TUDelft

Bachelor Thesis Report

Statistical Analysis of Geotechnical Parameters of Starnmeer and Alkmaardermeer

Zeid Bitar
zeid.bitar@hotmail.com

Bachelor Thesis Report
Starnmeer Polder Analysis

By

Zeid Bitar

In partial fulfilment of the requirements for the degree of

Bachelor of Science
In Applied Earth Science

At the Delft University of Technology,
To be defended publicly on Monday July 10th, 2018 at 08:30 A.M

Supervisors: Prof. Michael Hicks
Tom De Gast

Abstract

The aim of this thesis report is to analyze slope stability parameters given by HHNK for the dike around a polder located in the province of North-Holland, “Starnmeer”, and a lake located to the west of it called “Alkmaardermeer”. Most of the dike has been reported to be too low in some places and plans of restoration will be implemented in 2019. A RFEM (Random Finite Elemental Method) analysis is to be done in collaboration with TU Delft and HHNK on these dikes to identify slope failure probabilities, where the input data is to be addressed in this report. Starnmeer contains thick layers of peat and clay until a depth of approximately -5 meters NAP, where a large sand aquifer begins to much larger depths. For geotechnical analysis, the first few meters are of primary concern, so only the thick peat and clay layers will be assessed. Rising water levels proves problematic for the polder, but fortunately drainage opportunities to adjacent water bodies is possible. The addition or reduction of water heavily influences soil behavior by affecting underlying pore water pressures and stress states. Parameters to be analyzed include wet bulk density, dry bulk density, water content, cohesion, and the friction angle. A statistical analysis and distribution of each parameter provide meaningful insight on probabilities and disseminations of measurement data. The clay layer was found to have a significant amount of sand and silt, affecting the porosity, sorting, cohesion, and bulk densities resulting in an uncompacted porous clay layer. The peat layer is also highly saturated, highly organic, and rather loose. The layers are hence prone to heaving and rapid primary consolidation. By classifying each of the five parameters for both regions to be normally distributed or lognormally distributed, the data can be then standardized and correlated to each other to determine linear dependencies. The relations heavily vary depending on the environmental setting and other physical aspects. For instance, the wet and dry bulk densities are found to have a strong positive correlation with each other in Starnmeer but not in Alkmaardermeer. The layers were found to be heavily saturated, which negatively influences the correlation between water content and bulk densities. Cohesion and friction angle are primarily dependent on the consistency, shape, and packing of the soil grains, so low to no correlation was established.

Table of Contents

1. Introduction:

Hoogheermraadschap Hollands Noorderkwartier (HHNK) is a government agency meant to protect, manage, and clean water sources in North Holland above the North Sea Canal. HHNK works tightly in collaboration with other municipalities and stakeholders to ensure safety regulations concerning water management are met. The area being investigated in this report, is a polder in the North Holland province called “Starnmeer”, and the adjacent lake “Alkmaardermeer”. The dikes surrounding the polder has been rejected and needs improvement to meet safety regulations. The dike is too low in some areas and is inwardly unstable for the majority of the dike according to the most recent study done in 2015 (Voorstel D&H, 2015). The dike has a high risk to push inwards towards the base with a lot of water inflow from rain or from groundwater flow. This increases the tendency of overtopping to occur. The dikes indicated in red in Figure 1 are in need of restoration:

Figure 1: Dike Stability (HHNK)

An obvious solution is to widen the inner support berm in the hinterland, but in many areas, there is very little space available for this. Plans have already been made to lay an extra layer of soil on the bank (polder side) to raise the quay about 30 cm. These plans will be implemented in 2019- 2021. However, raising the quay will not ensure dike stability, so further analysis will have to be done.

The goal of this BSc thesis is to analyze geotechnical data of dikes surrounding the polder Starnmeer and the lake located west of Oostwoulderpolder, called Alkmaardermeer. For this, HNKK will provide data for classic slope stability analysis techniques, where probability density functions and the data's correlations can be obtained to show a probabilistic overview for important geotechnical parameters. These parameters will be used in a software that carries out a more advanced analysis approach called RFEM (Random Finite Elemental Method) which fully accounts for spatial correlation and averaging and does not require assumptions for the shape and location of the failure mechanism. For this thesis, it is expected to work on these input parameters for the RFEM analysis. The input parameters to be assessed will be:

- Volumetric Saturated Weight ($\gamma_{wet}: \frac{kN}{m^3}$)
- Volumetric Dry Weight ($\gamma_{dry}: \frac{kN}{m^3}$)
- Gravimetric Water Content ($\theta: \%$)
- Cohesion ($c': \frac{kN}{m^3}$)
- Friction Angle ($\phi/tan\phi$)

These parameters were obtained in a laboratory by carrying out consolidated undrained (CU) triaxial tests on soil samples extracted from Starnmeer and Alkmaardermeer. The parameters are to be analyzed, for corresponding means, standard deviations and distributions are used to generate a correlation between each parameter. More about these parameters will be discussed in section 2.4.

In this document, the geography, geology, water balance, and soil data will be discussed to give a general overview of the region. This will be done in accordance to data and preceding research given by HHNK and TU Delft. Next, the methodology of the statistical analysis of the data is to be carried out, and, in the Results section, the corresponding outcomes of the statistical analysis will be displayed and discussed. Further discussion and concluding statements about this research and certain concerns will be addressed in the final chapters 4 and 5. Attached at the end of the document are references and supplementary figures in the Appendix.

2. Background Information:

2.1. Location:

Starnmeer is located in the municipality of Alkmaar, in the province of North Holland. It spans approximately 861 ha and bounds the large lake Alkmaardemeer. Situated close by, are villages including West-Graftdijk, Oost-Graftdijk, Oostknollendam, Westknollendam, Markenbinnen and De Woude [1]. Starnmeer consists of the former Kogerpolder, the v.m Oostwouderpolder, and Markveld. Bounding the polder, the North Holland Canal runs on the north side, Knollendammervaart on the east side, Tapsloot on the south side, and Markervaart on the west side [2]. These can be seen in the schematic map shown in Figure 2:

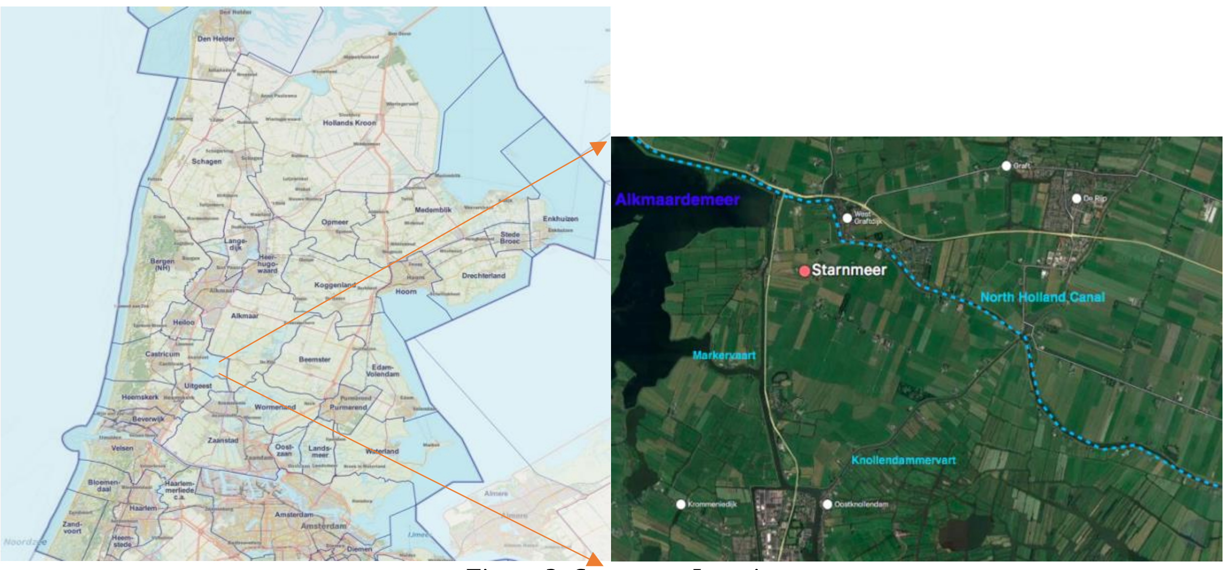


Figure 2: Starnmeer Location

Starnmeer consists of mainly grassland, with the urbanized areas being largely farms and hamlets. The soft peat and clayey soil provide ideal conditions for agriculture, where livestock farming is commonly practiced. A map showing the different land uses is provided in Appendix A under Figure 1A.

Climate conditions in the municipality of Alkmaar are rather warm and temperate (according to the Köppen-Geiger system), reaching a minimum of 0°C in winter and a maximum of 20 degrees in the summer. The rainfall is quite significant as it usually rains all year with the wettest months being September to November. The amount of rainfall hence affects the water balance in Starnmeer, which could alter several geotechnical parameters discussed in section 2.4. The climate conditions are given in Table 1:

	January	February	March	April	May	June	July	August	Sep-tember	October	November	December
Avg. Temperature (°C)	2.5	2.7	4.7	4.8	11.6	14.5	16.4	16.5	14.4	11.1	6.5	3.9
Min. Temperature (°C)	0.1	0.1	1.8	4.3	8	10.9	12.9	13	10.9	8	4	1.6
Max. Temperature (°C)	4.9	5.3	7.6	5.3	15.2	18.1	19.9	20.1	17.9	14.2	9.1	6.2
Avg. Temperature (°F)	36.5	36.9	40.5	40.6	52.9	58.1	61.5	61.7	57.9	52.0	43.7	39.0
Min. Temperature (°F)	32.2	32.2	35.2	39.7	46.4	51.6	55.2	55.4	51.6	46.4	39.2	34.9
Max. Temperature (°F)	40.8	41.5	45.7	41.5	59.4	64.6	67.8	68.2	64.2	57.6	48.4	43.2
Precipitation / Rainfall (mm)	67	45	55	42	47	52	72	78	82	87	83	73

Table 1: Temperature and Rainfall [3]

2.2. Geology:

The Starnmeer used to be, like a large part of Holland, a high peat marsh. The high peat content arose from a number of peat bogs, which contains approximately 15 cm thick top layer of hummus rich peaty clay. Over time people began to settle in the region where they removed vegetation and dug ditches to deviate the water flow to the rivers to make the land usable. The increase in urbanization and land reclamations over the years have led to peat settlement, thus causing greater concerns with the surrounding water levels. It was not until 1632 that a patent for draining was issued, and in 1636 a ring canal was built around the east side of Starnmeer to reclaim the land [4].

