The yearly mean NDVI timeseries, averaged for 50 plots including the yearly average precipitation.*

The relief of the area is potentially an important factor in determining SOC due to its relationship with erosion. The effect of relief can well be seen in the PCA (Fig. 3), where slope explains both Dim1 (*appendix*, Table A11; cor.=0.72) as well as Dim2 (cor.= 0.32). Opposite to the expectations of the effect of slope, a steeper slope is observed to have a positive effect on SOC (*appendix*, Table A10; cor.=0.50). This effect can be understood when looking at the topographic wetness index (TWI), which is similarly describing the variability in the PCA, but oppositely directed. This implies that the effect of a lower slope contributing to the accumulation of water is far more important than a steeper slope causing instability. When also considering the role of stone fraction is this process, one can explain a situation in which the high stone fractions cause stabilization of SOC at steeper slopes is this region. At the accumulative positions there is both erosion of the top layer, but also deposition of carbonates causing higher pH (cor.=-0.19) at gentle slopes. Consequently, Duric horizons are formed, which are found at the most degraded plots in this region. Eventually the TWI and Slope can be used interchangeable due to their similarities in the PCA caused by their similar origin.

The analysis of the contribution of slope in driving SOC development is limited due to the influence of slope on tree height. *Spekboom* is naturally growing on steep slopes causing a better establishment at these positions. Larger *Spekboom* height, as previously observed, means higher SOC stocks. Through this mechanism the slope is positively contributing to SOC concentrations, while it has not so much to do with its physical effects on erosion. Keeping in mind also the other interrelated variables previously described, it is hard to quantify accurately the contribution of slope to SOC development as a result of ecological restoration

### 3.2.5 Climate

The results showed climate factors (solar radiation, summer precipitation and heavy precipitation) are mostly explaining the second dimension of the PCA (Fig. 3), both negatively as positively. Albedo is closely related to these climatic parameters in its role of determining soil temperature. The land surface temperature (LST), in the contrary, is not situated close to those variables since it is a combination of the opposite effects of albedo and solar radiation on temperature. Due to this effect the LST becomes an important variable in describing dimension 1 (*appendix*, Table A11; cor.= -0.70) and therefore SOC (*appendix*, Table A10; cor.=0.59).

Precipitation variables are mostly located at the negative axis of dimension 2 (Fig. 3). What is noteworthy is the difference between summer and winter precipitation in explaining variability in the timeseries. Winter precipitation is weakly explaining any variability, but is directed perpendicular to the vegetation variables, which relates to the previously shown effect on NDVI. The summer precipitation however is mainly contributing negatively to SOC (cor.=-0.16) as well as heavy precipitation (cor.=-0.25). Both are explaining erosional effects and have similar effects as the TWI on SOC. The interrelation between variables holds for climatic variables such as rainfall, soil moisture and temperature, which will be included in the model, but its variables not always showing such high correlations with SOC. With their effects on erosion and mineralization, respectively, climate variables impose a theoretical importance (Jenny, 1991), which cannot be ignored.

### 3.2.6 Parent material

Since there are 5 different types of parent material present in the study area, one would expect it causes differences in soil development. In the PCA this soil forming factor, represented as rock type, is just weakly represented in Dim1 (*appendix*, Table A11; cor.=0.05) and therefore mainly situated in the third dimension (cor. = 0.45). The weak contribution to the prediction of the variability of environmental parameters is most likely because of the categorical character of parent material as a variable. Nevertheless, when comparing the development of SOC as a result of ecological restoration between the different parent materials, the effect of parent material is visible.

Both the lines (Fig. 7) categorized by Shale and Quartizitic sandstone show higher average levels of SOC, while Quartizitic Sandstone and Conglomerate lines show faster buildup of SOC. These differences could either be caused by 1) random variability and uncertainty represented by low accuracy (low P-values). 2) the distribution of texture classes, where a higher clay content stabilizes organic carbon and sandy fractions stimulate the buildup 3) a possible gradient in grazing pressure caused by remotely positioned parent materials relative to the grazing areas 4) a difference in albedo and emissivity between parent materials causing differences in LST and therefore mineralization rates 5) higher stone fractions preventing heavy erosional events during extreme rainfall.

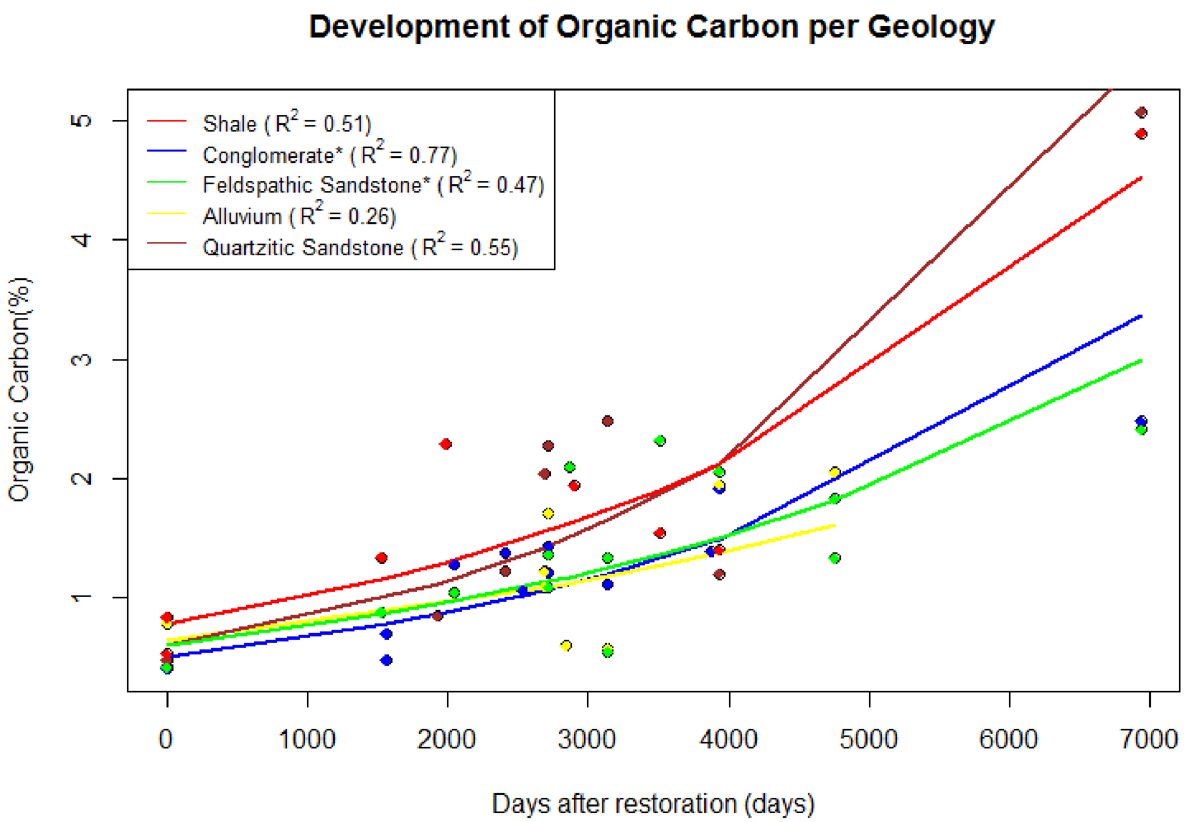


*Fig. 7 Development of organic carbon separated for the different parent materials including shale, conglomerate, feldspathic sandstone, alluvium and quartzitic sandstone.*

The second hypothesis is supported by a significant difference (*appendix*, Fig. A14; P=0.02) in mean clay fraction of Shale and Feldspathic Sandstone (*appendix*, Table A11; 15.6%) compared to Alluvium, Conglomerate and Quartzitic Sandstone (13.0%). Where Quartzitic Sandstone in its case has a sand fraction of 76% (*appendix*, Fig. A15), which is almost significantly different (P=0.06) from other Parent materials containing on average 71% of sand. A high clay fraction could be beneficial in the stabilization of organic matter through aggregates (Six et al.,2004). Thereby, the influence of aggregates on soil functioning is known to be more prominent for Shale parent materials than for the quartzites (Barskale and Itani, 1989), where other parent materials in the area belong to. Stabilization of SOC and microbial communities in the aggregates during the degradation process could cause the starting point for these plots to be substantially higher in terms of microbial communities' diversity and magnitude (Zhang et al.,2010) and therefore cause divergence in its development (Walker et al., 2010).

The influence of texture can also be understood by the fact higher sand/lower clay fractions might be beneficial when looking to the main driver of turnover according to (Coleman and Jenkinson, 1995). Turnover is driven by the fraction of clay present, which negatively contributes to litter turnover in water limited systems (Rietkerk et al.,1997). Consequently, the observed faster increase of SOC in case of the Quartzitic sandstone parent material can be explained by a higher mean sand fraction.

Both the two mechanisms in the parent materials are caused by physical properties, however, the effect is not necessarily evident when looking to the environmental variables available. Texture classes between parent materials differ heavily with a standard deviation of up to 30% (*appendix*, Table A13;), but also the contribution of aggregates is very much dependent on the genesis of the parent material (Harris, 2003), which is beyond the scope of research.

### 3.2.7 Other

The third hypothesis of parent material as driver of SOC is strengthened by looking at the effect of grazing on the development of organic carbon. Grazing does affect the tree height by reducing growth and keeping the height under the 70 cm length line, at least for the available observations (Fig. 8). The same negative effect of grazing is visible in the PCA where livestock has a negative contribution to Dim1 (*appendix*, Table A11; cor.=-0.46)

However, other observations seem to contradict the positive effect of livestock derived from the PCA. In fact, livestock grazing seems to have an average positive effect on SOC for the first 3000 days of soil development, as can be seen from the grey and black line in Fig. 8. The large diversity present, is not surprisingly, caused by differences in parent material. The lower points in the "livestock inclusive" graph do represent a parent material, which is very close to road and farmland, while the more remotely situated Shales and Quartzitic Sandstones are higher on the graph. These observations can sketch a situation where a somewhat lower grazing pressure does not cause the dieback of vegetation, but microbiological life to be stimulated due to the presence of livestock (Raiesi and Asadi, 2006). However, the contribution of microbial activity in relation to grazing pressure has not been measured and cannot be taken into account in the current prediction of SOC, nor the quantification of grazing pressure be compared to SOC measurements.

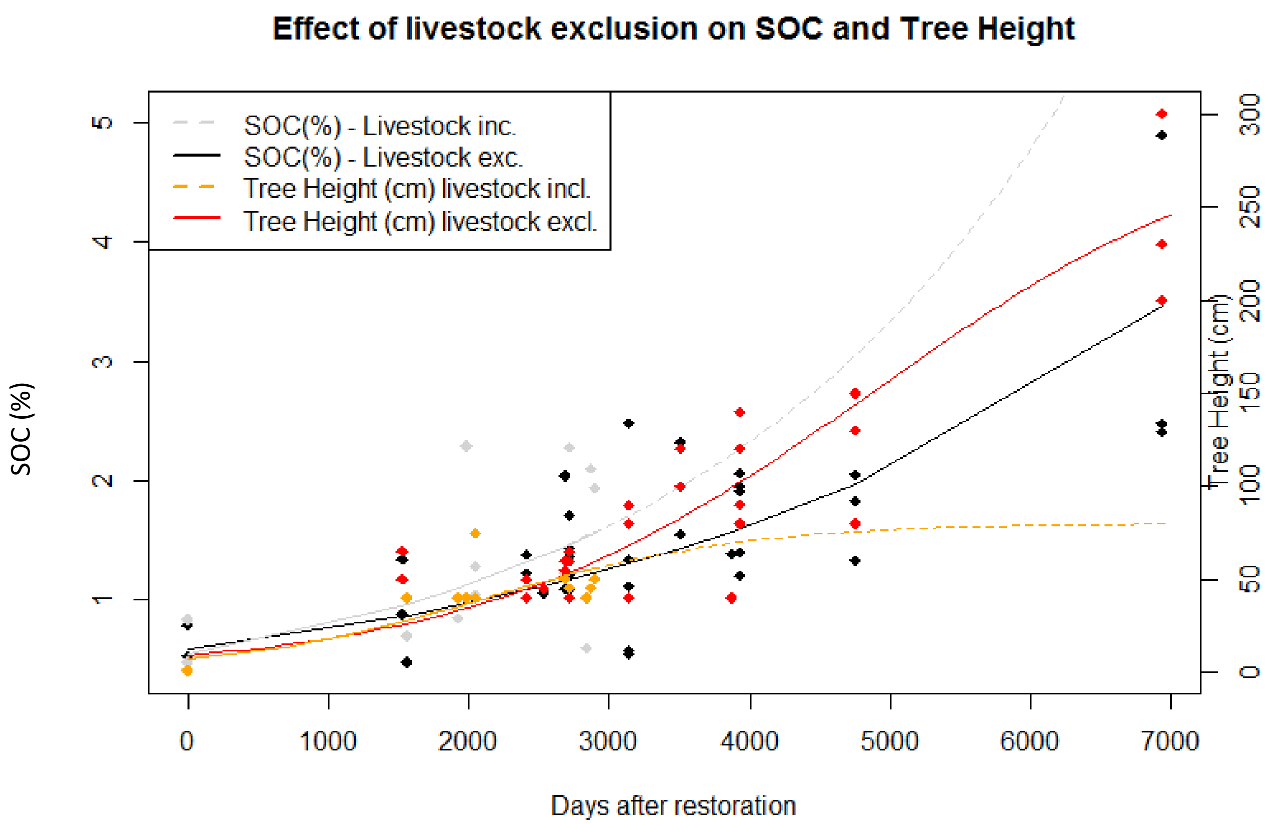


*Fig. 8 The development of tree height and SOC for days after restoration of 0 to 7000. This includes a separation of SOC*

The contribution of drivers in this research can be summarized when looking to the points associated with the minimum scenario in section 3.2.2. In comparison with average statistics of plots, one can observe reduced tree height (*appendix*, Table A12; 40cm), gentle slopes (6 deg.), specific parent materials (Conglomerate, Feld. Sandstone and Alluvium) and low stone fractions (20%). Obviously, reduced vegetation activity has caused these plots to barely have developed compared to degraded plots. Plot 6 has inclusion(0), but extra accumulation, which could have caused extra degradation. Other plots were exposed to livestock grazing, which could have been the cause of the reduced tree growth. However, livestock grazing has not been observed as being dominant over environmental conditions (section 3.2), although a gradient in grazing pressure was associated with parent material locations. Assuming grazing pressure was dominant in the past, and physical conditions at water accumulative plots are less favorable, a new situation can be sketched. Thereby, the combination of both erosion and grazing pressure driving extreme degradation at these plots, have been leading to conditions characterized by Duric horizons and or absence of organic carbon and microbes. Restoration activities were not successful, at these plots, in introducing vegetation nor increasing carbon stocks due the alternative degraded conditions. This implies initial conditions tend to dominate at an early stage and therefore its drivers are important to consider.

## 3.3 Mechanistic model

### 3.3.1 Model calibration

In order to quantify the current soil development state, a prediction of the SOC stock was established using a mechanistic model. Based on the least RMSD generated by including multiple different proxies in the model runs, a set of proxies and constants has been chosen to relate to the observed SOC stocks (Table 6). The proxies that resulted in the lowest RMSD in the model, did confirm the observed scale of contribution of those proxies in describing the principal components. More specifically, variables were chosen which did show the highest correlation with SOC in the first dimension of the PCA. The exception is found in the prediction of temperature and soil moisture, which are based on the model results which are not most accurately done by using respectively the LST and soil moisture. Instead using respectively solar radiation multiplied by albedo, and summer precipitation to predict the mineralization rate, gave the best results regarding least RMSD (Table 6). Most probably because soil moisture and LST are a product of the NDVI and therefore its variability is already represented in the first equation of the model. Moreover, the tree height was predicted on catchment scale by fitting a regression model to the relation of tree height to NDVI (Table 6)

*Table 6 Overview of the implementation of proxies compared to two different model calibrations (Hendriks and the current research)*

| Process | Proxy Hendriks 2018 | This research |
|---|---|---|
| **Litter input** | Tree height | NDVI ~ Tree height |
| **Turnover rate** | Clay (%) | Rocktype ~ Clay (%) |
| **Mineralization rate** | Temperature | Temperature ~ albedo*solar radiation, Soil moisture ~ summer rainfall |
| **Erosion** | Slope | $Slope^{-1}$, heavy rainfall |

The calibration of the model resulted in a RMSD of 2972 kg/ha, which is about 30% of the total SOC stock (Table 7). The model's PCC was 0.82 when comparing observed and modelled SOC stock. On average the litter input was 1362 kg/ha, which is about 50% of pristine *Spekboom* yearly litter input (Mills and Cowling, 2010). In addition, the modelled mineralization rate of 0.01 corresponds to earlier findings (Hendriks, 2018). Compared to a model simulation with the original parameters of the mechanistic model, the newly developed model performed better in explaining the variance, but also showed smaller residuals (Table 7). That might be due to the fact the original model underestimated the erosion rate, because the slope was wrongly inversibility related to erosion in this model. This causes the tendency of the model to reject the erosion process. The new model performed better in explaining this and other processes, therefore showing similar magnitudes for process values as reported by Hendriks (2018) (Table 7). However, the relative RMSD (RMSD/SOC stock = 0.3) was much higher than originally reported for another study area (0.049), which is related to the large residuals generated by the current model.