The lithology of the region consists of a few hundred meters of water bearing coarse sandy layers interbedded with thin clay layers. Above it resides the “Westland Formatie” which are marine and peri-marine deposits deposited during the Holocene. This formation consists of clastic marine deposits such as “Zeeklei” (sea clay) and peat layers. The peat layers from this formation are quite oxidized and weathered. Deposited on top of this is the “Hollandveen” formation which contains the largest part of peat in all of the North Holland province, then heavy clay and a top layer of peat completes the lithology to the surface.

For slope stability assessment, the upper layers are of primary concern, although when analyzing groundwater flow, the sand layers can play a crucial role. The top soil layers generally consist of peat and sea clay soils. Peat is more concentrated from the northwest of the polder to the southwest and is scarcer throughout the rest of the polder due to reclamation [5]. Therefore, clay dominates as the main soil type for the first 5 meters under the subsurface for these regions. To demonstrate the expected layers, a cross section in Figure 3 is provided from the north of Starnmeer to the south looking east, and another cross section extending from west to east looking north:

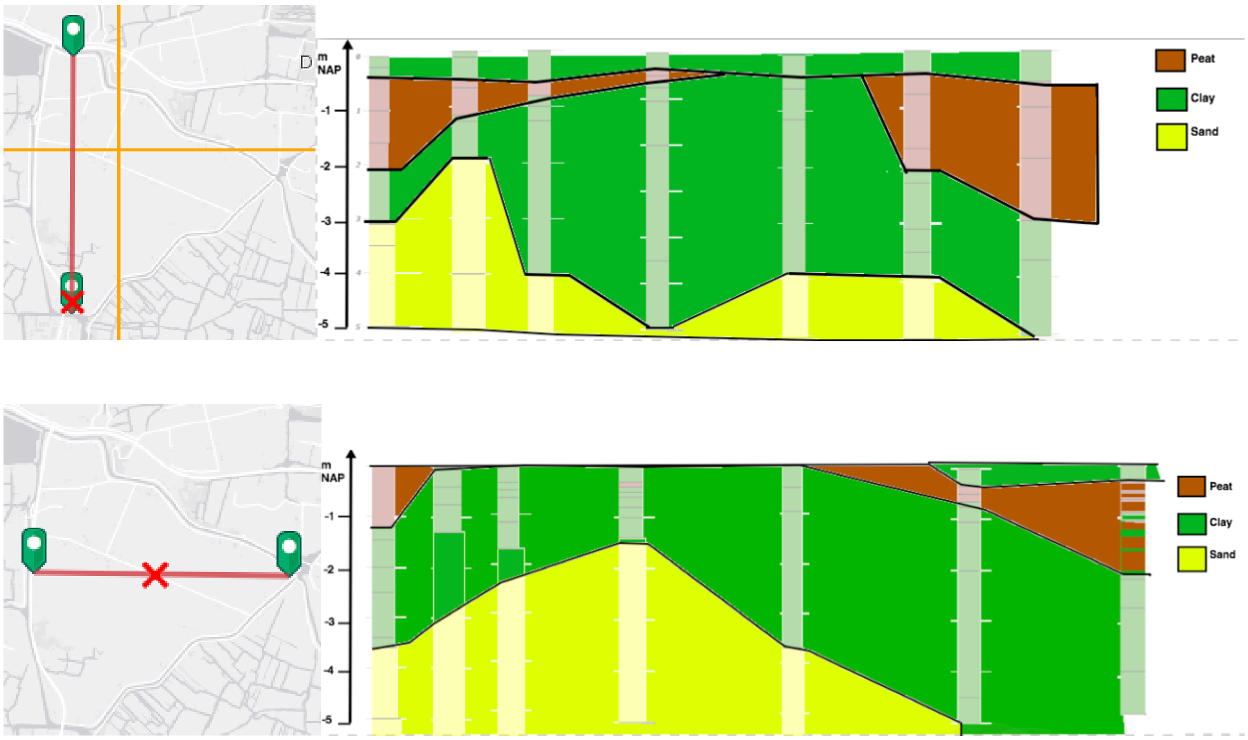


Figure 3: Cross Sections Starnmeer (Dinolocket)

This cross section has been interpolated from borehole samples using the website dinoloket.nl. It has been estimated to have linear continuous layers between each borehole sample. From this example we can see that clay is the most abundant constituent of the subsurface, with significant peat layers to the sides of Starnmeer.

2.3. Water Assessment:

Starnmeer like a large portion of Holland is under sea level, so dike construction and its stability are vital to the reclamation of the land. The dikes are needed to protect the hinterland from the water levels in nearby water bodies such as the Alkmaardermeer, Knollendammervaart canal, and the North Holland canal. The initial task therefore would be to observe the topographic height of Starnmeer in comparison to the surrounding water bodies. Through laser altimetry, a topographic map with the corresponding elevations can be generated, with more accurate elevation estimations when less obstacles obstruct the signals from reaching the ground surface (such as dense vegetation and urbanization). This is given in Appendix A.

The altitude of the area varies in the v.m. Kogerpolder from -1.40 m NAP in the northwest to -1.80 m NAP in the southeast. In the v.m. Markerpolder, variation of elevation extends from -1.00 m NAP in the west to -1.60 m NAP in the east. In the v.m. Oostwouderpolder, the topographic variation goes from -1.40 m NAP in the north to -2.40 m NAP in the south. In the v.m. Starnmeerpolder varies the height of -2.80 m NAP in the west to -4.40 m NAP in the East. [6]. These topographic variations are mostly, if not all induced by human activity.

Water levels surrounding the polder range from a height of about -1.00 m NAP to about -4.50 m NAP. The water level locations in comparison with that of the topographic elevations are given in the schematic illustration on Figure 4:

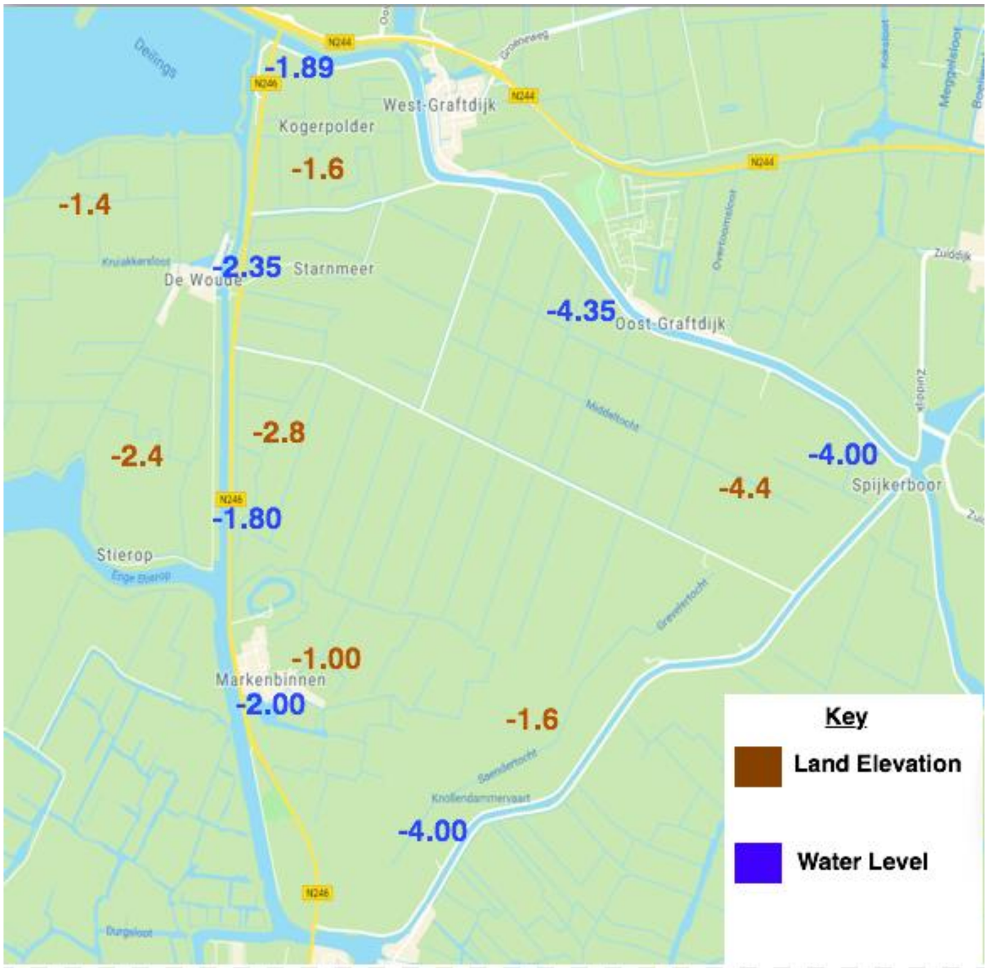


Figure 4: Water Level Relative to Topography

Note that these values are average water levels based on the information conducted by HHNK [6] in 2009. However, more recent information received by HHNK in the database had an average water level of -1 m NAP. The differences in water level around Starnmeer are controlled by a system of dams and water gates situated throughout Holland's inland waters. However, from the data received, an average water level of -1 m NAP is estimated in this document. We can see that dike stability is crucial in the in some regions of the Starnmeerpolder, due to higher water levels with respect to the land elevation. Most of the dikes around the polder are in need of maintenance as mentioned previously.

The dike's tendency to withhold surrounding water levels is not the only concern, as the geotechnical stability of the dikes is important to study. The stability of the dike has a lot to do with the hydrological balance in the region. Water through capillarity and/or seepage can affect the stress state in the soil causing instabilities by localizing strains into narrow zones inducing a slip plane to occur (macro-instability). The total stress σ_t is given by:

$$\sigma_t = \sigma' + p \qquad Eq\ (1)$$

where σ' is the effective stress and p is the pore pressure. An increase of water level induces an increase of pore pressure, which in turns causes a decrease in effective stress. When handling clay material, due to significant capillarity effect and the presence of a permeable sand aquifer below, saturated conditions are expected in the dike throughout the year. If we were to widen the support berm of the dike, we would be loading the soil. When a soil is loaded the pore pressure is then given by:

$$p = p_0 + \Delta p \qquad \textit{Eq (2)}$$

where p0 is the hydrostatic or steady state pore pressure and Δp is often referred to as the excess pore pressure. It is the additional pore pressure, relative to hydrostatic or steady state conditions, and is due to changes in load. When there are no more changes in external loading, the excess pore pressures dissipate into the underlying sand aquifer until p goes back to the steady state p0 (either to the original steady state, or to a new one arising from the engineered construction).

We can see how water content in our soil effects the stress states of the dikes, but where is this water coming from and how much do we expect? These questions can be answered with looking at the water balance in Starnmeer. The water influxes in relation with the outflows can give a notion of water levels and what to expect in the near future. From the geology, a large sand aquifer located approximately at -6 meters depth NAP supplies light brackish groundwater originating from the adjacent open waters of the North Sea, Markermeer, Alkmaardermeer, and Ijsselmeer. This was found to have an infiltration rate of 0-0.1 mm/day [7]. About 24 inlets supply the polder with water, where 10 inlets are supplied from Knollendammervaart, 6 from Noorhhollands Kanaal, 6 from Markervaart, and 2 inlets from Tapsloot. The water outflow is primarily through drainage systems and groundwater flow. The 869 ha of the inner land in Starnmeer drains by means of dike ditches, interflow, and groundwater discharge with a capacity of 88 m3/min. Due to the impermeable nature of the clay layer, the amount of interflow could have an increased effect on the drainage rate.

Further groundwater analysis has not been done in the area, where insufficient attention was paid to the prevailing water head and discharge rates due to basing the information on the 1989 water level decisions. No current monitoring wells to measure groundwater head are available as of now, but some wells are available to calculate the head only in the top confined sand aquifer at -6 meters NAP. However, more measurement data have been requested by NITG-TNO (Netherlands Institute for Applied Geosciences) and plans for this further research will be implemented by 2019.

2.4. Soil Parameters

The areas to be analyzed in this section is Starnmeer and the neighboring lake, Alkmaardermeer. The parameters given by HHNK and to be analyzed in this section are:

- Volumetric Saturated Weight ($\gamma_{wet}: \frac{kN}{m^3}$)
- Volumetric Dry Weight ($\gamma_{dry}: \frac{kN}{m^3}$)
- Gravimetric Water Content (θ: %)
- Cohesion ($c': \frac{kN}{m^3}$)
- Friction Angle ($\phi / tan\phi$)

These parameters are the main constituents for classical slope stability techniques such as Fellenius and/or Bishop methods. They give significant insight on the properties and behavior of the different soils, and how the soil behaves under different stress states. Soil material behavior is crucial in determining probabilistic failure mechanisms. The borehole sample data supplied by HHNK are all in the range of -0.5 m to about -7 m NAP with the majority of boreholes reaching about -3.2 m NAP. Consequently, very few data were given about the underlying sand layer in which the top of the layer is found to be around -5 m NAP. Therefore, only peat and clay will be analyzed, as these layers are the most abundant in both regions and are much more likely to fail.
The peat layers are about 90% organic material rather high hydraulic conductivity. If more material is to be added to the dike, the peat layers are expected to have a rapid primary consolidation and a large secondary consolidation. The clay material as expected are very fine grained, ranging from about 2 μm to about 63 μm. The grain size distribution of the clay material can be estimated from averaging a few sample points from the sieve test, by simply calculating the Trask coefficient. The original Trask Sorting Coefficient is fundamentally defined as the square root of the ratio of the 75th percentile to the 25th percentile. This is a suitable method of determining the grain size distribution for a normally distributed dataset because “the standard deviation is directly related to the percentiles because the central two thirds of the data set should lie within one standard deviation from the mean” [8]. Due to the lack of sieve data points, the Trask sorting coefficient is estimated assuming the whole clay layer has similar properties to that of the given sieve data. The Trask Coefficient is then given by:

$$\sigma = \sqrt{\frac{P75}{P25}} = \sqrt{\frac{45\ \mu m}{16\ \mu m}} = 2.8 \quad Eq\ (3)$$

We can see from the Trask coefficient that the clay material is not very well sorted. About 45% of the clay material are larger than 63 micrometers along with 37% being smaller than 2 micrometers. Therefore, we could also classify 8% of the “Clay” layer as silt, as silt is usually classified as having grain diameters between 2 and 60 micrometers. Having a large sand aquifer below the sampling boreholes and having a significant portion of the clay layer with grain diameters larger than 60 micrometers, it is appropriate to classify this layer as sandy clay.

Various lab experiments are used to calculate the different parameters mentioned in this section. The wet and dry bulk densities as well as the water content are simply measured by measuring the mass and volume of the sample and recording the resulting mass after water has been expelled from the sample. This will give the mass of the solid sample as well as the water content, so the bulk densities can simply be determined by:

$$\gamma_{wet} = \frac{m_{water}+m_{solids}}{V_{sample}} \qquad \gamma_{dry} = \frac{m_{solids}}{V_{sample}} \qquad Eq\ (4)$$

For cohesion and friction angle measurements, consolidated undrained tests (CU) are made on several soil samples, which initially entails isotropic compression to consolidate the sample, then allow drainage of the sample before applying a deviator stress (on the closed sample). The average values for Clay in Starnmeer and Alkmaardermeer and their corresponding standard deviations are given in the Table 2:

Starnmeer	$\boldsymbol{\gamma_{wet}}$ $(\frac{kN}{m^3})$	$\boldsymbol{\gamma_{dry}}$ $(\frac{kN}{m^3})$	θ (%)	$\boldsymbol{c}$ $(\frac{kN}{m^2})$	Ø (°)
Clay					
Mean Value	15.03	9.56	68.74	3.35	29
Standard Deviation	1.66	2.13	37.05	0.87	0.04
Peat					
Mean Value	10.41	2.29	467.90	3.47	0.44
Standard Deviation	0.97	1.31	226.90	1.53	0.04

Alkmaardermeer	$\boldsymbol{\gamma_{wet}}$ $(\frac{kN}{m^3})$	$\boldsymbol{\gamma_{dry}}$ $(\frac{kN}{m^3})$	θ (%)	$\boldsymbol{c}$ $(\frac{kN}{m^2})$	Ø (°)
Clay					
Mean Value	15.58	9.56	68.74	3.35	29
Standard Deviation	1.99	3.13	21.52	1.27	0.06
Peat					
Mean Value	10.58	2.17	543.70	4.20	25
Standard Deviation	1.16	1.65	190.50	1.30	0.07

Table 2: Starnmeer and Alkmaardermeer Mean and Standard Deviation

To begin with, the volumetric weight of the soil or the bulk density, is the weight of the soil in a given volume. Bulk density is different to the normal density, as it takes into consideration the pore space. Therefore, bulk density is generally less than that of the normal density of the material. Wet bulk density (γ_{wet}) gives the weight of a fully saturated unit volume of a soil sample, while the dry bulk density (γ_{dry}) is theoretically the dry weight per unit volume, even though in practice, a soil sample will always have some moisture. Bulk density is an indicator for the soils ability to function for structural support, water movement, and plant growth. As expected, the peat layers have a lower bulk density than that of clay, due to the presence of organic material, which tend to have larger pores. Is it this very reason why the peat's wet bulk density is much larger than its dry bulk density, due to water filling up the sizable pore spaces.

The dry bulk density of peat is rather low in comparison to the average bulk density of 0.5 g/cm3 most probably due to the high organic content present. The average bulk density of the clay layer is also quite low relative to average values online. The sorting of this layer is rather poor causing larger pore spaces to occur in comparison to finely packed well sorted clay grains.

The Gravimetric water saturation (∅) is the ratio of the mass of the water to the mass of the solids. Starnmeer being a polder with topographic heights lower than the water table in some areas, is expected to have soil layers with significant saturation levels. This is evidently the case, as most of the data received have saturation percentages higher than 50% for the clay layers. For the peat layers in both regions, the saturation level exceeds 100%, signifying that the water mass is larger than that of the solid. Theoretically speaking, this signifies that the water levels are higher than the level which the sample was taken from, which is obvious in the case of Alkmaardermeer (being a lake) but less so for Starnmeer. The particle density of organic soils is typically very low and is less than that of the water density of about 9.78 kN/m3. Therefore, it is not surprising to have saturation levels above 100% due to the very minute density and hence mass of the peat material.