*Table 7 Results of the mechanist model calibrations, including comparison between the original Hendriks model, the Baviaanskloof Hendriks application and the current model (Koster) on the different processes and model statistics.*

| Variable/process | Value Hendriks Cantabria (Spain) | Value Hendriks Baviaanskloof | Value Koster Baviaanskloof |
|---|---|---|---|
| **RMSD/SOC stock** | 0.049 | 0.526 | 0.305 |
| **Pearson correlations** | 0.44 | 0.64 | 0.82 |
| **Average Balance (kg/ha)** | -821 | -3.27 | -3E-15 |
| **Litter input (kg/ha)** | 18000 | 6556 | 1362 |
| **Turnover rate (-)** | 0.16 | 0.16 | 0.16 |
| **Mineralization rate (-)** | 0.015 | 0.05 | 0.01 |
| **Erosion rate (%SOM)** | 0.7 | 4.5E-5 | 0.8 |

The largest residuals were found in the lowest and highest range of the modelled SOC stock values, where the plots with equilibrium conditions (pristine and degraded) are located (Fig. 9). This is normal when considering the general behavior of models in the upper and lower tail of the prediction. Nevertheless, this is not the only factor that possibly explains the observed residuals involved with the modelling methods. The unpredictability of the presence of livestock is also an important factor contributing to the residuals. In fact, plots without livestock exclusion showed larger residuals in their prediction (Table 8). This can be related to the effect of livestock presence on small scale biological soil processes, as explained in previous sections. These microbiological processes were not taken into account in the model. The other residuals can be partly be explained by microbiological processes related to the chronosequence scenarios explained in section 3.2.1. Additionally, the prediction of clay fraction per parent material introduced some uncertainties, since there was a significant standard

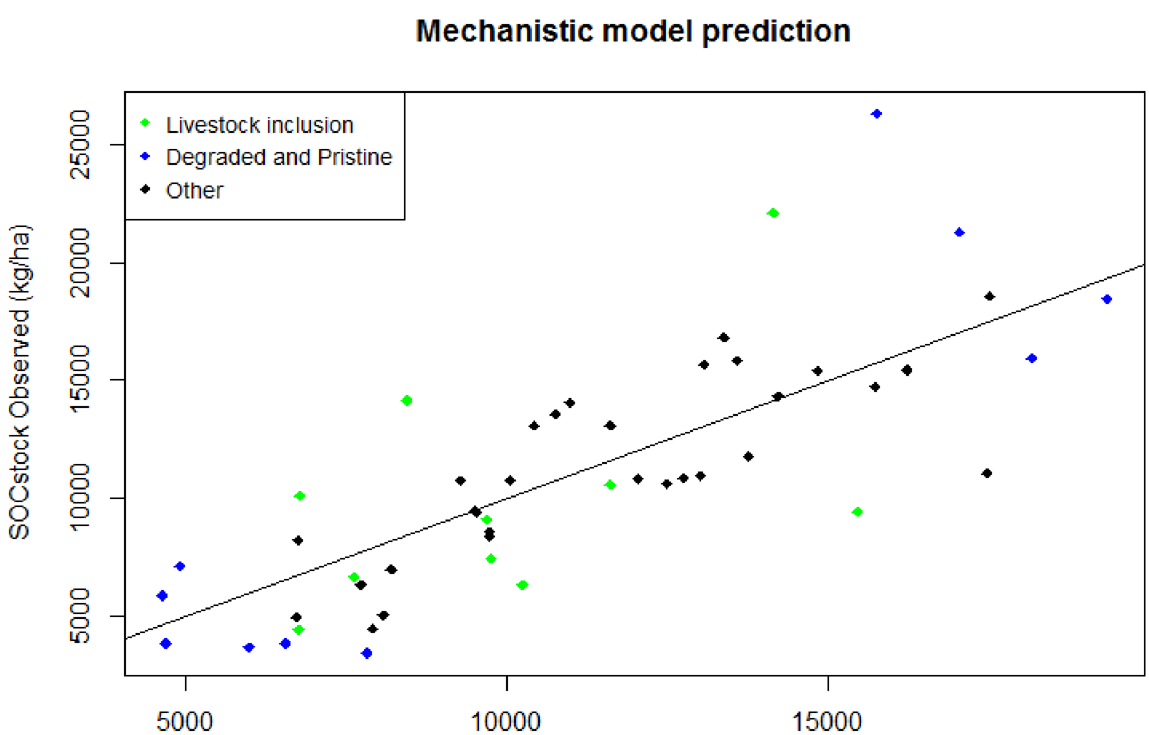


*Fig. 9 The prediction of SOCstock compared to the observed SOCstock. Livestock included, degraded and pristine plots are indicated.*

deviation associated with this assumption and the texture classes did not follow a linear effect on SOC build up (section 3.2.6).

Moreover, the limitations of the model regarding the stability of the system could have been a factor that influenced the model predictions. The model might have failed in explaining equilibrium conditions (degraded and pristine) due to the fact its parameters were applied to a dynamic system, while the original model was designed for steady state conditions. The observation of this limitation is strengthened by the observations of a decreasing relative error with time (*appendix,* Fig. A16). Most likely this occurs because equilibrium conditions were approaching proceeding with time. Another explanation for the decreasing relative error with time is the dependency of the temporal development on initial conditions, which were not taken into account in this research, but identified as important (section 3.2). This phenomenon also strengthens the hypothesis of alternative states, especially when looking at the location of the highest relative errors in the graph (*appendix,* Fig. A16), which corresponds to the points associated with degraded conditions in section 3.2. These degraded points were also the most important residuals of the chronosequence trendline (*appendix*, Fig. A16).

Another explanation for the poor prediction is related to the NDVI. The index has been shown to be very variable in time, with no clear trend. Therefore, the choice of the time period for the NDVI calculation determines the magnitude of the constants in the calibration of the model. In fact, using the NDVI for future analysis of soil development with the same mechanistic method, requires a new relationship between NDVI and tree height, since the NDVI is very variable. Combining this sensitivity for the time period and the already described uncertainty in the NDVI tree height prediction (section 3.2.3), the prediction of tree height becomes highly uncertain.

Not only the spatial distribution, but also the temporal variability was important for the derivation of the current soil development status. First of all, the temporal variability in rainfall was barely taken into account due to the fact only the time period of 2015 until 2016 was available for almost all stations. The difference of the average of these years with the all-time average has potentially caused a mis-calibration. Moreover, not all *Spekboom* vegetation on the plots have been living in the same amount of years, while the same time period was used for all plots. Given the fact the most extreme situations might have an impact on the vegetation development, this approach might have been causing uncertainties since some extreme cases have possibly been excluded. Keeping in mind, the rainfall data was interpolated between stations of which each had its own unknown uncertainty, the derivation of rainfall statistics for the designated plots were limited in accuracy. This uncertainty cannot only be dealt with by looking only at the large-scale variability, since the large-scale variability in rainfall has also its uncertainties.

The climate variables were also doubtful in their contribution to the processes included in the model. Due to the fact temperature influences mineralization, tree growth, litter decomposition and water availability (Hevia et al., 2013, Alvarez and Lavado,1998, Homann et al.,1995), its contribution as well as the contribution of precipitation was very much dependent on the individual case (Yang et al.,2008) (Hevia et al., 2013). This theoretical understanding of climatic variables having multidirectional effects is in line with the observation of a low correlation between climate variables and SOC (*appendix*, Table A10). An exception would be the relation with heavy rainfall and number of dry days, which are theoretically single directed (Jin et al.,2009).

Finally, the stone fraction, soil depth and bulk density were necessary to include in the model to convert SOC fractions to SOC stocks. However, these variables were not accurately established and therefore caused uncertainty in the SOC stock derivations. In case of the soil depth this means stocks were reduced where only 7 cm of soil was sampled, while this soil depth was not certain in the first place. The choice for 10 cm depth sampling has also its implications on the uncertainty related to below and upperground exchange. More specifically, the limited soil depth sampled did not cover the full SOC stock present, which is most likely exceeding beyond 10cm depth (Zhang and Hartemink, 2017). Consequently, the variability in below-upper ground exchange might be relatively large relative to the limited SOC stock sampled.

### 3.3.2 SOM map and accuracy

The generated SOC map shows SOC fractions between 0.2% and 5% (Fig. 10), which are similar to the range of sampling conditions. The spatial distribution of SOC is in line with the hypothesis that the natural areas contain the highest SOC fractions and are closer to pristine (Fig. 4). Moreover, one can see slight improvements in soil functioning compared to the majority of the area when looking to the area situated in between both natural areas. These slightly improved areas are known to be restored between 2005 and 2016 and are close to turning pristine in the future (Fig. 4). However, these higher SOC areas tend to follow drainage patterns, which might have caused higher NDVI and therefore higher SOC. Moreover, the small differences between the restored and other areas are not larger than the standard error of the prediction (Fig. 10) and therefore not certain.

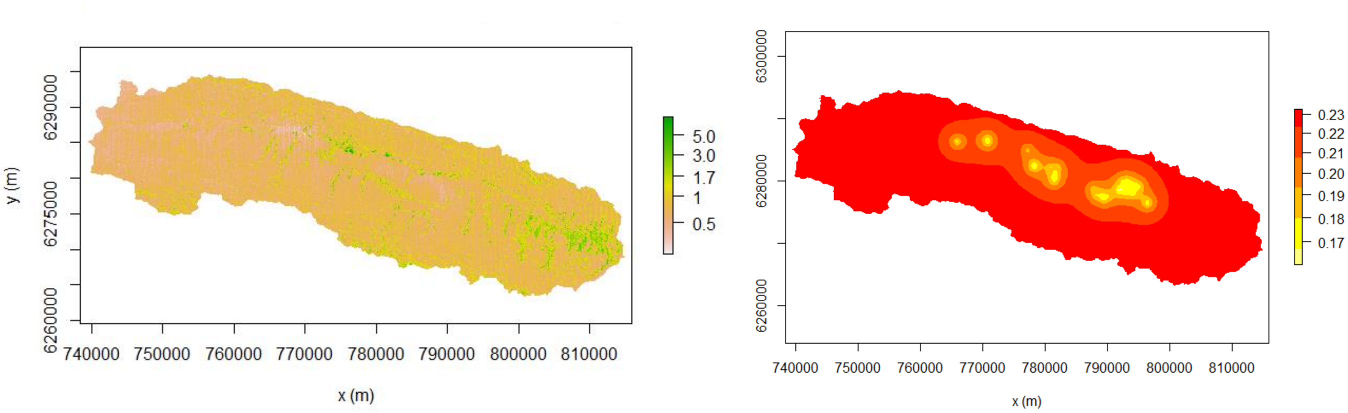


*Fig. 10 left: predicted map of SOC for the Baviaanskloof catchment, right: standard error of the predicted map of the Baviaanskloof catchment.*

Another uncertainty in the derived maps can be related to the sampling method used. More specifically, the SOC map is based on SOC samples taken within the influential tree range. What is known however, is that the SOC outside of the tree range does not follow the same development (Mills and Cowling, 2010). This means only using the current sampling sites for spatial mapping, creates an overestimation of the SOC content. This overestimation is barely present for degraded and pristine areas, since their assumed tree cover is close to 0% or 100%, respectively.

When looking to the generated uncertainty of the maps, the standard error of the prediction ranges between 0.17% and 0.23% SOC (Fig. 10), which is an average relative standard error of about 30%. The largest standard error is found for the areas closer to the corners of the catchment, which is explained by the absence of measurements in those regions. Since the lowest SOC fractions are also situated in those areas, the relative error could potentially exceed 100%, which makes the values highly uncertain, especially when looking to small scale spatial distribution in those corner areas. Consequently, the standard error it generates on the actual maps is acceptable in the absolute sense, but represents uncertainties associated with the SOC buildup of tens of years (section 3.2.2)

For some of the uncertainty factors mentioned in the previous section, the small-scale variability is unknown, while the average values and large-scale distribution is rather well known due to the derived relationships and existing maps. Therefore, the derived maps of prediction and standard error of the model, on a large scale, can be justified regarding these uncertainty factors in the calibration of the model. However, the misfit caused by absence of an accurate representation of soil physical processes and livestock effects has the potential to significantly influence the large-scale spatial distribution of soil development. The reason for that can be found in the understanding of the causes for differences in these two factors, which source is found in both the parent material as well as a gradient in grazing pressure, which might not be randomly distributed. Therefore, there is a potential for a large-scale misfit in the soil development maps, both in the distribution of the prediction results, but also regarding the standard error distribution.

## 3.4. Comparison

### 3.4.1 Spatial

When comparing the generated map of both mechanistic models with the resulting maps of SOC over the whole Baviaanskloof catchment, the differences are distinct. The order of magnitude in SOC values is similar (Fig. 11; 0.4%-6%).

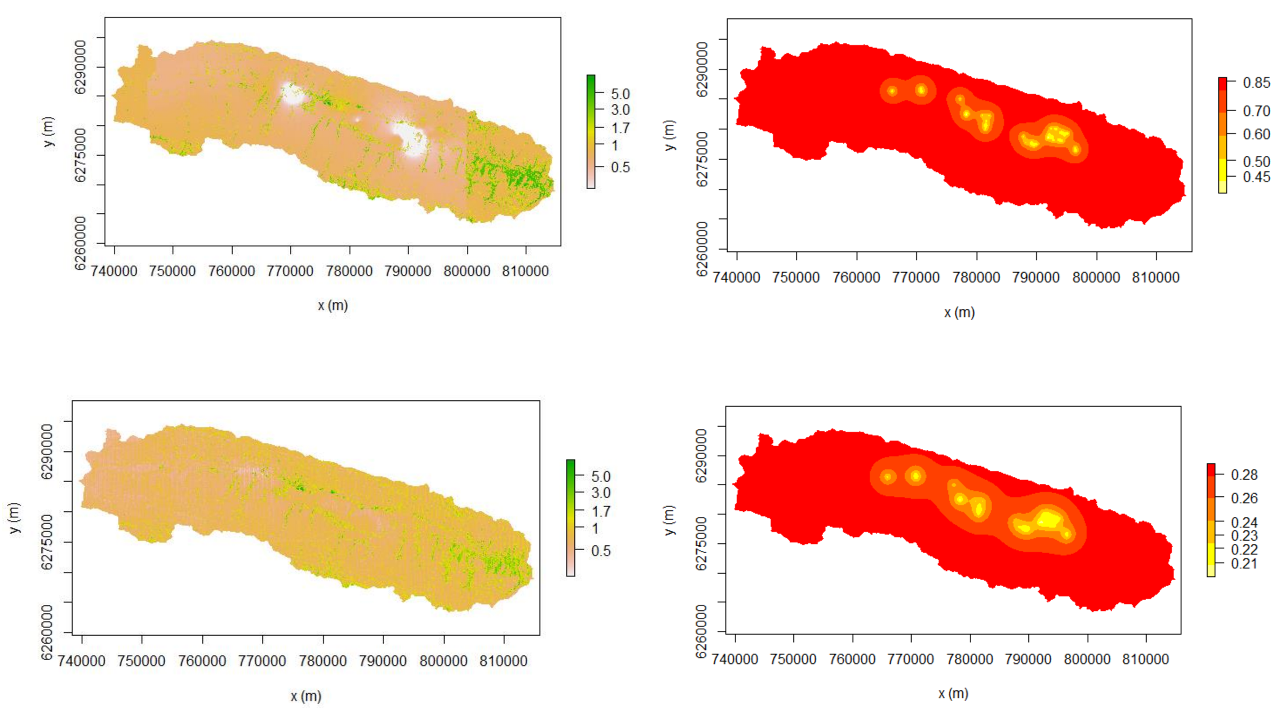


*Fig. 11 topleft: generated SOC map with the Hendriks/Stoorvogel model, topright: generated uncertainty map with the Hendriks model, bottomleft: generated SOC map with the Koster model, bottomright: generated uncertainty map with the Koster model..*

The spatial distribution of SOC however includes distinct differences, which are related to the large residuals generated by both models shown in section 3.3.1. Moreover, the relation between slope and SOC is differently directed in both models, which also causes distinct spatial differences between both maps. The average standard error of the newly developed model (Fig. 11; ~0.25 SOC%) is substantially lower than the original (~0.70 SOC%), which sources from the difference in variograms (*appendix,* Fig. A17 and A18). This ultimately indicates a more accurate spatial prediction of the current model compared to the mechanistic model with the original set of proxies.