The shear strength of the soil primarily depends on two properties, the cohesion and the frictional resistance between the particles. For finer grained softer material like peat and clay, cohesion plays a more important role in determining geotechnical stability. The cohesion is defined as the sticking together of particles of the same substance. It results from the common attraction on a microscopic level of similar materials. The clay material has undoubtedly small values for cohesion, and in Alkmaardermeer, even lower than that of peat. A logical explanation for this could be due to the variability of soil material in the clay layer. As defined before, the clay layer contains a significant amount of sand and silt, which can therefore reduce the cohesiveness of the clay layer. Furthermore, an assumption can be made that the layer is uncompacted, as compaction regularly upsurges the cohesiveness of clay. From multiple TC and DSS tests of peat soils around the Netherlands [9], peat cohesion values typically hover between 0 -10 kPa for the single stage tests shown on the graph B1 in Appendix B. The high moisture content can further increase cohesiveness of the organic material.

Along with cohesion, the friction angle between the soil particles is used in many geotechnical methods for stability, the most prominent one being the Mohr-Coulomb mathematical model. The friction angles of both peat and clay in both regions are rather low, which also exemplifies the assumption of loose soil:

Soil Packing	**Friction Angle (°)**
Very loose	<30
Loose	30-35
Compact	35-40
Dense	40-45
Very Dense	>45

Table 3: Correlation between Soil Packing and Friction Angle (Meyerhoff 1956)

The graph in Appendix B does show typical values of peat to hover around 30. Fibers in the peat usually causes a higher friction angle [10].

From the soil parameters, a general conclusion about the soil layers can be made. Note that chemical and physical factors such as diagenetic processes or human induced compaction can have an adverse effect on soil behavior that is not taken into consideration in this analysis. The generalizations made in this section are purely based on geotechnical data received from HHNK and researched information of the area. The clay layers in both regions seem to be uncompacted, mostly saturated, and have a high variability of grain sizes. The peat layers are highly organic, fibrous, and loose soil that has a high saturation. With further influx of water from surrounding water bodies can alter the physical and chemical nature of these soils.

3. Methodology:

3.1. Theory

Now that the region is familiarized, and raw soil parameters are assessed, a detailed statistical analysis can be done. This analysis is firstly meant to obtain more insight on the range, average, and unconventionalities of the data. This can provide more detailed information of how the probability of how a soil will behave depending on the factor being analyzed. The distribution of real data is generally not known, so it is apparent that the data must be in the right distribution in order to compare and/or correlate between different parameters. According to the “Central limit Theorem”, any independent identically distributed data have approximately normal distributions [11]. The normal distribution arranges the data in such a way that the majority of values cluster in the middle of the range and the rest tamper off symmetrically towards either extreme. Normalizing and reconstructing data in this way helps reduce data redundancy and improve data integrity. A normal distribution can be represented by a probability density function (PDF) for continuous variables:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Eq (5)

The probability that a variate has the value x is expressed by the are underneath the pdf or the integral between two points:

$$\int_a^b f(x)dx = \Pr[a \leq X \leq b] \qquad \textit{Eq (6)}$$

If handling discrete variables, the distribution becomes a probability mass function (PMF) which is the probability that the function takes the value x. However, it is often preferred to use a continuous function to retrieve approximated values or probabilities rather than no information between data points.

Another way to represent data is the integral of the probability density function, called the cumulative density function (CDF). The CDF is the probability that the variable takes a value less than or equal to x:

$$F(x) = \Pr[X \leq x] \qquad \textit{Eq (7)}$$

Sometimes, raw data do not coincide very accurately with their normal distributions. It is therefore convenient in this case to normally distribute the logarithms in a Log-Normal Distribution. If a random variable X is lognormally distributed, then Y=ln(X) is normally distributed. Likewise, X=exp(Y) has a lognormal distribution. A lognormal distribution probability density function has a probability density function different to that of the normally distributed PDF. See Figure 5:

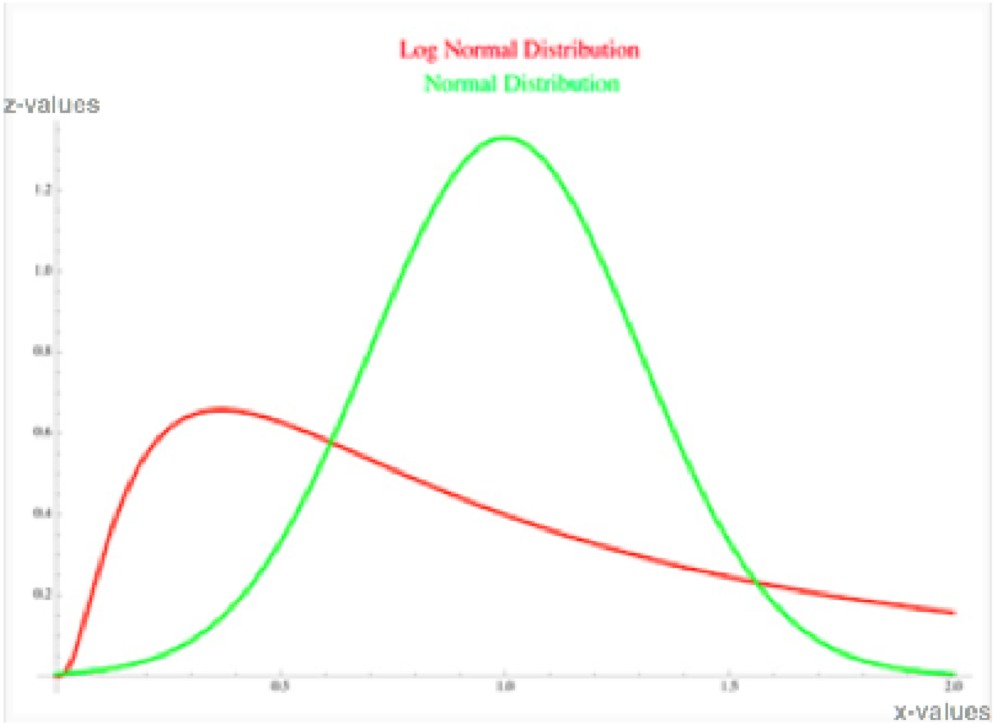


Figure 5: Log Normal and Normal Distributions (Instream solutions [12])

We can see that the curve is positively skewed, so it is more probable to return a value below the average value for the normal distribution but are also more likely to return a value above the average value for the normal distribution. Skewed distributions with low mean values, a large variance, and all positives values often fit this distribution [13]. Normal distributions take into consideration both positive and negative values. Therefore, if the error of the data is large in comparison to its absolute value then the data would probably be normally distributed. On the other hand, the logarithm of negative values is not valid, so it is very unlikely to return a lognormal distribution from negative data values. However, the data being analyzed in this report all have positive values, so the distribution selection for different data values will need to be determined by another approach.

To test if the data are normally distributed or log-normally distributed we need to determine how well the data coincide with each distribution accordingly. This can be done for example by visually observing the CDF of the data and determining if the log distribution or the normal distribution fits best. However, most probably the log distributed function is very similar to that of the normal distribution, so another method has to be used to test whether the sample data is consistent with the hypothesized distribution. Fortunately, there are many methods to test whether a given distribution is suited to a data set, where the Chi-Squared Goodness of Fit test will be implemented in this reports analysis. The Chi-Squared test makes use of expected and observed frequency counts for each parameter. Before beginning the test, two hypotheses are usually made:

- H0: Data are normally distributed
- H1: Data are not normally distributed (in this case, lognormally-distributed)

The chi-squared test is essentially the sum of differences between observed and expected frequency counts squared, divided by the expected frequency:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \qquad Eq\ (7)$$

where:

- Oi is the observed frequency count for bin i
- Ei is the expected frequency count for bin i

We can see by the difference in the formula that the closer the value is to zero, the more suitable the distribution is to the data. Sampled data will never be negative because of the squared term. By relating the normal distribution's chi-squared to the log-distribution' chi-squared, the best type of distribution can be determined.

Along with the distributions, the data values are often standardized to obtain z values which can be compared to each other to obtain a more meaningful correlation between different parameters. Correlation gives insight on how strongly pairs of variables are related [14]. This is often represented by a correlation coefficient, which is in between -1 and 1. The closer the value is to either bound, the more strongly correlated the variables are, with 0 showing no correlation, -1 showing perfect negative correlation, and +1 having perfect positive correlation. In this evaluation, a linear correlation called the "Pearson's correlation" will be used, where the coefficient is the covariance between the variables divided by the product of their standard deviations:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \qquad Eq\ (8)$$

Correlation can also be described by the r-squared method. This is simply the square of the correlation coefficient and measures how close the data are to the fitted regression line. It is measured in percentage, where the higher the percentage, the better the correlation. 0% indicates that the model indicates no variability about its mean, while 100% shows all variability about the mean [15]. A disadvantage of r-squared in comparison to the Pearson Correlation method is that it does not signify whether the variables in question are positively or negatively correlated and is not used in the RFEM Analysis. However, for purely visualization purposes, both the Pearson's correlation and the r-squared are displayed in Chapter 4 to give a visual representation of the correlations. Once the data is distributed and correlated correctly, the input parameters for the RFEM Analysis will be ready.