### 3.4.2 Temporal

*Tool setup*

Based on observations in section 2.2.2 two carbon sequestration tools have been setup. Based on average land cover classes, both the baseline and 25% cover scenarios were established (Table 8). Where "pristine grasses" were assumed to represent the observed 14% shrubs in the tool, the 11% grasses were implemented as "moderately degraded". "Severely degraded" corresponds with observed bare soil, whose fraction is equal to the remainder of the area.

*Table 8 Overview of the implemented land use cover scenarios using the Carbon Benefit Project (CBP) tool.*

| Land cover | Baseline | CBP 100% cover | CBP 25% cover |
|---|---|---|---|
| **Forest (ha)** | 0 | 113946 | 28486 |
| **Total grass (ha)** | 113946 | 0 | 85459 |
| **Grass pristine/shrubs** | 17100 | 0 | 17100 |
| **grass moderately degraded/grass** | 12540 | 0 | 12540 |
| **grass severely degraded/bare soil** | 84306 | 0 | 55819 |

*Tool comparison*

The carbon benefit project tool reports annual carbon sequestration between 436370 tons of CO2e.year$^{-1}$ and 1694000 tons of CO2e.year$^{-1}$. This equals an interval of 0.06-0.24 (%) SOC.year$^{-1}$ (Fig. 2). When applying this to the previously constructed chronosequence one can find the point level SOC curves included in the range simulated by the GBP tool. This simulated range has an average relative error of the mean curve of more than 50%. This would be the relative error in a situation one does not have any knowledge about the tree cover or effects of grazing on the organic carbon development.

Both the CBP and RothC 100% tree cover scenario show similarities with the maximum scenario in the chronosequence curve. This makes sense keeping in mind the maximum scenario. This scenario is expected to describe an almost full potential of *Spekboom* stimulating SOC development. This, in combination with the sampling design dedicated to represent full tree cover, aligns well with the simulated model scenarios, since these models were implemented to simulate full cover and are not limited by any factor.

The 25% cover curve is similar to the minimum chronosequence scenario in the curve (Fig. 12), but slightly overestimated. This means, keeping in mind close to 100% tree cover is sampled, in that situation the *Spekboom* has only used less than 25% of its sequestration potential.

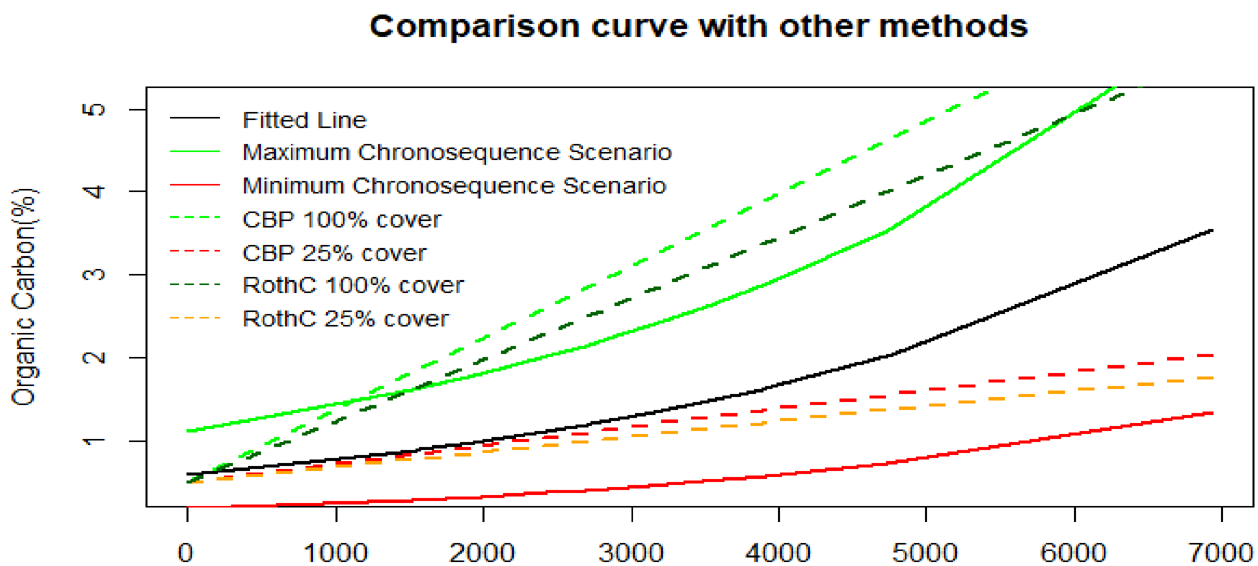


*Fig. 12 The development of SOC between 0 and 7000 days after restoration. This includes the fitted curve in section 3.2.2, its associated scenarios, the CBP 100% and 25% scenarios and the same scenarios ran with the RothC carbon model.*

*<u>Limitations</u>*

The model scenarios differ from the chronosequence scenarios in their temporal behavior which is linear in contrast to the exponential behavior of the SOC chronosequence. Considering the theoretical framework, which explains a mechanism of a positive feedback involving the development of SOC (Suding et al.,2004), this assumption might be incorrect. This could potential mean an initial overestimation followed by an underestimation of the actual SOC buildup. However, the errors associated with this assumption do not seem to exceed 1% of SOC.

Since sampling was biased in the sense that only surviving *Spekboom* was chosen to sample and the sampling was done right next to the tree, the maximum chronosequence scenario was not likely representable for the full study area. This is supported by fact this pathway is similar to the 100% tree cover scenarios simulated by RothC and CBP, which is not likely when looking at the actual coverage of 25% on average for the sampling sites. The same holds for the minimum scenario which offers still an overestimation of the carbon sequestration by assuming 100% survival rate. The actual area wide sequestration is dependent on the lower remaining tree biomass, which is still unknown in magnitude.

*<u>Tools</u>*

The range of values generated by the established models corresponds with a tree cover difference of 75%, while the chronosequence was constructed for a relatively constant tree cover scenario by sampling close to trees. The fact that the observed differences had to be mimicked by changing tree cover implies the most important drivers causing the variation in the actual development of carbon, are not taken into account in both models. An exception exists for the clay fraction included in RothC, but the actual contribution of clay to SOC is very doubtable, given both the correlation with SOC and the graphic results in this research. The failure in explaining the variability in case of the carbon models implies there is no valid factor that can be changed in the model to mimic the range of temporal development, and therefore only an average be assumed.

A corresponding average tree cover scenario both RothC and GBP would lie somewhere between 100% and 25% cover, given the graphs in section 3.4.2. In contrast, the actual coverage sampled is 100% since sampling is done at actual trees. The fact the carbon models need lower tree covers for mimicking observations represented by 100% scenarios implies a strong overestimation of the carbon sequestration by the established models. Moreover, to simulate the actual SOC values one needs to know the actual tree cover, which is very hard to establish. This is due to the lack of available indicators of tree cover (section 3.2.3) and the scarce knowledge on the success rate of establishment. Even if this success rate is known, a failure of growth does not necessarily mean that the soil is not affected. For instance, in this research a bunch of plots has been sampled where livestock has caused a hold to tree growth, but where there is an indication soil organic carbon has developed substantially. Such effects are not been taken into account in the models and make a difference between the maximum and minimum chronosequence sequestration scenario, which do differ in such magnitude that these differences can cause a mismatch between expected and actual circumstances using the available carbon models.

# 4. General Discussion

## 4.1 Environmental drivers

***<u>SOC</u>***

The first questions in this study sought to determine the most important drivers of soil development as a result of ecological restoration. The results indicate that SOC can be considered the most important indicator of soil development during ecological restoration. This finding is supported by SOC being the most important factor in deciding the degraded-pristine transition in the PCA (Fig. 3). Consequently, SOC is more important compared to other well-known indicators of ecological state transitions.

Although this study has successfully applied SOC as indicator for soil development, its application has certain limitations that need to be discussed. Firstly, an inappropriate simplification of reality might have been assumed, because SOC as a single indicator is being assumed valid for the complex soil state. This assumption might not be appropriate since previous studies have considered multiple environmental indicators when assessing the soil development state. They did that in their successfully application of a soil health assessment strategy (Idowu et al., 2008; 2009). However, other studies contradict by noting that the timescale considered is more important for deciding the best indicator of soil development (Carpenter and Turner, 2000), because response times of the diverse indicators of soil development are different. Applying this response time theory to the current observations, teaches us that the decadal scale that applies here allows SOC changes to be accurately visible. This ultimately implies that the current study assumptions of organic carbon as an appropriate soil development indicator are most likely justifiable and therefore do agree with the conclusions of Constantini et al. (2016) and Muñoz-Rojas et al. (2016).

A second limitation of the use of SOC as indicator for soil development arises from the use of total SOC instead of different SOC fractions. The current analysis of total SOC cannot detect changes in SOC composition, while these changes can have implications for soil development (Constantini et al., 2016). In fact, differences in SOC composition can be present in this region, since it is known that the SOC composition is sensitive to the occurred soil degradation (Franzluebbers, 2002). This would imply that some important changes in soil development, might not have been detected by the analysis of total SOC. However, the impact is expected to be minor, since the majority of samples is situated in the same state of degradation (Fig. 3) and therefore expected to show a stable SOC composition. Although SOC has concluded to be an appropriate indicator of soil development, it is considered to be multidisciplinary and thus complex in its behavior. Therefore, understanding the behavior of the system regarding environmental drivers requires a focus on the drivers involving the individual soil forming factors, which are considered less complex in their effects on the soil state.

*Tree height*

Vegetation height is representing the organisms in the development of soils and is an important factor during soil development as a result of ecological restoration. The results have shown that it correlated better to SOC ($R^2$=0.88) compared to other environmental factors. This observed importance of tree height in deciding soil functioning agrees with the theoretical framework described by Suding et al. (2004), where it is raised as the main source of organic carbon input via litter. Also, Andel and Grootjans (2005) stated the vegetation is essential in increasing soil stability, litter input and microbial activity and thus important for increasing SOC. Not only vegetation is known to be important for building up SOC, but also the other way around, SOC is an essential factor promoting tree height growth by increasing nutrient and water availability (Suding et al, 2016). Due to this mutual dependence between SOC and tree height, their correlation between has turned out to be relatively large ($R^2$ = 0.88) compared to other drivers. The observed strong correlation increases the indication vegetation should be considered as a simplified indicator of the effectiveness of restoration, such as SOC, instead of the main driver of its success (Bullock et al., 2011, Chen et al., 2007). Since vegetation cannot be defined as the main driver, other factors should be considered.

*Other drivers*

Interestingly, parent material is a driver which has a considerable effect on SOC development. This study showed parent material be the factor mainly determining the temporal divergence of SOC development (Fig. 7). This has been theoretically explained in section 3.4 by relating parent material to soil traits such as aggregate type and magnitude (Barskale and Itani,1989), texture class distribution and stoniness. When linking these soil traits with well-known SOC drivers such as microbiology (Zhang et al.,2010), litter turnover (Rietkerk et al.,1997) (Chen et al., 2016) and erosion, respectively, parent material can be related to SOC development. In reviewing the literature, no data was found associating parent material to drive the development of soils during ecological restoration. This absence might have been caused by the weak understanding of the role of microbes in restoration ecology (Singh and Gupta, 2018, Zhang et al.,2016) associated with parent material. Another possible explanation can be the lack of an observed significant effect of parent material on SOC. These observations might lack because the generally high variance in texture classes per parent material (section 3.2.6) are causing the non-linear effect of texture differences on SOC (Rietkerk et al.,1997, Chen et al., 2016) to be hard to trace back to parent material. Despite, the absence of the relation between parent material and SOC in literature, this study indicates that parent material may be directly associated to SOC and therefore counts as an essential driver in the development of soil as a result of ecological restoration

*Limitations*

The analysis of the most important driver of soil development involved some uncertainties. Many of the analyzed drivers did have both a positive as well as a negative effect on processes influencing SOC. Moreover, they have potentially influenced other drivers as well. This phenomenon could have lowered the significance of correlations between drivers and SOC, made the correlation absent or made it different in direction. The first relationship affected by the described phenomenon is those of slope and SOC. Slope behaved counterintuitively by stimulating SOC, instead of decreasing it. This behavior, which has been identified in other areas before (Zhang and Dong, 2010, Jian-Bing et al., 2006) most likely occurred in this region due to correlation of slope with vegetation dynamics and water accumulation. Due to the counterintuitive behavior of slope, the theoretically large importance of slope processes for erosion and therefore SOC in this region (Van Luijck et al., 2013), can hardly be confirmed with practical evidence. The second hard to establish relation was related to the climatic variables in this study (section 3.2.5). These variables were hardly correlating with SOC but are known to be important for soil functioning (Hevia et al., 2003). An underestimation of the importance of climatic drivers for SOC development might have occurred. The last example refers to the effect of grazing pressure on SOC development, which was observed to be weakly positive. Although some positive influences of grazing pressure on SOC are reported (Skarpe, 2000, Raiesi and Asadi,2006), grazing of livestock is considered to decrease SOC, because it causes a hampering in vegetation growth. Therefore, grazing theoretically largely reduces litter input of leaves and ultimately SOC development. The described uncertainties caused by the complex effect of variables are worsened by the limited sampling size. Limited sampling size namely, causes a higher relative impact of small effects of a different environmental variable on the overall relationship between SOC and the analyzed variable. This is also the reason that the relations between different parent materials and SOC could not be significantly proven, while the importance of parent material is substantial in this study.

The final uncertainty in establishing the main drivers of SOC development after ecological restoration is related to the complex system faced. The results of Fig. 4 and the indications of SOC's dependency on initial conditions (section 3.2.2), helps to indicate that past conditions are still influencing the current soil status and thus SOC values. This agrees with the findings of Anand (2004), which describe the process of ecological restoration by applying the complex system theory. This theory is thoroughly described by Runge et al. (2015). Analyzing this theory for the current case means initial conditions of SOC drivers do have a profound effect on the currently measured SOC values. More specifically, legacy of SOC itself, rainfall, but also texture class might still have their impact on the current SOC values. This means that the current SOC values are also influenced by remnant values of drivers, while the current values of drivers might be very different. This complex system behavior causes the suggested importance based on the correlation of drivers with SOC to be different than their actual physical importance. Ultimately the complex system behavior hardens the identification of the main drivers of SOC, which causes stable parameters in the temporal domain, like parent material, to be easier identifiable and therefore potentially overestimated in importance.

## 4.2 Current soil development state

With respect to the second research question, it was possible to find the current soil development state in the region, indicated by the current SOC value. The mechanistic model applied to predict the SOC content observed in soils was successful (cor. = 0.82) in explaining the overall variance of the samples taken. Consequently, this method turned out to be successful in quantifying large-scale differences in SOC (Fig. 10) associated with the applied management practices. The current soil development state was derived by situating the current SOC content in the derived SOC development curves and asses its horizontal proximity to the pristine areas in the figure (Fig. 7). By generating a 30x30 m resolution SOC map of the Baviaanskloof catchment, this method offered a noteworthy improvement compared current SOC maps generated by (Minasny et al., 2017, Arrouays et al.,2017). In contrast to the current method, these maps only provided single ranges (0-5%) of SOC for the full study area. As a result, the current method significantly improved the overview of soil development state in this region. Moreover, the fact all soil forming factors are included in the applied prediction of SOC observations indicates the mechanistic model involves a multidisciplinary approach. This multidisciplinary in combination with cost-effective SOC mapping, has not been applied before to cases of ecological restoration. Therefore, this method improves, for instance, the locally applied carbon benefit project approach tools, which lack multidisciplinary by neglecting soil erosion and parent material as input (GEF, 2010). These are soil physical factors, this study and other studies highlight (Van Luijk et al.,2013), as being highly important for spatial differences in SOC.

Despite the ability to predict most of the variability in the SOC observations, the current model performs less compared to other mechanistic models. In fact, the current model captured significantly less variance compared to the original setup of Hendriks (2018). Firstly, this can be explained by the current methodological assumption of steady-state conditions. This assumption might not have been appropriate for predicting SOC in a non-equilibrium case associated with ecological restoration (Choi, 2004). While this assumption is frequently applied in other cases (Karhu et al. 2012, Ortiz et al. 2013), also Wu et al. (2015) observed a poor prediction of a well-established mechanistic model applied to restoration cases. However, the assessment of the impact of the equilibrium assumption to a dynamical complex system, as described before, requires environmental legacy data (Anand, 2004), which was not available in this study.

The second reason for the lower explained variance is the poor contribution of the NDVI to the spatial distribution of SOC. While this variable explained SOC variation in some other study areas well (Abu-hashim et al.,2016, Wang et al., 2018), its contribution has been questioned in this and other studies (Peng et al., 2015). This research explains the possible poor contribution of NDVI to SOC dynamics and therefore overall monitoring of ecological restoration. This explanation is related to the observed influence of rainfall variability and legacy on the temporal variability in the NDVI (Davenport and Nicholson, 1993). Since, the variability in rainfall does not necessarily follow the same trend as the tree height variability, the prediction of tree height by NDVI is disturbed in the current situation, especially for a combination of water limiting areas and a particular vegetation type (Davenport and Nicholson, 1993). *Spekboom* is one of the vegetation types (Harris et al., 2018) due to its smaller leave area, which in combination with a water scarce environment, offers a poor example of the assessment of the vegetation development trend with the NDVI.

The prediction of the soil development state with SOC, has also been very much limited to sampling design in this research. Firstly, the design is not fully representative for the spatial distribution of SOC on plot level. This is due to the decision that samples were going to be taken close to trees, which is made because this research aimed at detecting a significant change of SOC as a result of *Spekboom* planting. However, sampling close to tree implies SOC at other land covers types are neglected, while differences in SOC between land cover types are present due to limited change in SOC outside the influential zone of *Spekboom* (Mills and Cowling, 2010). Ultimately, the current sampling design causes the generated SOC maps to be relatively unrepresentative for the actual soil development state in the area, although spatial differences are minimally affected by the described limitations. Secondly, sampling size, compared to the mechanistic approach of Hendriks (2018), has been a major issue regarding the uncertainty and validation of the derived SOC map. Thereby, there was a suboptimal spatial coverage of samples due to the availability of only 50 samples for covering the full catchment. Moreover, there was limited possibility to compare the current approach with regular kriging techniques, which require a higher spatial coverage of samples to be performed optimally (Hughes and Lettenmaier,1981). Because of the lack of comparison, the applied mapping technique cannot be fully validated, besides assessing the model calibration on its possible limitations. Nevertheless, the limitations caused by sampling design can only be regarded limiting for the accuracy and representativeness of the SOC map. It is not limiting for achieving the actual goal of determining the suitability of the proposed methods for identifying the soil development state. In fact, it was very much possible to create a map using the mechanistic approach for quantifying the development state and thus the current success of ecological restoration.

## 4.3 Temporal development

***Chronosequence***

The following section discusses the possibility of predicting the temporal behavior of the soil development state after restoration has been taking place, using the chronosequence time period. The findings have shown that the SOC chronosequence, which is established at the short-term timescale of the observations (Fig. 7) showed an exponential increase in SOC with time, which is well predictable ($R^2$=0.58). Further analysis showed that the maximum and minimum scenario curves diverged from the overall mean curve proceeding over time. Where the majority of plots including the maximum scenario could be associated to “restoring conditions”, the minimum scenario showed some alternative “degraded conditions”. At these plots belonging to the minimum scenario, *Spekboom* planting did not result in any change in SOC nor other important environmental parameters (Fig. 3), in contrast to other restored plots.

The observations of divergence in results of chronosequences are not new to scientists (Walker et al., 2010). Yet, divergence has never been associated with SOC chronosequences associated with land use changes. In this study, divergence is related to the occurrence of alternative states in restoration ecology. On this topic, literature suggests the likely occurrence of such states in restoration ecology (Bestelmeyer, 2006, Suding et al., 2004, Scheffer et al, 2001), but scientists have not yet been able to prove its occurrence using actual timeseries or chronosequences. They suggest that an alternative degraded state, as observed and associated here with the lower chronosequence scenario, occurs when the system, earlier on, has tipped from a “restoring state” to a “degraded state” (Bestelmeyer, 2006). Grazing pressure, erosion, but also drought (Vetter, 2009) and extreme rainfall (Ahmed, 2017) can be drivers which reduce the resilience of the system and could have triggered such tipping points in the current region. The degraded plots agree with having such conditions associated with heavy grazing pressure, high rainfall accumulation and low physical protection to erosion. The degraded state is hard to overcome by planting vegetation (Bestelmeyer, 2006, Suding et al., 2004, Scheffer et al, 2001) and therefore no increased SOC detected.

The occurrence of divergence and alternative states makes the currently derived chronosequence less suitable for predicting the behavior of SOC after restoration. As an addition, this prediction involves other uncertainties. These mainly arise from spatial variability in environmental factors. Their influence on the observed SOC fractions hardens the ability of getting a clear indication of divergence and possible alternative states. Therefore, in combination with a limited sample size, the points which could be clearly qualified for alternative conditions, were limited. Consequently, limited amount of available points causes it to be hard to statically prove the existence of a different temporal response of the soil state to ecological restoration. The uncertainty on the actual response ultimately leads to an uncertainty in the prediction of the development path of SOC. Thus, although alternative states could be associated and theoretically strengthened, they should be interpreted with caution.

*Future development*

The previous section focused on the chronosequence curve and the predictability of SOC after restoration up to the chronosequence timescale. To predict future conditions, the findings in this study should be extrapolated over time. These findings have shown that the oldest samples are close to pristine in terms of SOC values. This indicates the majority of plots are close to turning pristine in the future. This would agree with the hypothesis of Mills and Cowling, (2010) which explains 17 years of restoration is needed to approach pristine conditions in this ecological region, associated with *Spekboom* development.

To come to the described findings some uncertainties were involved. These uncertainties are related to the extrapolation of the chronosequence. The extrapolation was needed since this study did not collect any data about the development of SOC after 12 years (Fig. 7). Extrapolation can be very uncertain in case of non-linear behavior. A tipping point from restoring to pristine, associated with the period after 12 years in this research, would contribute to this non-linearity and has been established before (Scheffer et al., 2001). However, other research contradicts by stating the transition from restoring to pristine conditions lacks any additional threshold associated with tipping point behavior (Suding et al, 2006). Moreover, the same behavior is assumed before in the implementation of temporal behavior in organic carbon models (Coleman and Jenkinson, 1995). The extrapolation of the chronosequence becomes more uncertain when considering the impact of the observed divergence (Fig. 7) for future predictions of SOC. Keeping this in mind, the extrapolation of the chronosequence would create uncertainties associated with decades of carbon build up. However, this current study offered some potential factors influencing the divergence, which, in combination with the assumed absence of additional tipping points, allows decent predictability associated with future extrapolation of the chronosequence.

In order to offer any alternative approach for predicting future development of SOC, this study was very much limited to the application of methods which were unsuitable for this purpose. First, the findings show that the mechanistic approach could not be temporally extrapolated. This is caused by the fact the NDVI could not accurately predict tree height development, which is the main external factor influencing SOC in the temporal domain. Another failure of this model for the purpose of predicting temporal development is related to the fact the mechanistic model includes large residuals associated with the prediction of SOC observations, these residuals would only increase by temporally extrapolating the approach. Finally, the results show that the known carbon models for predicting carbon sequestration, do not capture the SOC dynamics well (Fig. 12). This is due to assumption of linear behavior, a lack of the implementations of important drivers and a general overestimation of SOC development. Given the fact these models also heavily depend on the correct implementations of land cover fractions, the known carbon models are deemed unsuitable for predicting the future SOC stock development. This contributes to the overall conclusion that the future conditions could not be accurately established, which makes it difficult to determine when pristine conditions might be reached, associated with a potential different state (Suding et al., 2004). Due to the necessity of chronosequence extrapolation and the lack of a better alternative, it is still uncertain how the current system is developing towards a future pristine condition.

## 4.4 Future perspective/recommendations

### 4.4.1 Restoration practices

The conclusions of this study regarding the main drivers of soil development as a result of ecological restoration do have important consequences for the future perspective of restoration practices. This study concluded parent material and livestock to be two important underestimated drivers. There influences on other environmental traits, such as aggregates and microbial activity, can potentially decide the ultimate faith of soil development state after ecological restoration. However, parent material and livestock presence are associated with some uncertainties due to spatial variation. Therefore, it is deemed important to overcome uncertainties regarding the contribution of these drivers first, before concluding on their consequences. To do this, purposeful sampling is recommended as a tool to determine the contribution of environmental drivers in influencing the effect of ecological restoration on soil functioning. This sampling design means areas should be assessed on soil development status by comparing plots with different magnitude of the driver desired to investigate, but a similar magnitude of other important environmental parameters. Preferably the same plots are repeatedly sampled over time, such that, multiple restoration curves can be made differing in the magnitude of the investigated driver. Only then an accurate determination of the impact of the drivers can be assessed and consequently be applied into the practice of ecological restoration.

The conclusions related to the temporal behavior of soil development after ecological restoration has implications for restoration practices, as well. This study indicated the occurrence of multiple SOC development scenarios, of which one was characterized by a lack of vegetation growth and SOC development (degraded state). The presence of such a degraded system state implies that applying ecological restoration might not be beneficial in returning ecosystem services in these areas (Bestelmeyer, 2006). This knowledge has implications for the future perspective of ecological restoration which is very much aimed to return ecosystem services, which benefits from a large success rate of vegetation introduction. These success rates of vegetation planting are currently low, both in the Baviaankloof as well as other areas around the world (Cao et al., 2011). Consequently, in most parts of the study area, ecological restoration has failed in improving soil functioning, despite the promising SOC development showed in this research in case of the successful sites. To improve success rates the restoration of vegetation in a new area should be carefully planned. Thereby, one should take into account the most important drivers determining the possible occurrence of restoration failure. These drivers have been poorly understood in the past in semi-arid areas (Bullock et al., 2011) but are indicated in this study. Eventually this may imply that applying ecosystem restoration, in terms of natural development, might not turn out to be beneficial in areas where initial soil conditions are negatively deciding the ultimate success of vegetation establishment. This means initial conditions, in terms of main drivers, should be improved first, as is done more frequently nowadays by applying for example sedimentation traps (Thomaz and Luiz, 2012). Another possibility is applying restoration only to areas where its benefits are optimal (Marrs et al., 2000), which corresponds with finding optimal conditions in terms of soil development drivers.

### 4.4.2 Monitoring

The application of the current methods for predicting the soil development state in the Baviaanskloof catchment involves high potential for monitoring purposes. SOC turned out to be an appropriate indicator of soil functioning. Combined with the fact this study was successful in identifying the current soil development state by applying the mechanistic model, the findings involve high potential for monitoring the successfulness of ecological restoration practices. Especially because the approach is cost-effective and addresses multiple disciplines associated with soil development due to the usage of SOC as indicator of soil functioning. Such a combination of cost-effectiveness and multidisciplinary has only been scarcely applied for monitoring restoration practices. Moreover, it highly improves the currently applied models, which exclude such processes. Therefore, those models fail on the temporal domain of SOC development, which also means they would be unsuccessful in quantifying spatial development. In contrast, the mechanistic approach is more reliable.

Despite the effectiveness of this method for quantifying the SOC development in case of successful restoration, unsuccessful restoration was harder to quantify. The importance of identifying the unsuccessful situations becomes important when keeping in mind that by far the most vegetation planted does not establish due to unfavorable initial conditions. Where previously the NDVI was the prime indicator of the effectiveness of ecological restoration, in this study it turned out to have trouble capturing the spatial variation in vegetation activity and soil development. The failure of the NDVI for this purpose implies that the current method could not reliably give a full overview of the success of the restoration efforts, next to the fact initial conditions could hardly be predicted with the mechanistic model. Therefore, policy makers should reconsider using indices such as the NDVI in areas where they might not be appropriate. To determine the possible suitability of the index, the vegetation type and climate zone have been indicated important. To accurately establish an overview of suitable cases and to perform consultancy on third parties in the application of the NDVI, more research on various cases is needed.

The future perspectives for the application of the current mechanic model approach to monitoring are dependent on the improvements of limitations. These limitations up to now prevent accurate determination of the development state. The already discussed poor contribution of the NDVI should be improved by selecting a proper vegetation indicator instead. A good cost-effective alternative might be the RS-derived CASA netto primary production (NPP), which incorporates climate and land use cover next to the NDVI (Potter et al., 1993). Subsequently the vegetation indicator can be used as input for soil development models, as is successfully being applied by Feng et al. (2013). Finding a proper indicator for vegetation is a first step for improving the current soil development assessment and therefore the monitoring strategy. A next step can be found in applying a sampling design which has an optimized sampling distribution for accurate spatial coverage, which potentially hugely increase the performance of the model as can be seen from Hendriks (2018). Moreover, sampling design should be adjusted to the goals to be achieved. Thereby, current study focused on establishing the relation with soil and vegetation and therefore sampled at trees. In case of monitoring however, full spatial coverage should be sampled, unrelated to tree cover locations. The same holds for sampling depth, which should be adapted to aim at capturing the full desired physical influence of land use change (Zhang and Hartemink, 2017), instead of selecting sampling depth based on available equipment or convenience. The last, more advanced step is finding a solution for applying a more appropriate dynamical model instead of the current steady-state mechanistic model to a non-equilibrium case. This would increase complexity but might be an important to improve prediction of SOC, mainly related to the effect of initial conditions on the current SOC present.

Finally, one can discuss the applicability of mechanistic models to other restoration cases. More specifically, since the relationships are optimized and therefore not applicable to other cases, the application of this model needs baseline measurements, which might cause other sampling methods to be more effective regarding financial constraints. Nevertheless, despite the limited sampling size and input parameters the mechanistic model approach can accomplish an indication of the current soil development state. Therefore, it is recommended very useful for monitoring purpose, given the recommendations are followed

### 4.4.3 Temporal SOC modelling

The analyses on the temporal development of SOC in the restoration curve have profound implications for future modelling of SOC development as a result of ecological restoration. This study has identified multiple scenarios and system states, which have only scarcely been found before in restoring ecosystems. This is the case because identifying them is widely beyond scope of most research (Suding et el.,2004). However, identifying the possible system states is considered highly important. In fact, research on restoration ecology has acknowledged that the response of the complex ecological system (Anand, 2004) to land use change is important for the understanding for further recommendations on restoration ecology (Suding et al., 2016). For the current situation this means the results obtained by applying the chronosequence might have the granted policy makers involved, a better indication of the potential soil development of plots exposed to ecological restoration. Thereby, the possible time it takes to approach pristine conditions has been identified.

Another implication for modelling scenarios relates to the results of the comparison between the chronosequence scenarios and the already known carbon models. This comparison offered an important realization about the suitability of those models. This includes that the carbon models were found to be unsuitable for quantifying the temporal development of SOC. Therefore, at catchment scale, the carbon sequestration assessment they pursue involves a considerable uncertainty. This has large implications for the current application of such models. More specifically it implies that these models are only usable for larger national scale assessments of carbon sequestration. Thereby large caution should be taken when implementing land cover fractions and vegetation type into these models, which are most often very uncertain, while potentially highly influencing the outcomes.

to improve the future perspective of soil carbon predictions, the vast limitations of both carbon models and the methods in this study, should be overcome. A method, which would overcome the temporal limitations described, requires to obtain repeated measurement of SOC over time. These should be conducted on plots differing in environmental conditions associated with the occurrence of alternative states such as suggested by Alday and Marrs (2010). However, Alday and Marrss (2010) focuses on ecological state variables, which have not been proved to be the ultimate indicators of soil system development. Instead focusing on the main drivers of soil development as described in this study, is highly recommended.