3.2. Practical

The previously mentioned statistical analysis can all be done in Microsoft Excel. It is here where various functions are utilized in order to obtain the desired results. HHNK has supplied a large database consisting of various parameters, categorizations, and information about the Starnmeer dikes where the data was extracted and sorted. The mean and standard deviations were then calculated using the AVERAGE() and STDEV() functions. Next the cumulative probability, that is the count number divided by the number of counts, was taken as an input to the NORM.INV() and LOGNORM.INV() functions, where they generate an inverse cumulative distribution function. Contrary to the CDF, the inverse functions return an x value given a probability. Excel uses an iterative method to calculate the Norminv function and seeks to find a result, x, such that [16]:

NORMDIST(x, mean, standard_dev, TRUE) = probability *Eq (9)*

For the LOGNORM.INV() function, the data's mean and standard deviation needs to be transformed before implementing the function. The transformations make use of the mean and standard deviation of the data:

$$\mu\_trans = \ln(\mu^2/\sqrt{\sigma + \mu^2}) \qquad Eq\ (10)$$

$$\sigma_{trans} = \sqrt{\ln(\frac{\sigma}{\mu^2+1})} \qquad Eq\ (11)$$

The CDF and PDF functions could then be easily plotted. Next is to test whether the a normal or log-normal distribution can fit each parameter. For this we perform the Chi-Squared goodness of fit test, where the data is firstly segregated into bins. The first bin contains the minimum value of the normal/log distribution minus a very small number, and the last bin contains the maximum value of the normal/log distribution plus a very small number. The purpose subtracting and adding the small number is to ensure that all data is taken into account by omitting rounding errors. The number of bins was taken to be 10, where for each bin, a narrow range between two values is compared to that of the data value. If the data value falls in this range, a frequency count of 1/(# of counts of NORM.INV or LOG.INV), if not; the value is taken to be zero. The summation over the data set for each bin is taken to be the observed frequency count. The expected frequency count is determined by making use of the NORM.DIST() function that calculates the probability of obtaining a certain value. For each bin the probability of obtaining the higher value minus the probability of obtaining the lower value gives the expected frequency count. Finally, the chi-squared can easily be calculated and compared between the normal distribution and the log distribution.

The built-in function CORREL is used to find the Pearson's correlation coefficient between two data sets. If the data is lognormally distributed, the logarithms of the data set are used in the correlation. The correlation value can simply be squared or using the built-in function RSQ to obtain the r squared percentage.
A schematic flow diagram showing the process of the statistical assessment is shown in Figure 6:

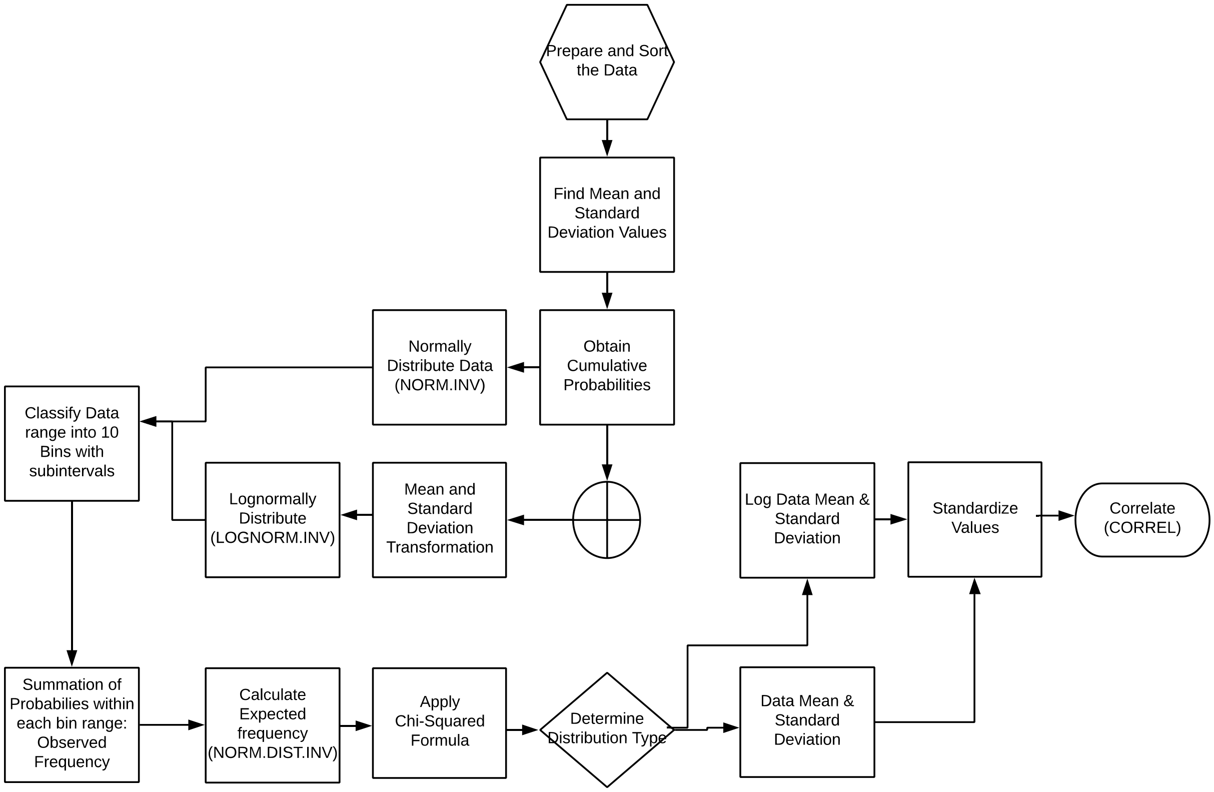


Figure 6: Practical Methodology Flow Chart

4. Results and Discussion:

4.1. Distribution:

The distributions of the five parameters are determined to be either normally distributed or log-normally distributed by performing the chi-square goodness of fit test. The resulting chi-squared values for Starnmeer and Alkmaardermeer is shown below in Tables 4 and 5:

Starnmeer Chi-Squared Clay	**Normal Chi-Squared**	**Log-Normal Chi-Squared**	**Resulting Distribution**
γ_{wet}	0.23	0.31	Normal
γ_{dry}	0.19	0.66	Normal
∅	0.17	0.11	Log-Normal
c	1.68	2.04	Normal
tanθ	2.77	2.69	Log-Normal

Starnmeer Chi-Squared Peat	**Normal Chi-Squared**	**Log-Normal Chi-Squared**	**Resulting Distribution**
γ_{wet}	0.19	0.27	Normal
γ_{dry}	0.87	0.71	Log-Normal
∅	0.47	0.37	Log-Normal
c	4.31	3.12	Log-Normal
tanθ	3.99	3.86	Log-Normal

Table 4: Starnmeer Distributions

Similarly, for Alkmaardermeer:

Alkmaardermeer Chi-Squared Clay	Normal Chi-Squared	Log-Normal Chi-Squared	Resulting Distribution
γ_{wet}	0.19	0.17	Log-Normal
γ_{dry}	0.23	0.26	Normal
∅	0.10	0.03	Log-Normal
c	4.54	3.15	Log-Normal
tanθ	3.46	2.68	Log-Normal

Alkmaardermeer Chi-Squared Peat	Normal Chi-Squared	Log-Normal Chi-Squared	Resulting Distribution
γ_{wet}	0.42	0.40	Log-Normal
γ_{dry}	1.64	1.08	Log-Normal
∅	0.37	0.24	Log-Normal
c	4.61	4.58	Log-Normal
tanθ	5.02	4.16	Log-Normal

Table 5: Alkmaardermeer Distributions

4.2. Correlation

To see how one parameter changes in variation of another parameter, the correlation coefficient is calculated. Only linear correlations can be represented by these coefficients. The correlation coefficients between the different parameters are calculated and represented in Tables 6 and 7:

- For Starnmeer:

Correlation Starnmeer Clay	γ_{wet}	γ_{dry}	Water Content	Cohesion	Friction Angle
γ_{wet}	-	0.99	-0.97	0.12	0.35
γ_{dry}	-	-	-0.99	0.03	0.44
Water Content	-	-	-	0.00	0.48
Cohesion	-	-	-	-	-0.75
Friction Angle	-	-	-	-	-

Correlation Starnmeer Peat	γ_{wet}	γ_{dry}	Water Content	Cohesion	Friction Angle
γ_{wet}	-	0.87	-0.84	0.34	0.24
γ_{dry}	-	-	-0.99	0.55	-0.04
Water Content	-	-	-	-0.56	0.04
Cohesion	-	-	-	-	-0.78
Friction Angle	-	-	-	-	-

Table 6: Starnmeer Correlation

- For Alkmaardermeer:

Correlation Alkmaardermeer Clay	γ_{wet}	γ_{dry}	Water Content	Cohesion	Friction Angle
γ_{wet}	-	0.76	0.59	-0.43	0.68
γ_{dry}	-	-	0.17	-0.44	0.66
Water Content	-	-	-	-0.10	0.07
Cohesion	-	-	-	-	-0.91
Friction Angle	-	-	-	-	-

Correlation Alkmaardermeer Peat	γ_{wet}	γ_{dry}	Water Content	Cohesion	Friction Angle
γ_{wet}	-	0.43	0.80	-0.73	0.30
γ_{dry}	-	-	-0.28	-0.71	0.32
Water Content	-	-	-	-0.67	0.66
Cohesion	-	-	-	-	-0.99
Friction Angle	-	-	-	-	-

Table 7: Alkmaardermeer Correlation

Firstly, the dry and wet bulk densities are expected to be more or less positively correlated. An increase in the weight of the soil should result in an increase in both dry and wet bulk densities. This is seen in clay more-over than in peat for Starnmeer. Due to the light weight property of the solid organic material in peat, any change in weight of the soil mass will most likely result from water weight. This results for a lower correlation percentage between dry bulk density and wet bulk density. In Alkmaardermeer, the dry and wet bulk densities show to have little to no correlation, due to the oversaturation of both soil layers, and it is this reason that the water content and bulk densities for Alkmaardermeer show no correlation. The samples taken for this region are already submerged, so no variation of water content can take place naturally. On the other hand, in Starnmeer the wet and dry bulk densities show a negative correlation with the water content. This is obvious for γ_{dry}, because the denominator total volume will change with changing water volumes, but the numerator solid soil mass will not vary. In Triaxial Tests however, the volume of the sample in the sample holder remains fixed, so this explanation is not applicable. The negative correlations for the bulk densities are therefore quite unexpected. An increase in the amount of gravimetric water content lubricates the soil particles and brings the soil particles closer together (compaction) reducing the soil-water-air mixtures into a denser state hence increasing the dry bulk density [17]. A logical explanation of why this is not the case with this clay layer is that the optimum moisture content has been reached and surpassed. Addition of water beyond the optimum moisture content will in fact decrease the bulk density due to the water occupying the voids in which the clay particles could have occupied:

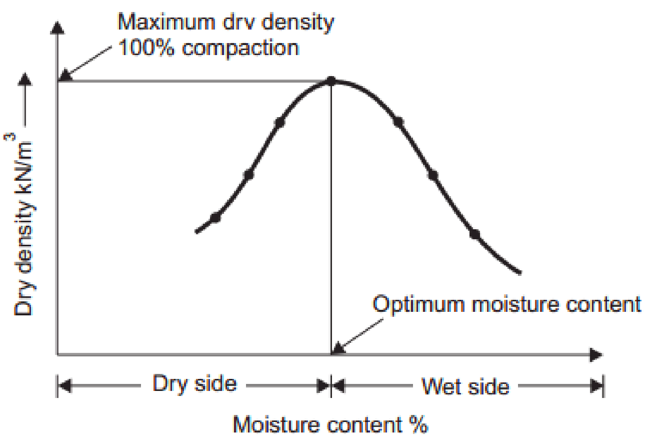


Figure 7: Moisture content vs dry bulk density

Water typically having a lower density than that of clay particles, would decrease the overall dry bulk density of the soil. The dry bulk density and wet bulk density are proportional to each other represented by the formula:

$$\gamma_{dry} = \frac{\gamma_{wet}}{(1+\theta)} \qquad Eq\ (12)$$

The bulk density and water content of both peat and clay show low to no correlation with cohesion and friction angle. Cohesion is mainly influenced by other physical properties such as aging, consistency and water suction. Therefore, it shows minute correlation to all parameters except with friction angle, with the strongest correlation in Alkmaardermeer. Cohesion and friction angle are two independent physical properties and depend on other various factors. A negative correlation between these two parameters cannot symbolize a direct physical relationship, but according to a geotechnical study of an interpretation of cross-correlation of cohesion and friction angle for cohesive slopes [18], a negative correlation was found to decrease the probability of failure. This conclusion was drawn from numerous stability charts, Monte-Carlo simulations, and slope stability techniques.

Some limitations to this correlation analysis include:

- All correlations are represented by correlation coefficients, so only linear correlations can be done. This means that the ratio of change is constant, so inconstant change cannot be interpreted by this method.
- Numerous considerations are not taken into account. The correlation only relates how one parameter changes linearly with another even though other physical and chemical factors heavily influences the relationship. Therefore, a thorough analysis of soil behavior is limited.
- The correlation coefficient does not indicate causality, so a strong relationship between two parameters does not necessarily imply that these parameters are responsible for each other.

5. Conclusion:

The scope of this Thesis is to assess borehole data in Starnmeer and Alkmaardermeer received by HHNK by carrying out a statistical analysis of the dry and wet bulk densities, gravimetric water content, average cohesion, and average friction angle. These parameters are the basis to carry out a RFEM slope stability analysis of the dikes surrounding the polder. From each soil parameter, certain assumptions can be drawn about the geotechnical characteristics of each soil layer which can further be justified by making an overall interpretation of all the parameters together. The distributions of these parameters were then interpreted to be normally or lognormally distributed to reduce data redundancy, then were correlated to each other to identify any linear relationships.

The depth of soil investigation was limited to a maximum of -6 meters NAP. Peat and Clay layers take up the majority of this depth. Both layers are highly saturated with peat having a higher water to solid ratio in Alkmaardermeer. The clay can be described as poorly sorted with relatively large pores resulting in lower bulk density and cohesion values. Similarly, the peat layers contain high organic content with large pores, so care will need to be taken when widening the dike in 2019. Peat and Clay are especially prone to consolidation.

The distribution each parameter for the peat and clay layers were determined by observing their respective goodness of fit the measurement data. The distributions are characterized as either normal or lognormal distributions. The distributions provide insight on the average values and how the data deviates from this value. This reduces the workload when working with data as it helps to understand the probabilities of obtaining a certain value. The distributions are then used to correlate between different parameters. Lognormally distributed data have their standardized logarithms correlated to other data. From the correlation results, dry and wet bulk densities for Starnmeer show positive linear relations in contrast with Alkmaardermeer which show almost no correlation. The environmental setting of the soil layer hence affects the soil behavior considerably. Peat and clay layers in Starnmeer show to have high saturation levels which negatively influences the soil’s dependency on bulk densities as the correlation coefficient tends toward negative one. Cohesion and friction angle of Alkmaardermeer shows strong correlation, larger than that of Starnmeer.

References:

1- "Starnmeer" (April 9th, 2018). Retrieved from URL: http://www.starnmeer.nl/
2- Mantel, C., Hoogheemraadschap Hollands Noorderkwartier Peilbesluit Starnmeer May 1st, 2009.
3- Climate-data, Openstreetmap, Retrieved from URL: https://en.climate-data.org/location/11254/
4- Mark, C., Collection Regional Archive Alkmaar, 1997.
5- Stichting Bodem Kartering, De Bodemkaart van Nederland, 1:50.000, 1984,
6- Actueel Hoogtebestand Nederland (AHN). Retrieved from URL: http://www.ahn.nl/index.html
7- Van Oostrum, A.E.M. , HNKK Peilbesluit Starnmeer, October 2009
8- Conway Valley Systems, "Correlation", 2016. Retrieved from URL: http://ws2.petrog.com/articles/trask.html
9- Tomczak, A., "Peat dikes and strength parameters" University of Groningen, 1996
10- Al-Ani, Haider, Erwin, Chai, Gary, "Engineering properties of peat in estuarine environment". Foundation & Soft Ground Engineering Challenges in Mekong Delta, 2013.
11- Investopedia, "Normal Distribution". Retrieved from URL: https://www.investopedia.com/terms/n/normaldistribution.asp
12- Instreamsolutions 2008, Retrieved from URL: https://support.instreamwealth.com/hc/en-us/articles/202386009-What-is-the-difference-between-a-Normal-Distribution-and-a-Lognormal-Distribution-
13- Kalecki, M. On the Gibrat Distribution. Econometrica. Log-normal Distributions across the Sciences: Keys and Clues. Vol. 13, No. 2 pp. 161-170, Apr. 1945
14- Creative Research Systems, "Correlation" 2016. Retrieved from URL: https://www.surveysystem.com/correlation.html
15- The Mini Tab Blog, "Regression Analysis" Retrieved from URL: http://blog.minitab.com/blog/adventures-in-statistics-2/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit
16- Al-Ani, Haider, Oh, Erwin, Chai, Gary. (2013). "Engineering properties of peat in estuarine environment". Foundation & Soft Ground Engineering Challenges in Mekong Delta
17- The constructor Civil Engineering Home, "Soil Moisture and dry density relationship", Retrieved from URL: https://theconstructor.org/geotechnical/soil-moisture-content-dry-density-relationship/6947/
18- Javankhoshdel, Sina; Queen's University, Civil Engineering Dept Bathurst, Richard; Queens University/Royal Military College. Canadian Geotechnical Journal, "Influence of cross-correlation between soil parameters on probability of failure of simple cohesive and c-phi slopes", 11-Aug-2015.

Appendix A:

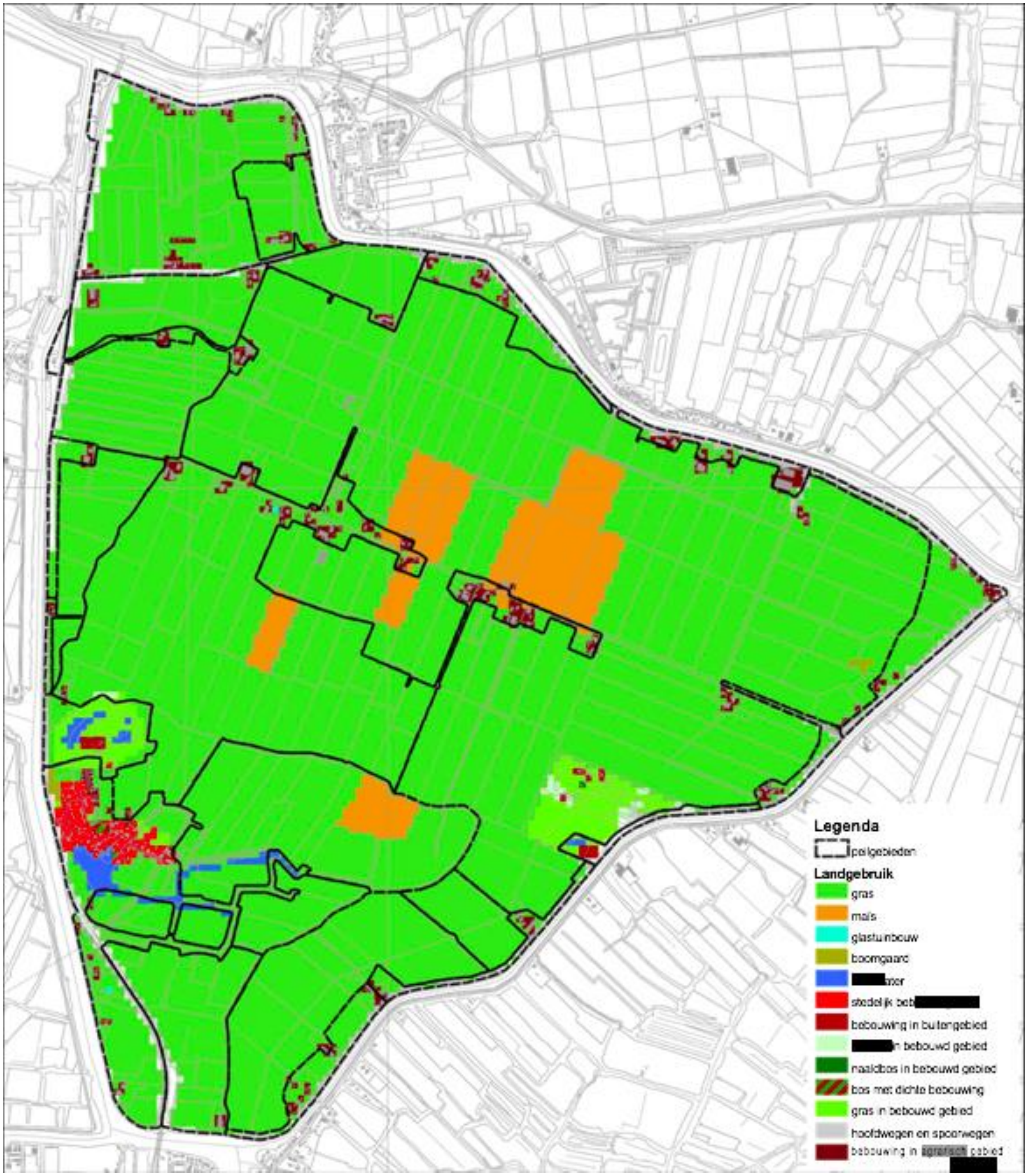


Figure A1: Land use (LGN5)

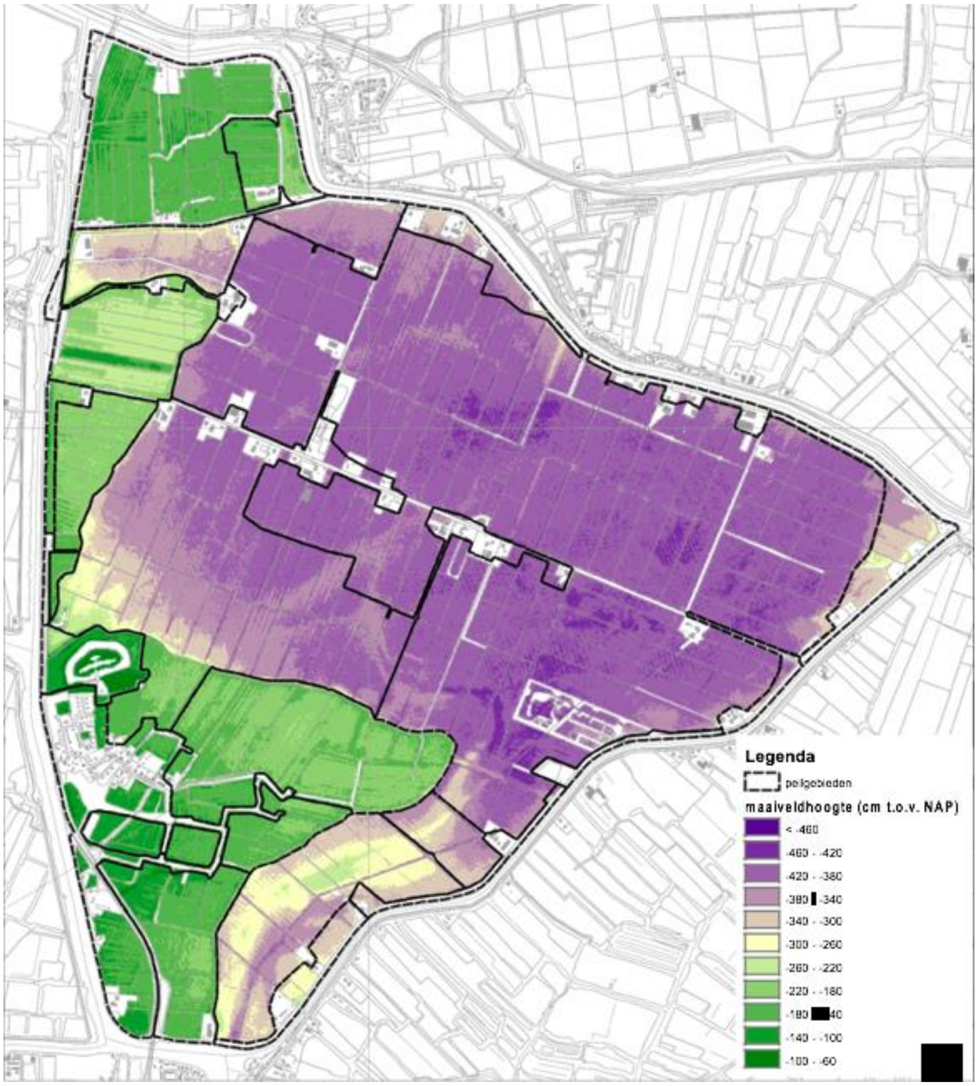
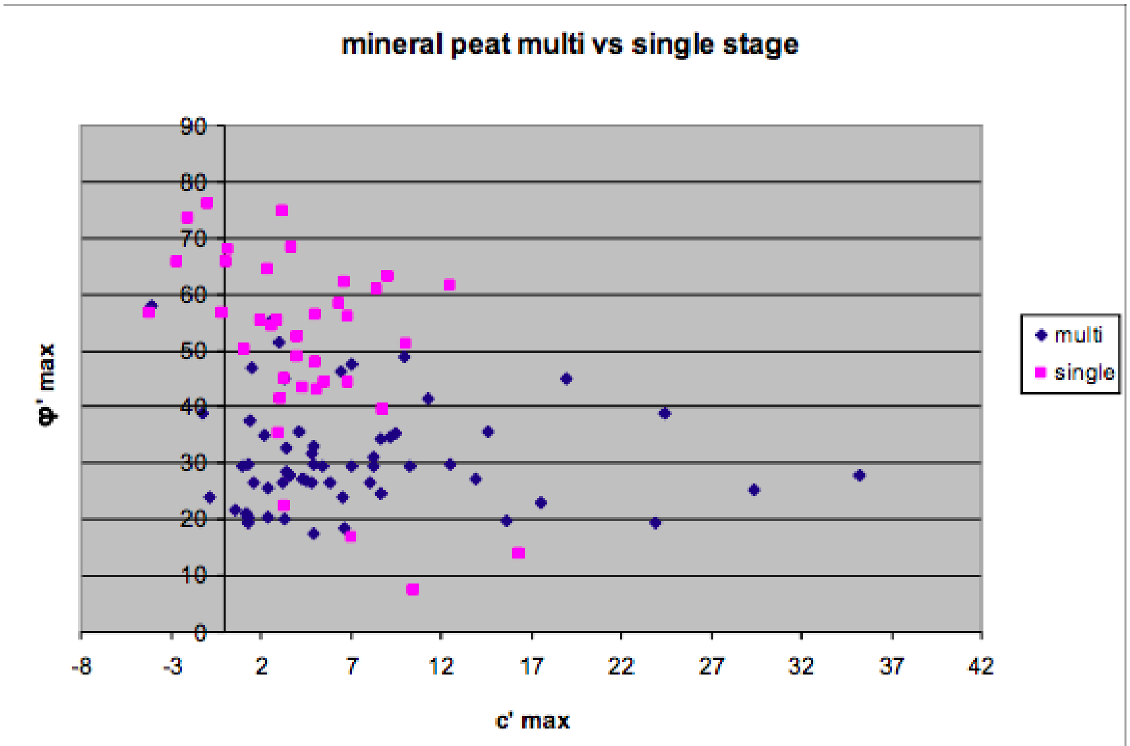


Figure A2: Starnmeer Height Level

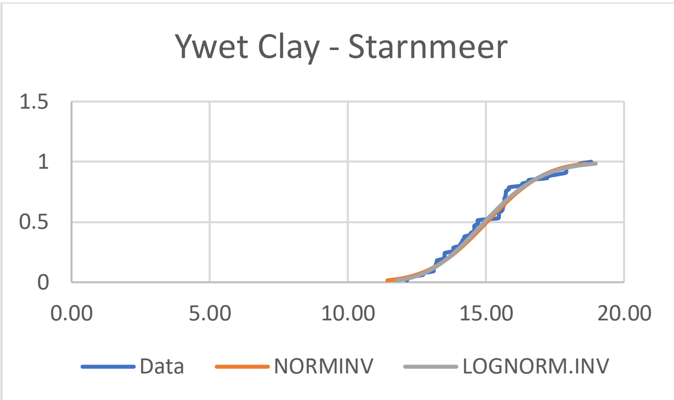
Appendix B:

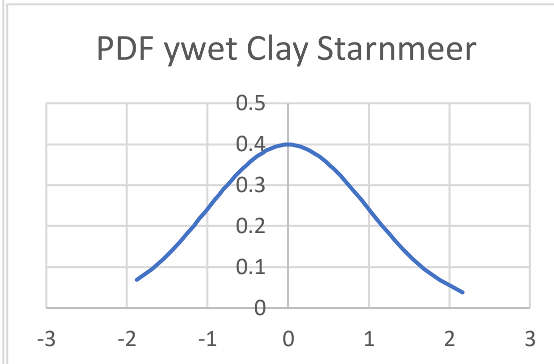


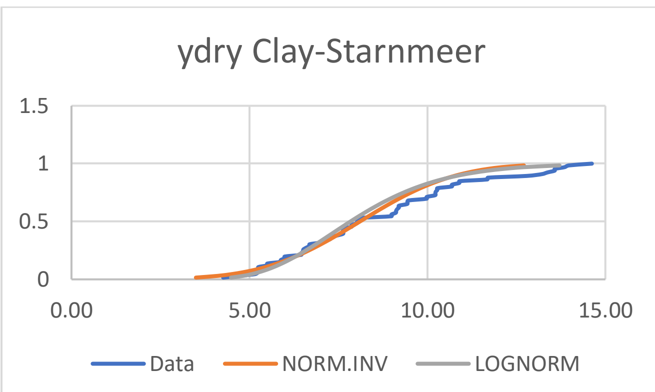
Graph B1: Results from multi-stage and single stage TC-Tests (A.Tomczak)