Another option to overcome limitations would be creating a new chronosequence with an increased number of samples. Preferably samples with a longer restoration time should be included as well. This new approach would allow a significant differentiation between the development scenarios and a more accurate prediction of future scenarios. However, the initial conditions, mainly related to aggregates, livestock grazing, and most likely microbial biomass, have been identified as being potentially highly important for causing divergence (Walker et al., 2010). Consequently, these Initial conditions should be considered as well, when constructing a chronosequence for temporal SOC prediction. Thereby, either the collection of knowledge to predict those conditions or taking few repeated measurements is recommended. This contributes to the accurate establishment of the possible temporal development scenarios of SOC as a result of ecological restoration.

Despite the large uncertainties associated with the establishment of scenarios, the current research raises awareness for the possible scenarios soil development may imply. In relation with the current established models for carbon sequestration assessments, it shows these models bear uncertainties which are hard to overcome, especially on the current spatial scale. Therefore, reports on carbon sequestration are highly dependent on scenarios followed and on success rates which are not always possible to obtain. Based on this, the mismatch between modelled and observed soil development conditions have not yet been overcome and should be addressed in order to increase global success rates of ecological restoration and to gain insight in the global potential of these practices.

# 5. Conclusion

Understanding the process of soil development after ecological restoration is deemed important for management decisions and the global potential of carbon sequestration. The suitability of SOC to monitor soil development is shown to be appropriate, while other indicators like the NDVI tend to fail for this purpose. Tree height is the main driver of SOC in this study area, but its establishment and development are strongly influenced by initial conditions. The influence of initial conditions was related to parent material type as well as susceptibly livestock presence and erosional processes. The effect of those drivers on the success rate of vegetation establishment and development raises awareness these drivers should be included in management decisions regarding ecological restoration, especially in semi-arid areas.

Monitoring requires a cost-effective assessment strategy for the quantifying current soil development state. The potential of the mechanistic model method for this purpose is identified, since it was successful in predicting the observed variability in SOC. However, it includes limitations regarding sampling design and coverage, vegetation indication and suitability for complex systems. Although a complex system theory would tackle the last limitation, the current study offered a more cost-effective strategy for quantifying the soil state. Therefore, the mechanistic model approach is potentially more applicable for monitoring given the fact recommendations are followed. It therefore the importance of including soil forming processes and initial conditions to improve predictions of soil development.

The goal of predicting temporal development of SOC in the Baviaanskloof, was moderately successful using the chronosequence approach. Multiple scenarios of development where established, which showed divergence and indicated the possible occurrence of an alternative degraded state. However, sampling size and influence of environmental conditions limited the prove. Identification of the future soil state was uncertain due to limited possibility to temporally extrapolate the measurements and the unsuitability of the known methods to predict SOC development. Since indications of temporal behavior and future states were established, this study contributes in narrowing the gap between modelled and observed soil development conditions.

The available carbon models for quantifying carbon sequestration and development did fail in capturing the complexity of the system and were overestimating the vegetation potential in building up SOC, when being compared to the chronosequence approach. Because of this general failure in the Baviaanskloof catchment, both models might not be suitable for quantifying soil organic carbon in smaller catchments. Moreover, these models include substantial sensitivity to its input parameters regarding land use coverage, which are most often uncertain. Consequently, the limitation described should be carefully assessed when applying the known models to a restoration case.

The identification of parent material as a driver in determining (temporal) soil development in combination with the successful application of a multidisciplinary model for predicting SOC distribution contributes to an important conclusion. Despite its complex behavior, the importance of soil forming processes in driving the ecological system, teaches us multidisciplinary methods such as applied in this research, are crucial in the application of restoration practices, monitoring and modelling scenarios.

# 6. Acknowledgements

I am particularly grateful for the assistance given by Livinglands B.V., which provided the opportunity to perform fieldwork and lab analysis in the Baviaanskloof.

Also, I would like to offer some special thanks to Commonlands B.V., in particular Simon Molenaar, for establishing some important contacts and involving in practical arrangements.

Moreover, Dr Eleanor Milne provided me with very valuable information regarding tools to model SOC.

I would also like to express deep gratitude to dr.ir. Titia Mulder and dr. Louise (Wieteke) Willemen, my supervisors in this research, for their patience, encouragements and critiques on the work performed in this study

# 7. References


1. Abu-hashim, M., Elsayed, M., & Belal, A. E. (2016). Effect of land-use changes and site variables on surface soil organic carbon pool at Mediterranean Region. *Journal of African Earth Sciences*, *114*, 78-84.
2. Ahmed, N. M. (2017). Earth System Disruption. In *Failing States, Collapsing Systems* (pp. 31-36). Springer, Cham.
3. Alday, J. G., & Marrs, R. H. (2014). A simple test for alternative states in ecological restoration: the use of principal response curves. *Applied vegetation science*, *17*(2), 302-311.
4. Alvarez, R., & Lavado, R. S. (1998). Climate, organic matter and clay content relationships in the Pampa and Chaco soils, Argentina. *Geoderma*, *83*(1-2), 127-141.
5. Andel, J., & Grootjans, A. P. (2005). Concepts in restoration ecology. *Restoration ecology: the new frontier*, 16-28.
6. Aplin, P. (2005). Remote sensing: ecology. *Progress in Physical Geography*, *29*(1), 104-113.
7. Arrouays, D., Minasny, B., McBratney, A., Grundy, M., McKenzie, N., Thompson, J., ... & Chen, S. (2017). GlobalSoilMap for Soil Organic Carbon Mapping and as a Basis for Global Modeling. In *Proceedings of the global symposium on soil organic carbon 2017* (pp. 27-30).
8. Avdan, U., & Jovanovska, G. (2016). Algorithm for automated mapping of land surface temperature using LANDSAT 8 satellite data. *Journal of Sensors*, *2016*.
9. Anand, M., & Desrochers, R. E. (2004). Quantification of restoration success using complex systems concepts and models. *Restoration Ecology*, *12*(1), 117-123.
10. Banning, N. C., Grant, C. D., Jones, D. L., & Murphy, D. V. (2008). Recovery of soil organic matter, organic matter turnover and nitrogen cycling in a post-mining forest rehabilitation chronosequence. *Soil Biology and Biochemistry*, *40*(8), 2021-2031.
11. Barskale, R. D., & Itani, S. Y. (1989). Influence of aggregate shape on base behavior. *Transportation research record*, (1227).
12. Bestelmeyer, B. T., Duniway, M. C., James, D. K., Burkett, L. M., & Havstad, K. M. (2013). A test of critical thresholds and their indicators in a desertification-prone ecosystem: More resilience than we thought. *Ecology Letters*, *16*(3), 339-345.
13. Bestelmeyer, B. T. (2006). Threshold concepts and their use in rangeland management and restoration: the good, the bad, and the insidious. *Restoration Ecology*, *14*(3), 325-329.
14. BEVEN, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, *24*(1), 43-69.
15. Blanksma, N. (2011). The Impact of land use on the provisioning of ecosystem services in the thicket biome of the Baviaanskloof, South Africa. Land and Water management. Wageningen, Wageningen University. Msc: 59.
16. Bobbins, K. L. (2011). Developing a form-process framework to describe the functioning of semi-arid alluvial fans in the Baviaanskloof valley, South Africa.
17. Brady, N. C., & Weil, R. R. (2010). *Elements of the nature and properties of soils* (No. 631.4 B733E.). Upper Saddle River, NJ: Pearson Educational International.
18. Bullock, J. M., Aronson, J., Newton, A. C., Pywell, R. F., & Rey-Benayas, J. M. (2011). Restoration of ecosystem services and biodiversity: conflicts and opportunities. *Trends in ecology & evolution*, *26*(10), 541-549.
19. Calzolari, C., Ungaro, F., Filippi, N., Guermandi, M., Malucelli, F., Marchi, N., ... & Tarocco, P. (2016). A methodological framework to assess the multiple contributions of soils to ecosystem services delivery at regional scale. *Geoderma*, *261*, 190-203.

20. Cao, S., Chen, L., Shankman, D., Wang, C., Wang, X., & Zhang, H. (2011). Excessive reliance on afforestation in China's arid and semi-arid regions: lessons in ecological restoration. *Earth-Science Reviews*, *104*(4), 240-245.
21. Carpenter, S. R., & Turner, M. G. (2000). Hares and tortoises: interactions of fast and slow variablesin ecosystems. *Ecosystems*, *3*(6), 495-497.
22. Chen, X., & Duan, Z. (2009). Changes in soil physical and chemical properties during reversal of desertification in Yanchi County of Ningxia Hui autonomous region, China. *Environmental Geology*, *57*(5), 975-985.
23. Chen, X., Duan, Z., & Tan, M. (2016). Restoration affect soil organic carbon and nutrients in different particle-size fractions. *Land degradation & development*, *27*(3), 561-572.
24. Chen, L., Wei, W., Fu, B., & Lü, Y. (2007). Soil and water conservation on the Loess Plateau in China: review and perspective. *Progress in Physical Geography*, *31*(4), 389-40
25. Choi, Y. D. (2004). Theories for ecological restoration in changing environment: toward 'futuristic'restoration. *Ecological Research*, *19*, 75-81.
26. Coleman, K., & Jenkinson, D. S. (1995). *ROTHC-26.3: A Model for the Turnover of Carbon in Soil: Model Description and Users Guide: K. Coleman and DS Jenkinson*. IACR.
27. Conant, R. T., Ogle, S. M., Paul, E. A., & Paustian, K. (2011). Measuring and monitoring soil organic carbon stocks in agricultural lands for climate mitigation. *Frontiers in Ecology and the Environment*, *9*(3), 169-173
28. Costantini, E. A., Branquinho, C., Nunes, A., Schwilch, G., Stavi, I., Valdecantos, A., & Zucca, C. (2016). Soil indicators to assess the effectiveness of restoration strategies in dryland ecosystems. *Solid Earth*, *7*(2), 397-414.
29. Davenport, M. L., & Nicholson, S. E. (1993). On the relation between rainfall and the Normalized Difference Vegetation Index for diverse vegetation types in East Africa. *International Journal of Remote Sensing*, *14*(12), 2369-2389.
30. Draaijer, A. G. (2010). *Monitoring Soil Water Interactions of South African Succulent Thicket Restoration Cost Effectively on Field Scale*. Sl: sn.
31. D'Urso, G., & Calera Belmonte, A. (2006, August). Operative approaches to determine crop water requirements from Earth Observation data: methodologies and applications. In *AIP conference proceedings* (Vol. 852, No. 1, pp. 14-25). AIP.
32. Elbasiouny, H., Abowaly, M., Abu_Alkheir, A., & Gad, A. (2014). Spatial variation of soil carbon and nitrogen pools by using ordinary Kriging method in an area of north Nile Delta, Egypt. *Catena*, *113*, 70-78.
33. FAO (2017). Voluntary Guidelines for Sustainable Soil Management, Food and Agriculture Organization of the United Nations, Rome, Italy
34. Feng, X., Fu, B., Lu, N., Zeng, Y., & Wu, B. (2013). How ecological restoration alters ecosystem services: an analysis of carbon sequestration in China's Loess Plateau. *Scientific reports*, *3*, 2846.
35. Ferwerda, W. (2015). *4 returns, 3 zones, 20 years: a holistic framework for ecological restoration by people and business for next generations* (No. 1).
36. Franzluebbers, A. J. (2002). Soil organic matter stratification ratio as an indicator of soil quality. *Soil and Tillage Research*, *66*(2), 95-106.
37. GEF (2010). Carbon Benefit Project offer innovative climate change solutions across complex landscapes global environment facility: http://www.thegef.org/gef/node/2263
38. Grant, R. F. (1995). Dynamics of energy, water, carbon and nitrogen in agricultural ecosystems: simulation and experimental validation. *Ecological Modelling*, *81*(1-3), 169-181
39. Harris, D., Vlok, J., & van Niekerk, A. (2018). Regional mapping of spekboom canopy cover using very high resolution aerial imagery. *Journal of Applied Remote Sensing*, *12*(4), 046022.
40. Harris, J. (2003). Measurements of the soil microbial community for estimating the success of restoration. *European Journal of Soil Science*, *54*(4), 801-808.
41. Hattori, T. (1988). Soil aggregates as microhabitats of microorganisms. *Reports of the Institute for Agricultural Research, Tohoku University.*, *37*, 23-36.

42. Haynes, R. J., & Naidu, R. (1998). Influence of lime, fertilizer and manure applications on soil organic matter content and soil physical conditions: a review. *Nutrient cycling in agroecosystems*, *51*(2), 123-137.
43. Hendriks, Chantal M.J. (2018). *Bridging the gap between the available and required soil data for regional land use analysis* (PhD dissertation). Retrieved from Wageningen University & Research – Library (dissertation no. – 6894)
44. Heuvelink, G. B. M., & Webster, R. (2001). Modelling soil variation: past, present, and future. *Geoderma*, *100*(3-4), 269-301.
45. 45.Hevia, G. G., Buschiazzo, D. E., Hepper, E. N., Urioste, A. M., & Antón, E. L. (2003). Organic matter in size fractions of soils of the semiarid Argentina. Effects of climate, soil texture and management. *Geoderma*, *116* (3-4), 265-277.
46. Homann, P. S., Sollins, P., Chappell, H. N., & Stangenberger, A. G. (1995). Soil organic carbon in a mountainous, forested region: relation to site characteristics. *Soil science society of america journal*, *59*(5), 1468-1475.
47. Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, *24*(6), 417.
48. Hughes, J. P., & Lettenmaier, D. P. (1981). Data requirements for kriging: estimation and network design. *Water Resources Research*, *17*(6), 1641-1650.
49. Hunt, R. (2012). *Basic growth analysis: plant growth analysis for beginners*. Springer Science & Business Media.
50. Idowu, O. J., van Es, H. M., Abawi, G. S., Wolfe, D. W., Ball, J. I., Gugino, B. K. & Bilgili, A. V. (2008). Farmer-oriented assessment of soil quality using field, laboratory, and VNIR spectroscopy methods. *Plant and Soil*, *307*(1-2), 243-253.
51. Idowu, O. J., Van Es, H. M., Abawi, G. S., Wolfe, D. W., Schindelbeck, R. R., Moebius-Clune, B. N., & Gugino, B. K. (2009). Use of an integrative soil health test for evaluation of soil management impacts. *Renewable Agriculture and Food Systems*, *24*(3), 214-224.
52. Insam, H., & Domsch, K. H. (1988). Relationship between soil organic carbon and microbial biomass on chronosequences of reclamation sites. *Microbial ecology*, *15*(2), 177-188.
53. Jansen, H. C. (2008). *Water for food and ecosystems in the Baviaanskloof Mega Reserve: land and water resources assessment in the Baviaanskloof, Eastern Cape Province, South Africa* (No. 1812). Alterra
54. Jenny, H. (1941). *Factors of soil formation; a sytem of quantitative pedology* (No. 631.4 J45).
55. Jia, G. M., Cao, J., Wang, C., & Wang, G. (2005). Microbial biomass and nutrients in soil at the different stages of secondary forest succession in Ziwulin, northwest China. *Forest Ecology and Management*, *217*(1), 117-125.
56. Jian-Bing, W., Du-Ning, X., Xing-Yi, Z., Xiu-Zhen, L., & Xiao-Yu, L. (2006). Spatial variability of soil organic carbon in relation to environmental factors of a typical small watershed in the black soil region, northeast China. *Environmental monitoring and assessment*, *121*(1-3), 597-613.
57. Jin, K., Cornelis, W. M., Gabriels, D., Baert, M., Wu, H. J., Schiettecatte, W., ... & Hofman, G. (2009). Residue cover and rainfall intensity effects on runoff soil organic carbon losses. *Catena*, *78*(1), 81-86
58. Karhu, K., Gärdenäs, A. I., Heikkinen, J., Vanhala, P., Tuomi, M., & Liski, J. (2012). Impacts of organic amendments on carbon stocks of an agricultural soil—comparison of model-simulations to measurements. *Geoderma*, *189*, 606-616.
59. Keesstra, S., Mol, G., de Leeuw, J., Okx, J., de Cleen, M., & Visser, S. (2018). Soil-related sustainable development goals: four concepts to make land degradation neutrality and restoration work. *Land*, *7*(4), 133.
60. Lebrija-Trejos, E., Meave, J. A., Poorter, L., Pérez-García, E. A., & Bongers, F. (2010). Pathways, mechanisms and predictability of vegetation change during tropical dry forest succession. *Perspectives in Plant Ecology, Evolution and Systematics*, *12*(4), 267-275.
61. Leifeld, J., & Kögel-Knabner, I. (2005). Soil organic matter fractions as early indicators for carbon stock changes under different land-use?. *Geoderma*, *124*(1-2), 143-155.

62. Lombard, A. T., Johnson, C. F., Cowling, R. M., & Pressey, R. L. (2001). Protecting plants from elephants: botanical reserve scenarios within the Addo Elephant National Park, South Africa. *Biological Conservation*, *102*(2), 191-203.
63. Lucero, M. E., Barrow, J. R., Osuna, P., & Reyes, I. (2006). Plant–fungal interactions in arid and semi-arid ecosystems: large-scale impacts from microscale processes. *Journal of Arid Environments*, *65*(2), 276-284.v
64. Marrs, R. H., Le Duc, M. G., Mitchell, R. J., Goddard, D., Paterson, S., & Pakeman, R. J. (2000). The ecology of bracken: its role in succession and implications for control. *Annals of Botany*, *85*(suppl_2), 3-15.
65. Mills, A. J., & Cowling, R. M. (2010). Below-ground carbon stocks in intact and transformed subtropical thicket landscapes in semi-arid South Africa. *Journal of Arid Environments*, *74*(1), 93-100.
66. Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., ... & Field, D. J. (2017). Soil carbon 4 per mille. *Geoderma*, *292*, 59-86.
67. Minasny, B., & McBratney, A. B. (2006). Mechanistic soil–landscape modelling as an approach to developing pedogenetic classifications. *Geoderma*, *133*(1-2), 138-149.
68. Muñoz-Rojas, M., Erickson, T. E., Dixon, K. W., & Merritt, D. J. (2016). Soil quality indicators to assess functionality of restored soils in degraded semiarid ecosystems. *Restoration Ecology*, *24*, S43-S52.
69. Mulder, V. L., De Bruin, S., Schaepman, M. E., & Mayr, T. R. (2011). The use of remote sensing in soil and terrain mapping—A review. *Geoderma*, *162*(1-2), 1-19.
70. Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, *97*(4), 558-625.
71. Ortiz, C. A., Liski, J., Gärdenäs, A. I., Lehtonen, A., Lundblad, M., Stendahl, J., ... & Karltun, E. (2013). Soil organic carbon stock changes in Swedish forest soils—a comparison of uncertainties and their sources through a national inventory and two simulation models. *Ecological Modelling*, *251*, 221-231.
72. Patton, M. (1990). Purposeful sampling. *Qualitative evaluation and research methods*, *2*, 169-186.
73. Pastorok, R. A., MacDonald, A., Sampson, J. R., Wilber, P., Yozzo, D. J., & Titre, J. P. (1997). An ecological decision framework for environmental restoration projects. *Ecological Engineering*, *9*(1-2), 89-107.
74. Peng, Y., Xiong, X., Adhikari, K., Knadel, M., Grunwald, S., & Greve, M. H. (2015). Modeling soil organic carbon at regional scale by combining multi-spectral images with laboratory spectra. *PloS one*, *10*(11), e0142295.
75. Pearson, K. (1896). Mathematical contributions to the theory of evolution.–III. Regression, heredity and panmixia. *Philos. Trans. Royal Soc. London*, Ser. A, vol. 187, pp. 253–318,
76. Post, W. M., & Kwon, K. C. (2000). Soil carbon sequestration and land-use change: processes and potential. *Global change biology*, *6*(3), 317-327.
77. Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*(4), 811-841.
78. Raiesi, F., & Asadi, E. (2006). Soil microbial activity and litter turnover in native grazed and ungrazed rangelands in a semiarid ecosystem. *Biology and Fertility of Soils*, *43*(1), 76-82.
79. Read, J. W., & Ridgell, R. H. (1922). On the use of the conventional carbon factor in estimating soil organic matter. *Soil Sci*, *13*, 1-6.
80. Rietkerk, M., van den Bosch, F., & van de Koppel, J. (1997). Site-specific properties and irreversible vegetation changes in semi-arid grazing systems. *Oikos*, 241-252.
81. Robinson, T. P., & Metternicht, G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and electronics in agriculture*, *50*(2), 97-108
82. Rossello, P., & Baret, F. (2007). Validation of Land European Remote sensing Instruments. *INRA, Davo*
83. Runge, J., Petoukhov, V., Donges, J. F., Hlinka, J., Jajcay, N., Vejmelka, M., ... & Kurths, J. (2015). Identifying causal gateways and mediators in complex spatio-temporal systems. *Nature communications*, *6*, 8502

84. Sadeghi, M., Babaeian, E., Tuller, M., & Jones, S. B. (2017). The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. *Remote sensing of environment*, *198*, 52-68.
85. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, *413*(6856), 591.
86. Shirato, Y., Paisancharoen, K., Sangtong, P., Nakviro, C., Yokozawa, M., & Matsumoto, N. (2005). Testing the Rothamsted Carbon Model against data from long-term experiments on upland soils in Thailand. *European Journal of Soil Science*, *56*(2), 179-188.
87. Shourzan, L. I. U., Shengrli, G. U. O., & Xiaorli, W. A. N. G. (2005). Effect of vegetation on soil organic carbon of slope land in gully region of Loess Plateau. *Journal of Natural Resources*, *20*(4), 529D536
88. Singh, J. S., & Gupta, V. K. (2018). Soil microbial biomass: A key soil driver in management of ecosystem functioning. *Science of The Total Environment*, *634*, 497-500.
89. Six, J., Bossuyt, H., Degryze, S., & Denef, K. (2004). A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. *Soil and Tillage Research*, *79*(1), 7-31
90. Skarpe, C. (2000). Desertification, no-change or alternative states: Can we trust simple models on livestock impact in dry rangelands?. *Applied Vegetation Science*, *3*(2), 261-268.
91. Smith, P. (2008). Land use change and soil organic carbon dynamics. *Nutrient Cycling in Agroecosystems*, *81*(2), 169-178.
92. Sparling, G. P. (1992). Ratio of microbial biomass carbon to soil organic carbon as a sensitive indicator of changes in soil organic matter. *Soil Research*, *30*(2), 195-207.
93. Suding, K. N., Gross, K. L., & Houseman, G. R. (2004). Alternative states and positive feedbacks in restoration ecology. *Trends in ecology & evolution*, *19*(1), 46-53.
94. Suding, K., Spotswood, E., Chapple, D., Beller, E., & Gross, K. (2016). Ecological dynamics and ecological restoration. In *Foundations of restoration ecology* (pp. 27-56). Island Press, Washington, DC.
95. Teng, H., Shi, Z., Ma, Z., & Li, Y. (2014). Estimating spatially downscaled rainfall by regression kriging using TRMM precipitation and elevation in Zhejiang Province, southeast China. *International journal of remote sensing*, *35*(22), 7775-7794.
96. Thomaz, E. L., & Luiz, J. C. (2012). Soil loss, soil degradation and rehabilitation in a degraded land area in Guarapuava (Brazil). *Land Degradation & Development*, *23*(1), 72-81.
97. Tiessen, H., Cuevas, E., & Chacon, P. (1994). The role of soil organic matter in sustaining soil fertility. *Nature*, *371*(6500), 783-785.
98. Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*, *8*(2), 127-150
99. Van Luijk, G., Cowling, R. M., Riksen, M. J. P. M., & Glenday, J. (2013). Hydrological implications of desertification: Degradation of South African semi-arid subtropical thicket. *Journal of arid environments*, *91*, 14-21
100. Vetter, S. (2009). Drought, change and resilience in South Africa's arid and semi-arid rangelands. *South African Journal of Science*, *105*(1-2), 29-33.
101. Viaud, V., Angers, D. A., & Walter, C. (2010). Toward landscape-scale modeling of soil organic matter dynamics in agroecosystems. *Soil Science Society of America Journal*, *74*(6), 1847-1860.
102. Victoria, R., Banwart, S., Black, H., Ingram, J., Joosten, H., Milne, E., ... & Baskin, Y. (2012). The benefits of soil carbon. *Foresight chapter in UNEP Yearbook*, *2012*, 19-33.
103. Walker, B. H., & Noy-Meir, I. (1982). Aspects of the stability and resilience of savanna ecosystems. In *Ecology of tropical savannas* (pp. 556-590). Springer, Berlin, Heidelberg.
104. Walker, L. R., Wardle, D. A., Bardgett, R. D., & Clarkson, B. D. (2010). The use of chronosequences in studies of ecological succession and soil development. *Journal of Ecology*, *98*(4), 725-736.
105. Walkley, A., & Black, I. A. (1934). An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil science*, *37*(1), 29-38.

106. Wang, S., Adhikari, K., Wang, Q., Jin, X., & Li, H. (2018). Role of environmental variables in the spatial distribution of soil carbon (C), nitrogen (N), and C: N ratio from the northeastern coastal agroecosystems in China. *Ecological Indicators*, *84*, 263-272.
107. Wang, Y. Q., & Shao, M. A. (2013). Spatial variability of soil physical properties in a region of the Loess Plateau of PR China subject to wind and water erosion. *Land Degradation & Development*, *24*(3), 296-304.
108. Wu, X., Akujärvi, A., Lu, N., Liski, J., Liu, G., Wang, Y., ... & Fu, B. (2015). Dynamics of soil organic carbon stock in a typical catchment of the Loess Plateau: comparison of model simulations with measurements. *Landscape ecology*, *30*(3), 381-397.
109. Yang, Y., Fang, J., Tang, Y., Ji, C., Zheng, C., He, J., & Zhu, B. (2008). Storage, patterns and controls of soil organic carbon in the Tibetan grasslands. *Global Change Biology*, *14*(7), 1592-1599.
110. Zhang, J. T., & Dong, Y. (2010). Factors affecting species diversity of plant communities and the restoration process in the loess area of China. *Ecological Engineering*, *36*(3), 345-350.
111. Zhang, B., Deng, H., Wang, H. L., Yin, R., Hallett, P. D., Griffiths, B. S., & Daniell, T. J. (2010). Does microbial habitat or community structure drive the functional stability of microbes to stresses following re-vegetation of a severely degraded soil?. *Soil Biology and Biochemistry*, *42*(5), 850-859.
112. Zhang, Y., & Hartemink, A. E. (2017). Sampling designs for soil organic carbon stock assessment of soil profiles. *Geoderma*, *307*, 220-230.
113. Zhang, C., Liu, G., Xue, S., & Wang, G. (2016). Soil bacterial community dynamics reflect changes in plant community and soil properties during the secondary succession of abandoned farmland in the Loess Plateau. *Soil Biology and Biochemistry*, *97*, 40-49.
114. Ziadat, F. M., & Taimeh, A. Y. (2013). Effect of rainfall intensity, slope, land use and antecedent soil moisture on soil erosion in an arid environment. *Land Degradation & Development*, *24*(6), 582-590.
115. Zucca, C., Wu, W., Dessena, L., & Mulas, M. (2015). Assessing the effectiveness of land restoration interventions in dry lands by multitemporal remote sensing–a case study in ouled dlim (Marrakech, Morocco). *Land degradation & development*, *26*(1), 80-91

# 8. Appendix

*Table A9 List of tests associated with the analyzed lab parameters (source: BEMLAB)*

## SOIL

| Code | Analyses | Description |
|---|---|---|
| **Profiles/packages** | | |
| EG001 | Standard | pH,resistance, Na, K, Ca, Mg, C, P Bray II, titratable acidity, stone fraction |
| EG002 | Full | pH, resistance, Na, K, Ca, Mg, P Bray II, titratable acidity, stone fraction, Fe, Mn, Cu, Zn, B, C |
| EG010 | Texture - 3 fraction | Clay-, silt-, sand-percentage |
| EG011 | Waterholding capacity (Texture - 5 fraction) | Clay, silt, fine- medium- & coarse sand |
| EG014 | Saturated paste extract | pH, Ec, Na, K, Ca, Mg, Cl, $SO_4^{2-}$, P, $NH_4^+$, $NO_3^-$ |
| EG018 | 1:2 Water extract | pH, Ec, P, Na, K, Ca, Mg, Fe, Zn, Mn, Cu, B |
| EG048 | Farming for the Future | |
| **Single parameters** | | |
| EG003 | Carbon (C) | Choice of Walkley-Black or total carbon (Leco) |
| EG004 | Nitrate-nitrogen | ($NO_3^-$ - N ) |
| EG005 | Ammonia-nitrogen | ($NH_4^+$ - N ) |
| EG006 | Chloride (Cl) | |
| EG007 | Sulphur (S) | |
| EG008 | Percentage ash | |
| EG009 | Percentage clay | |
| EG012 | Total nitrogen (N) | |
| EG013 | CEC | Cation exchange capacity |
| EG019 | Calcium carbonate | ($CaCO_3$ ) |
| EG020 | Heavy metals | Selection of any 5 elements, e.g. Cd, Cr, Hg, As, Pb |
| EG023 | Moisture percentage | Gravimetric moisture |
| EG026 | Bulk density | |
| EG028 | Phosphorus (P) | Choice of total P or an Olsen, Mehlich or Citric Acid extraction |
| EG030 | Single elements | Excluding heavy metals - per element |
| EG031 | Heavy metals | All types of heavy metals - per element |
| EG029 | Preparation | |
| **Microbiological analyses** | | |
| EG016 | Standard nematode analysis | OUTSOURCED |
| EG041 | Citrus nematode analysis | OUTSOURCED |
| EG051 | Full nematode analyses | OUTSOURCED |
| EG022 | Totale bacteria | Reported as total plate count |
| EG027 | Microbiology | Total bacteria, total coliforms, *E.Coli* |
| EG033 | Salmonella | |
| EG034 | Salmonella - Biological ID | Additional verification required, if detected |
| EG035 | Total fungi | |
| EG036 | *E.Coli* | |
| EG037 | Total coliforms | |
| EG046 | Plant diseases | OUTSOURCED |

*Table A10 correlations of environmental parameters associated with the different dimension in the PCA.*

| | Dim.1 | Dim.2 | Dim.3 | Dim.4 | Dim.5 |
|---|---|---|---|---|---|
| Livestock | -0.43614036 | -0.17286114 | 0.17506443 | -0.047356793 | 0.19280230 |
| Soil Moisture | 0.86827097 | -0.33756660 | -0.06931945 | 0.137170155 | 0.02687104 |
| Rock Age | 0.05711576 | 0.23066470 | 0.45167203 | -0.320493698 | 0.04067107 |
| fstone | 0.24731382 | -0.01529910 | -0.29521402 | -0.540837287 | 0.14739503 |
| Heavy Prec. | -0.33417014 | -0.50413793 | 0.48110906 | 0.354724509 | 0.31564254 |
| NDVI 2018 | 0.75698023 | -0.29924447 | -0.01016781 | 0.283531680 | 0.08545519 |
| Permeability | 0.83579076 | 0.33378437 | 0.22360651 | 0.166053804 | -0.09863812 |
| LST | -0.73873223 | 0.27871272 | 0.15814687 | 0.181664652 | 0.09117670 |
| Albedo | -0.32371350 | 0.70140329 | -0.08392019 | 0.338888058 | -0.23890339 |
| Temp. | 0.05637432 | 0.03423195 | -0.62596377 | 0.506645296 | -0.31511090 |
| Restoration Days | 0.83665546 | 0.21388453 | -0.02104087 | 0.001773416 | -0.11653152 |
| TWI | -0.49548985 | -0.36077077 | 0.36729594 | 0.004386010 | -0.32531235 |
| Slope | 0.69306817 | 0.34573086 | -0.18334982 | -0.123063236 | 0.13087012 |
| Solar Radiation | 0.02888506 | -0.71680918 | -0.18128841 | -0.266466277 | 0.17093820 |
| Summer Prec. | -0.08445270 | -0.46735532 | -0.13434303 | 0.673026901 | 0.26163071 |
| Winter Prec. | 0.30536463 | 0.08770207 | 0.55954978 | 0.033343323 | 0.31903915 |
| NDVI mean | 0.69235640 | -0.51266189 | -0.21523741 | 0.076057358 | 0.08679480 |
| C | 0.81291532 | 0.05824842 | 0.12038077 | -0.166056873 | 0.19160601 |
| Tree Height | 0.87898083 | 0.12607301 | 0.23282272 | 0.077201001 | -0.07210838 |
| fClay | -0.29749554 | 0.42844329 | -0.19381117 | 0.162908302 | 0.53898283 |
| fSilt | -0.04416750 | 0.45246438 | -0.25893486 | 0.083059591 | 0.70359422 |
| Total Nutrients | 0.48364051 | 0.24764421 | 0.44950496 | 0.412939323 | -0.08831565 |

*Table A11 Correlation table (pearson correlation) of all associated variables in this research*

| | Temp. mean | rocktype | rock age | DEM | Aspect | Slope | Solar Radiaton | Multi. flow acc. | P Overall Mean | P May-Oct | P Oct-May | Heavy rainfall | Number of dry days | NDVI mean | NDVI Sentinel | F stone | Albedo | TWI | Sol Radiation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Temp. mean | 1.0 | | | | | | | | | | | | | | | | | | |
| Rocktype | 0.2 | 1.0 | | | | | | | | | | | | | | | | | |
| Rockage | -0.2 | -0.1 | 1.0 | | | | | | | | | | | | | | | | |
| DEM | -0.6 | -0.2 | 0.3 | 1.0 | | | | | | | | | | | | | | | |
| Aspect | 0.2 | 0.2 | 0.1 | -0.1 | 1.0 | | | | | | | | | | | | | | |
| Slope | 0.1 | 0.2 | 0.2 | 0.2 | 0.3 | 1.0 | | | | | | | | | | | | | |
| Solar Radiaton | 0.0 | -0.1 | -0.3 | -0.1 | -0.1 | -0.6 | 1.0 | | | | | | | | | | | | |
| Multi. flow acc. | -0.1 | -0.2 | 0.1 | 0.1 | -0.1 | -0.3 | 0.2 | 1.0 | | | | | | | | | | | |
| P Overall Mean | -0.1 | 0.1 | -0.2 | 0.3 | 0.1 | 0.2 | 0.1 | -0.1 | 1.0 | | | | | | | | | | |
| P May-Oct | -0.4 | -0.1 | 0.1 | 0.7 | 0.0 | 0.3 | -0.1 | -0.1 | 0.8 | 1.0 | | | | | | | | | |
| P Oct-May | 0.3 | 0.2 | -0.3 | -0.2 | 0.2 | -0.2 | 0.4 | 0.1 | 0.6 | 0.0 | 1.0 | | | | | | | | |
| Heavy rainfall | -0.2 | -0.2 | 0.0 | 0.4 | 0.0 | -0.5 | 0.5 | 0.3 | 0.5 | 0.3 | 0.6 | 1.0 | | | | | | | |
| Number of dry days | 0.3 | -0.1 | 0.0 | -0.2 | 0.1 | -0.6 | 0.5 | 0.3 | -0.3 | -0.6 | 0.5 | 0.6 | 1.0 | | | | | | |
| NDVI mean | 0.2 | 0.2 | -0.2 | -0.1 | 0.1 | 0.4 | 0.2 | -0.2 | 0.3 | 0.1 | 0.2 | -0.1 | -0.1 | 1.0 | | | | | |
| NDVI Sentinel | 0.1 | 0.1 | 0.0 | 0.1 | 0.3 | 0.4 | -0.1 | -0.2 | 0.3 | 0.2 | 0.2 | 0.0 | -0.1 | 0.7 | 1.0 | | | | |
| Fstone | -0.1 | -0.4 | 0.1 | 0.1 | -0.1 | 0.3 | -0.1 | -0.1 | 0.0 | 0.1 | -0.1 | -0.2 | -0.2 | 0.1 | 0.2 | 1.0 | | | |
| Albedo | 0.2 | 0.2 | 0.0 | -0.1 | -0.2 | 0.0 | -0.5 | 0.0 | -0.2 | -0.1 | -0.1 | -0.2 | -0.1 | -0.6 | -0.3 | -0.1 | 1.0 | | |
| TWI | -0.1 | 0.0 | 0.1 | -0.1 | -0.1 | -0.6 | 0.3 | 0.5 | -0.2 | -0.2 | 0.0 | 0.4 | 0.5 | -0.2 | -0.3 | -0.3 | 0.0 | 1.0 | |
| SolRadiation | -0.1 | 0.1 | -0.2 | -0.2 | -0.1 | -0.2 | 0.8 | 0.1 | 0.0 | -0.2 | 0.2 | 0.2 | 0.3 | 0.4 | 0.1 | 0.1 | -0.6 | 0.2 | 1.0 |
| LST | -0.1 | -0.3 | 0.0 | 0.1 | -0.4 | -0.5 | 0.0 | 0.2 | -0.2 | -0.1 | 0.0 | 0.3 | 0.3 | -0.6 | -0.4 | -0.2 | 0.5 | 0.2 | -0.3 |
| restorationdays | 0.1 | 0.1 | 0.0 | 0.1 | 0.0 | 0.5 | -0.2 | -0.2 | 0.1 | 0.1 | -0.2 | -0.4 | -0.4 | 0.5 | 0.5 | 0.2 | -0.1 | -0.4 | 0.0 |
| Livestock | 0.1 | -0.1 | -0.2 | -0.1 | -0.2 | 0.4 | -0.2 | -0.2 | 0.3 | 0.3 | -0.1 | -0.3 | -0.6 | 0.3 | 0.2 | 0.2 | 0.1 | -0.3 | -0.1 |
| BD | 0.0 | -0.4 | 0.0 | -0.2 | -0.1 | 0.0 | -0.1 | 0.0 | -0.2 | -0.1 | -0.2 | -0.1 | 0.0 | -0.2 | -0.2 | 0.8 | 0.1 | -0.1 | 0.0 |
| pH | -0.2 | 0.2 | 0.1 | 0.3 | -0.1 | -0.2 | 0.0 | 0.3 | 0.0 | 0.1 | 0.0 | 0.3 | 0.1 | -0.2 | -0.2 | -0.6 | 0.1 | 0.3 | -0.1 |
| Resist | 0.1 | 0.0 | -0.4 | -0.3 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | -0.2 | 0.1 | 0.0 | 0.1 | 0.1 | 0.0 | 0.2 | 0.0 | -0.1 | 0.2 |
| PBrayII | -0.5 | 0.0 | -0.1 | 0.3 | -0.1 | -0.3 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.5 | 0.2 | -0.2 | -0.2 | -0.4 | 0.0 | 0.4 | 0.0 |
| K | 0.0 | 0.3 | 0.2 | 0.4 | 0.1 | 0.3 | -0.3 | 0.0 | 0.2 | 0.3 | -0.2 | -0.2 | -0.3 | 0.2 | 0.3 | -0.3 | 0.2 | -0.2 | -0.3 |
| Na | 0.0 | 0.2 | -0.1 | 0.0 | 0.1 | -0.1 | 0.2 | 0.3 | -0.2 | -0.1 | -0.1 | 0.0 | 0.2 | -0.2 | -0.1 | -0.2 | 0.1 | 0.0 | 0.1 |
| Ca | -0.4 | 0.1 | 0.1 | 0.5 | 0.1 | 0.2 | 0.0 | 0.0 | 0.0 | 0.2 | -0.2 | 0.1 | 0.0 | 0.2 | 0.2 | -0.1 | -0.3 | 0.1 | 0.1 |
| Mg | -0.3 | -0.1 | 0.2 | 0.5 | 0.1 | 0.2 | 0.1 | -0.1 | 0.0 | 0.3 | -0.3 | -0.1 | -0.3 | 0.2 | 0.2 | 0.0 | -0.4 | -0.1 | 0.2 |
| Cu | 0.2 | 0.4 | 0.0 | 0.0 | 0.1 | 0.0 | -0.1 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.1 | -0.5 | 0.2 | 0.2 | -0.2 |
| Zn | -0.1 | 0.1 | 0.2 | 0.3 | 0.2 | 0.3 | 0.1 | 0.1 | -0.1 | 0.0 | -0.1 | -0.1 | 0.0 | 0.2 | 0.2 | 0.2 | -0.3 | 0.0 | 0.3 |
| Mn | 0.1 | 0.4 | 0.2 | 0.1 | 0.1 | 0.2 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.0 | -0.1 | 0.2 | 0.2 | -0.3 | 0.1 | 0.1 | 0.0 |
| B | -0.2 | 0.0 | 0.2 | 0.5 | 0.1 | 0.2 | 0.0 | 0.0 | 0.1 | 0.3 | -0.2 | -0.1 | -0.3 | 0.2 | 0.3 | -0.1 | -0.2 | -0.2 | 0.0 |
| Fe | 0.4 | 0.1 | 0.2 | -0.2 | 0.1 | 0.2 | -0.2 | -0.1 | 0.1 | -0.1 | 0.2 | -0.2 | -0.1 | 0.1 | 0.2 | 0.2 | 0.0 | -0.1 | -0.1 |
| C | -0.1 | 0.1 | 0.2 | 0.4 | 0.1 | 0.5 | -0.1 | -0.2 | 0.1 | 0.2 | -0.2 | -0.3 | -0.3 | 0.5 | 0.5 | 0.2 | -0.3 | -0.4 | 0.1 |
| S | -0.2 | -0.1 | 0.1 | 0.1 | 0.0 | -0.2 | 0.3 | 0.3 | -0.3 | -0.2 | -0.1 | 0.0 | 0.2 | -0.2 | -0.1 | 0.0 | -0.1 | 0.1 | 0.2 |
| Clay | 0.0 | 0.0 | 0.0 | -0.1 | -0.3 | -0.1 | -0.1 | 0.0 | -0.2 | -0.2 | 0.0 | -0.1 | 0.1 | -0.3 | -0.2 | -0.2 | 0.3 | -0.1 | -0.1 |
| Silt | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | -0.1 | 0.0 | 0.0 | 0.1 | 0.0 | -0.1 | -0.1 | -0.1 | -0.1 | 0.1 | 0.2 | -0.3 | -0.1 |
| Sand | -0.1 | 0.0 | 0.0 | 0.1 | 0.2 | -0.1 | 0.1 | 0.0 | 0.1 | 0.1 | 0.0 | 0.1 | 0.0 | 0.2 | 0.2 | 0.1 | -0.3 | 0.2 | 0.1 |
| SOCstock | 0.0 | 0.1 | 0.1 | 0.3 | 0.1 | 0.6 | -0.2 | -0.2 | 0.2 | 0.2 | -0.1 | -0.3 | -0.4 | 0.5 | 0.5 | 0.2 | -0.2 | -0.5 | 0.0 |

| | LST | restoration days | Livestock | BD | pH | Resist | P BraylI | K | Na | Ca | Mg | Cu | Zn | Mn | B | Fe | C | S | Clay | Silt | Sand | SOCstock |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Temp. mean | | | | | | | | | | | | | | | | | | | | | | |
| Rocktype | | | | | | | | | | | | | | | | | | | | | | |
| Rockage | | | | | | | | | | | | | | | | | | | | | | |
| DEM | | | | | | | | | | | | | | | | | | | | | | |
| Aspect | | | | | | | | | | | | | | | | | | | | | | |
| Slope | | | | | | | | | | | | | | | | | | | | | | |
| Solar Radiaton | | | | | | | | | | | | | | | | | | | | | | |
| Multi. flow acc. | | | | | | | | | | | | | | | | | | | | | | |
| P Overall Mean | | | | | | | | | | | | | | | | | | | | | | |
| P May-Oct | | | | | | | | | | | | | | | | | | | | | | |
| P Oct-May | | | | | | | | | | | | | | | | | | | | | | |
| Heavy rainfall | | | | | | | | | | | | | | | | | | | | | | |
| Number of dry days | | | | | | | | | | | | | | | | | | | | | | |
| NDVI mean | | | | | | | | | | | | | | | | | | | | | | |
| NDVI Sentinel | | | | | | | | | | | | | | | | | | | | | | |
| Fstone | | | | | | | | | | | | | | | | | | | | | | |
| Albedo | | | | | | | | | | | | | | | | | | | | | | |
| TWI | | | | | | | | | | | | | | | | | | | | | | |
| SolRadiation | | | | | | | | | | | | | | | | | | | | | | |
| LST | 1.0 | | | | | | | | | | | | | | | | | | | | | |
| restorationdays | -0.5 | 1.0 | | | | | | | | | | | | | | | | | | | | |
| Livestock | -0.3 | 0.4 | 1.0 | | | | | | | | | | | | | | | | | | | |
| BD | 0.1 | -0.1 | 0.1 | 1.0 | | | | | | | | | | | | | | | | | | |
| pH | 0.2 | -0.4 | -0.2 | -0.5 | 1.0 | | | | | | | | | | | | | | | | | |
| Resist | 0.0 | 0.0 | 0.0 | 0.5 | -0.6 | 1.0 | | | | | | | | | | | | | | | | |
| PBraylI | 0.2 | -0.4 | -0.1 | -0.3 | 0.7 | -0.3 | 1.0 | | | | | | | | | | | | | | | |
| K | -0.3 | 0.5 | 0.1 | -0.5 | 0.4 | -0.4 | 0.0 | 1.0 | | | | | | | | | | | | | | |
| Na | 0.2 | -0.1 | -0.4 | -0.1 | 0.1 | 0.3 | -0.1 | 0.2 | 1.0 | | | | | | | | | | | | | |
| Ca | -0.3 | 0.2 | 0.0 | -0.4 | 0.5 | -0.5 | 0.4 | 0.4 | 0.0 | 1.0 | | | | | | | | | | | | |
| Mg | -0.5 | 0.5 | 0.1 | -0.4 | 0.1 | -0.3 | 0.0 | 0.6 | 0.2 | 0.6 | 1.0 | | | | | | | | | | | |
| Cu | 0.0 | 0.1 | -0.1 | -0.6 | 0.6 | -0.6 | 0.2 | 0.5 | 0.1 | 0.3 | 0.2 | 1.0 | | | | | | | | | | |
| Zn | -0.6 | 0.4 | 0.0 | -0.2 | 0.1 | -0.2 | 0.1 | 0.3 | 0.1 | 0.6 | 0.6 | 0.2 | 1.0 | | | | | | | | | |
| Mn | -0.2 | 0.2 | 0.0 | -0.5 | 0.3 | -0.4 | 0.1 | 0.4 | 0.0 | 0.1 | 0.1 | 0.6 | 0.2 | 1.0 | | | | | | | | |
| B | -0.4 | 0.6 | 0.1 | -0.5 | 0.3 | -0.5 | 0.1 | 0.7 | 0.2 | 0.6 | 0.8 | 0.3 | 0.5 | 0.3 | 1.0 | | | | | | | |
| Fe | -0.2 | 0.3 | 0.1 | 0.0 | -0.4 | 0.0 | -0.4 | -0.1 | -0.2 | -0.4 | -0.1 | 0.0 | 0.1 | 0.3 | 0.0 | 1.0 | | | | | | |
| C | -0.6 | 0.8 | 0.2 | -0.3 | -0.1 | -0.3 | -0.2 | 0.5 | 0.0 | 0.6 | 0.8 | 0.1 | 0.6 | 0.2 | 0.7 | 0.1 | 1.0 | | | | | |
| S | 0.1 | -0.1 | -0.4 | 0.0 | 0.0 | 0.2 | 0.0 | 0.1 | 0.8 | 0.1 | 0.3 | -0.1 | 0.3 | -0.1 | 0.3 | -0.1 | 0.1 | 1.0 | | | | |
| Clay | 0.3 | -0.2 | -0.2 | -0.2 | 0.2 | -0.1 | 0.1 | 0.0 | 0.0 | -0.1 | -0.2 | 0.1 | -0.2 | 0.0 | -0.1 | -0.2 | -0.1 | 0.0 | 1.0 | | | |
| Silt | 0.1 | 0.1 | -0.1 | 0.0 | 0.0 | -0.1 | -0.2 | 0.1 | 0.0 | 0.0 | -0.1 | 0.0 | -0.2 | 0.0 | 0.0 | -0.1 | 0.1 | -0.1 | 0.6 | 1.0 | | |
| Sand | -0.3 | 0.1 | 0.1 | 0.1 | -0.1 | 0.1 | 0.1 | -0.1 | 0.0 | 0.1 | 0.1 | -0.1 | 0.2 | 0.0 | 0.1 | 0.1 | 0.0 | 0.1 | -0.9 | -0.9 | 1.0 | |
| SOCstock | -0.5 | 0.8 | 0.3 | -0.1 | -0.2 | -0.2 | -0.3 | 0.5 | -0.1 | 0.4 | 0.6 | 0.1 | 0.4 | 0.2 | 0.6 | 0.1 | 0.9 | -0.1 | 0.0 | 0.2 | -0.1 | 1.0 |

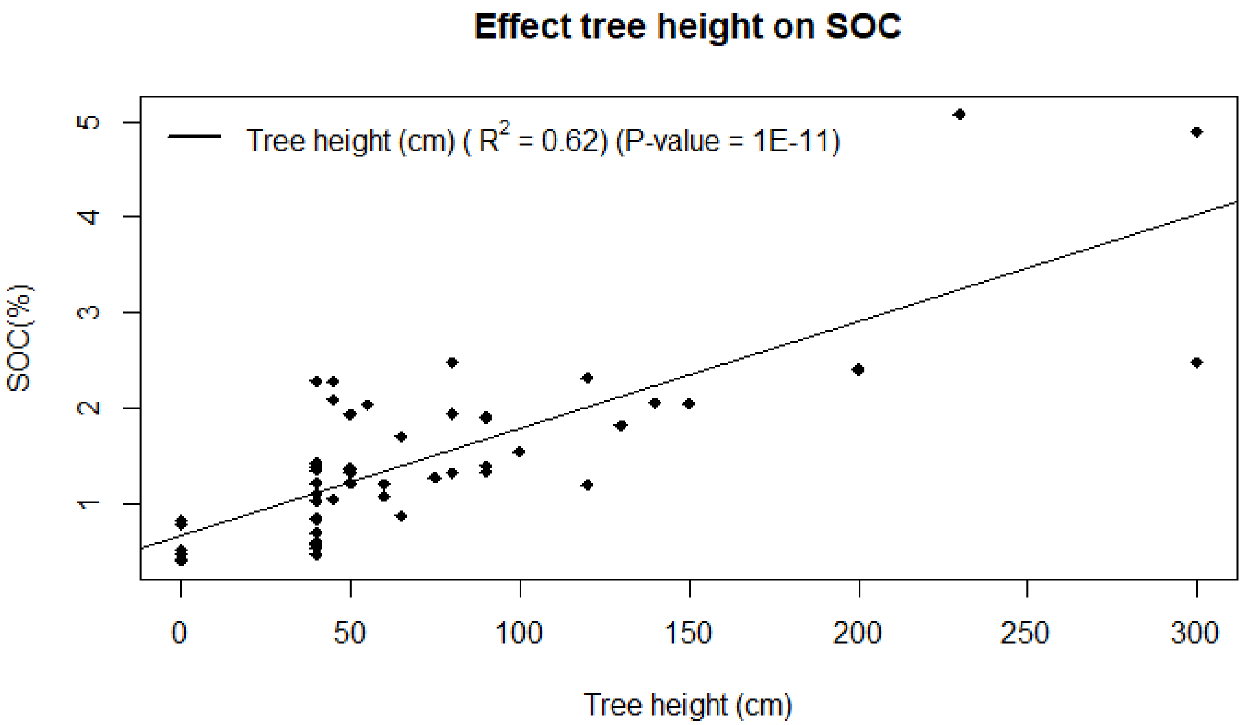


*Fig. A12 The effect of tree height on SOC fraction*

```
        Welch Two Sample t-test

data:  fClay.Other.Geologies and fClay.Shale.and.Feldspathic.Sandstone
t = -2.432, df = 34.945, p-value = 0.02028
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -4.7180871 -0.4247701
sample estimates:
mean of x mean of y 
 13.00000  15.57143 
```

*Fig. A14 Description of the ttest statistics comparing Shale with other parent materials on clay content.*

```
        Welch Two Sample t-test

data:  fSand.Other.Geologies and fSand.Quartzitic.Sandstone
t = -2.0781, df = 11.637, p-value = 0.06053
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -10.3825528   0.2635051
sample estimates:
mean of x mean of y 
 71.19048  76.25000 
```

*Fig. A15 Description of the statistics comparing Quartzitic Sandstone with other parent materials on sand content.*

*Table A12 Descriptive statistics of texture classes, stone fraction and SOC fraction for the 5 parent materials in the Baviaanskloof catchment*

| Geology type | fClay mean(St.dev) | fSilt mean(St.dev) | fSand mean(St.dev) | fStone mean(St.dev) | fSOC mean(St.dev) |
|---|---|---|---|---|---|
| Alluvium | 13.5 (3.34) | 12 (4.28) | 74.5 (7.15) | 0.2 (0.16) | 1.25 (0.59) |
| Conglomerate | 13.15 (2.38) | 12.46 (4.56) | 74.38 (5.97) | 0.4 (0.24) | 1.2 (0.57) |
| Feldspathic Sandstone | 15.46 (4.41) | 16.77 (5.2) | 67.77 (8.51) | 0.26 (0.17) | 1.44 (0.65) |
| Quartzitic Sandstone | 12.25 (3.85) | 11.5 (2.56) | 76.25 (6.04) | 0.36 (0.24) | 1.95 (1.44) |
| Shale | 15.75 (3.85) | 16 (3.7) | 68.25 (6.04) | 0.41 (0.13) | 1.85 (1.35) |

*Table A13 Values of the most important parameters determining conditions for the plots associated with an alternative degraded condition (6, 49, 3,10 and 36) compared to the mean of all plots.*

| Sample | 6 | 49 | 3 | 10 | 36 | Mean all plots |
|---|---|---|---|---|---|---|
| Tave_ave | 16.22 | 16.15 | 16.76 | 16.76 | 16.15 | 16.50 |
| rocktype | 10.00 | 4.00 | 9.00 | 10.00 | 4.00 | 6.44 |
| DEM | 430.99 | 528.01 | 371.07 | 359.62 | 504.50 | 497.44 |
| Aspect | 170.59 | 67.25 | 4.39 | 122.36 | 122.35 | 165.75 |
| Slope | 9.17 | 5.17 | 5.43 | 5.78 | 4.85 | 13.55 |
| Sol2018 | 146.88 | 135.68 | 131.73 | 128.59 | 142.41 | 130.21 |
| MUFF | 122.63 | 12.73 | 0.77 | 27.58 | 3.27 | 38.84 |
| P Mean | 239.12 | 274.27 | 273.88 | 269.84 | 273.82 | 279.41 |
| P Heavy | 164.86 | 191.30 | 163.40 | 161.93 | 191.32 | 173.42 |
| P Dry days | 311.58 | 314.54 | 294.64 | 293.14 | 314.97 | 299.79 |
| P Reg_DOR | 313.24 | 314.03 | 313.97 | 316.34 | 316.78 | 301.03 |
| NDVI mean | 0.26 | 0.25 | 0.22 | 0.24 | 0.34 | 0.29 |
| NDVI sent | 0.17 | 0.15 | 0.14 | 0.14 | 0.22 | 0.18 |
| Fstone | 0.23 | 0.09 | 0.32 | 0.06 | 0.04 | 0.33 |
| Drainage Class | 1 | 2 | 2 | 1 | 1 | 2.52 |
| Permeability | 2 | 0 | 2 | 1 | 1 | 1.72 |
| Boulder | 0 | 0 | 0 | 0 | 0 | 6.54 |
| Rock | 0 | 0 | 0 | 0 | 0 | 3.24 |
| Rooting depth | 20 | 10 | 20 | 35 | 10 | 19.98 |
| Bare soil | 35 | 65 | 30 | 20 | 90 | 49.82 |
| Tree cover | 5 | 5 | 10 | 10 | 10 | 23.58 |
| Tree height | 40 | 40 | 40 | 40 | 0 | 72.7 |
| restorationdays | 2859 | 1580 | 3163 | 3163 | 0 | 2902.92 |
| Livestock | 0 | 1 | 1 | 1 | 1 | 0.74 |
| Clay | 15 | 9 | 13 | 17 | 17 | 14.08 |
| Silt | 12 | 6 | 10 | 16 | 12 | 13.92 |
| Sand | 73 | 85 | 77 | 67 | 71 | 72 |
| pH | 5 | 7.9 | 6.6 | 7.5 | 7.9 | 6.472 |
| SOCstock | 6311 | 4463 | 5031 | 4946 | 3830 | 10995 |
| C | 0.6 | 0.48 | 0.55 | 0.58 | 0.41 | 1.49 |

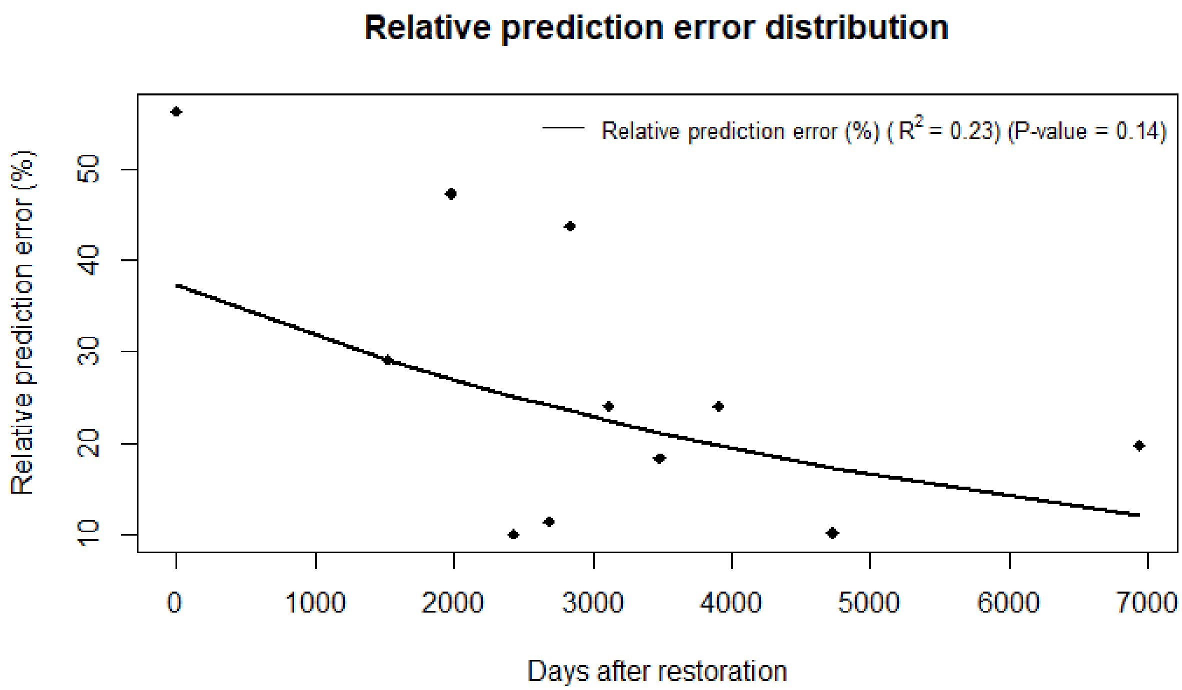


*Fig. A16 Relative prediction error of the Koster mechanistic model on plot level, between day 0 and day 7000 after restoration*

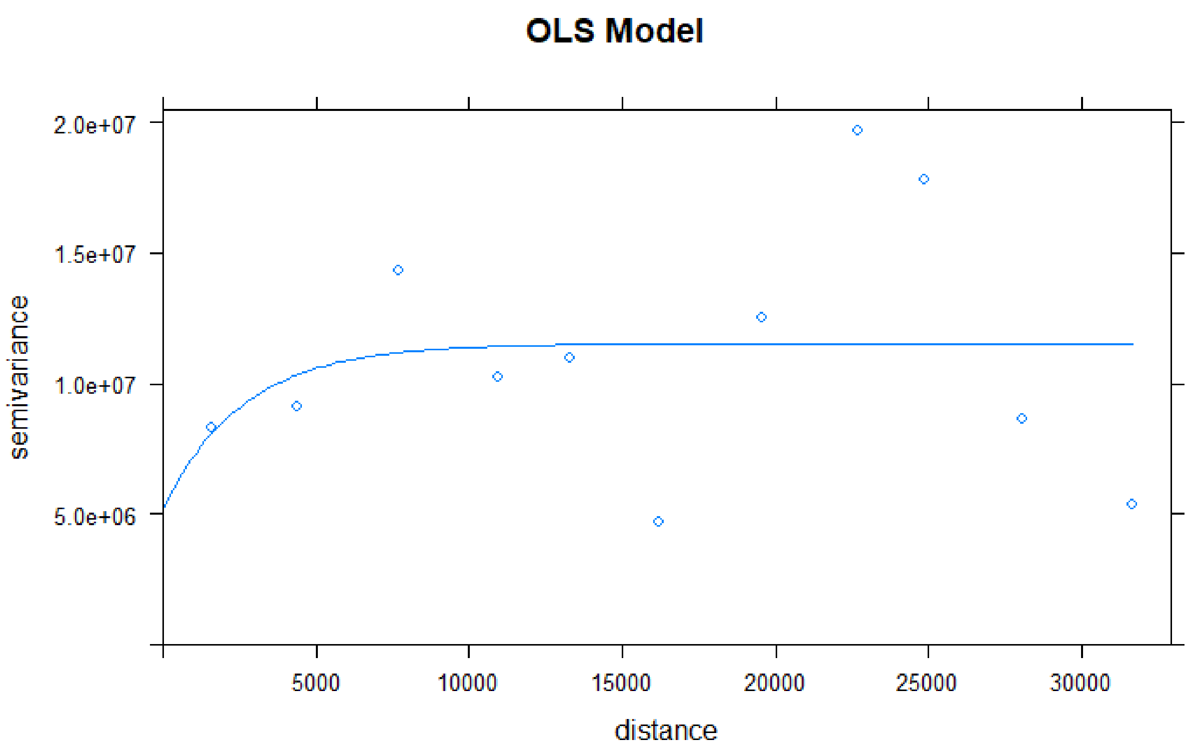


*Fig. A17 Semivariogram of residuals of the Koster model applied to the SOC values on plot level*

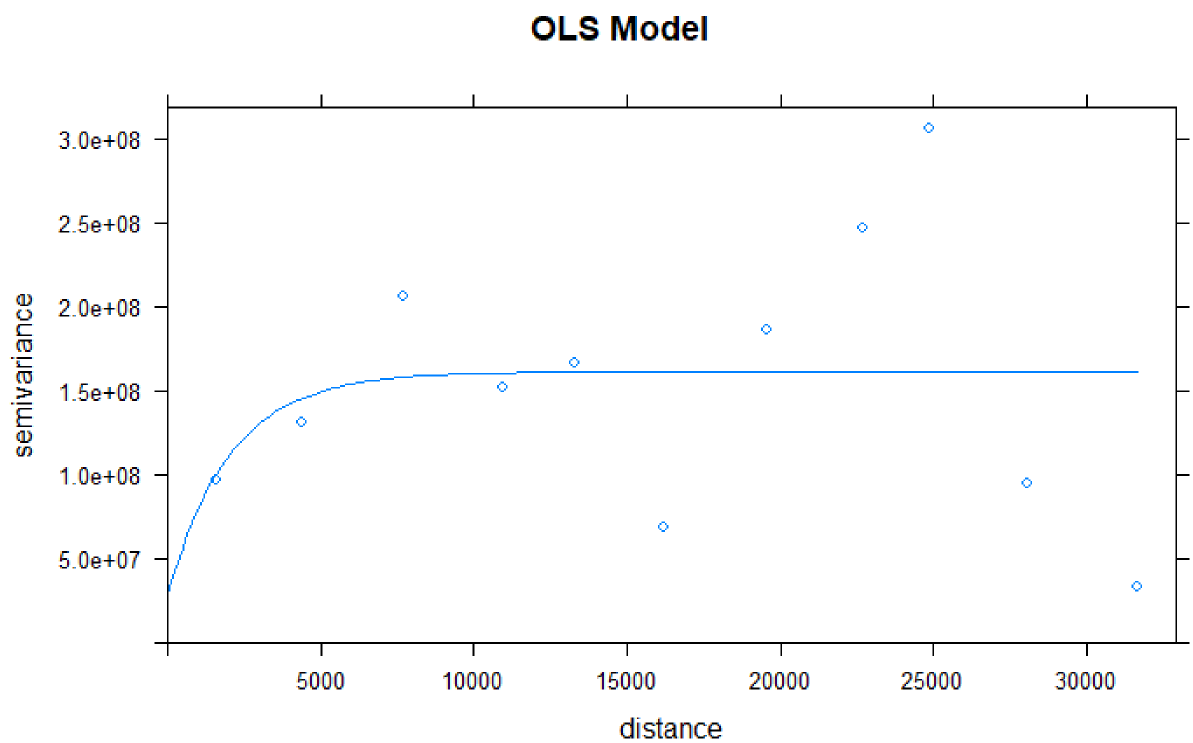


*Fig. A18 Semivariogram of residuals of the Hendriks model applied to the SOC values on plot level*