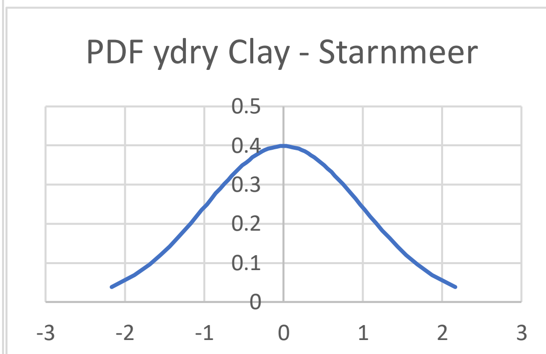
Appendix C:

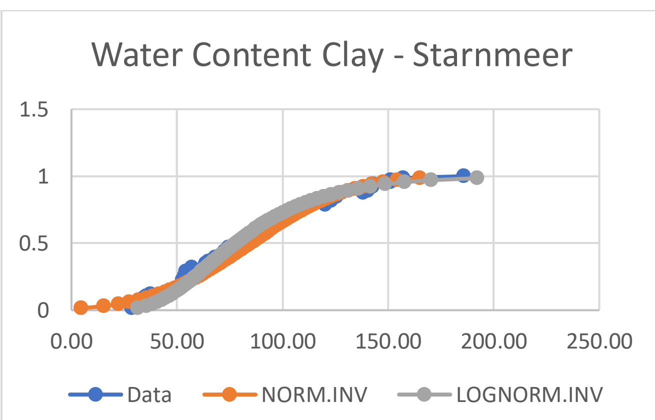
Starnmeer:

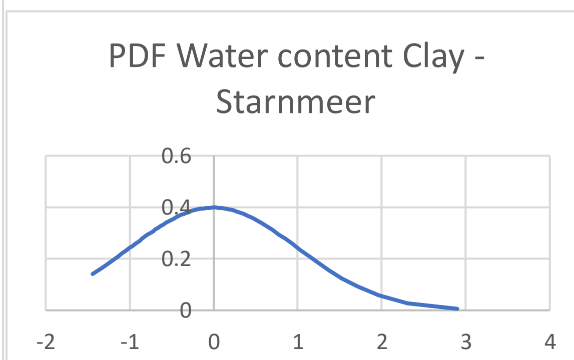


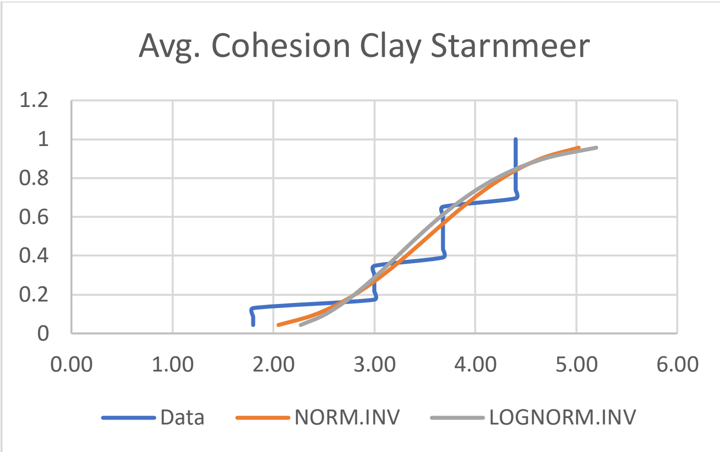


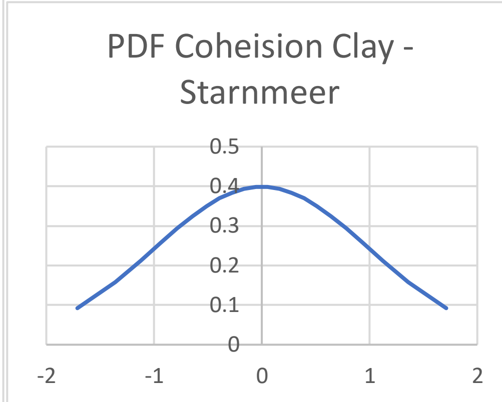


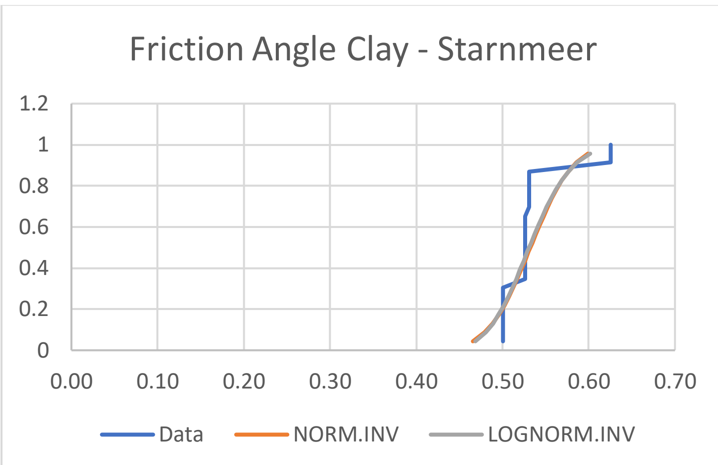


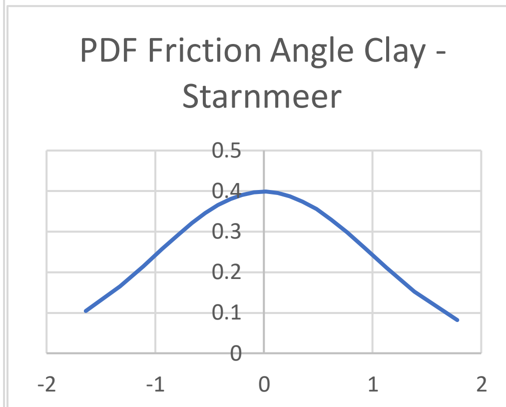




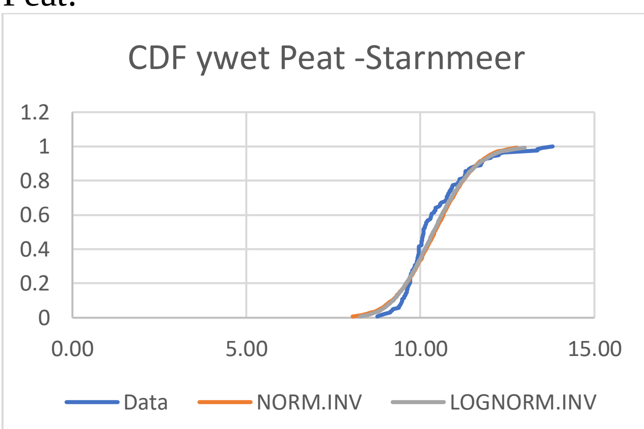

Avg. Cohesion Clay Starnmeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
1.00
2.00
3.00
4.00
5.00
6.00
Data
NORM.INV
LOGNORM.INV

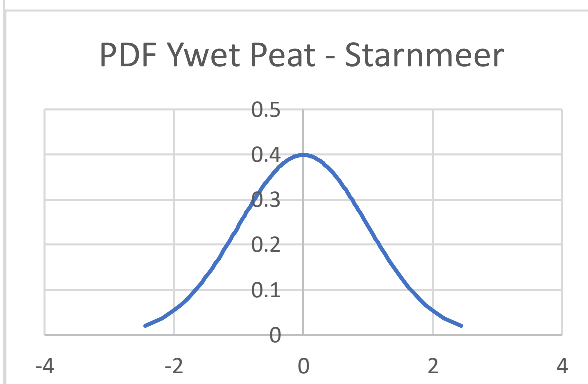

PDF Coheision Clay - Starnmeer
0.5
0.4
0.3
0.2
0.1
0
-2
-1
0
1
2

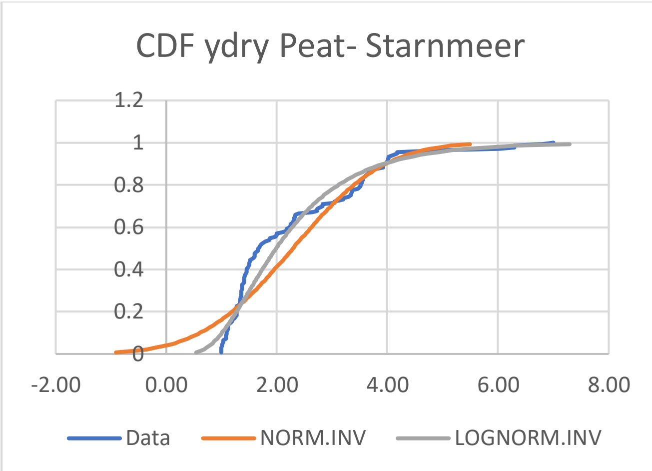

Friction Angle Clay - Starnmeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
Data
NORM.INV
LOGNORM.INV

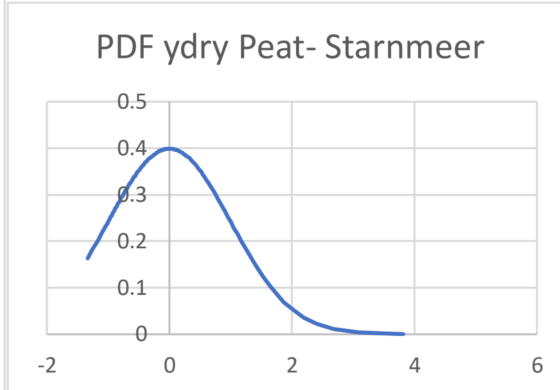

PDF Friction Angle Clay - Starnmeer
0.5
0.4
0.3
0.2
0.1
0
-2
-1
0
1
2

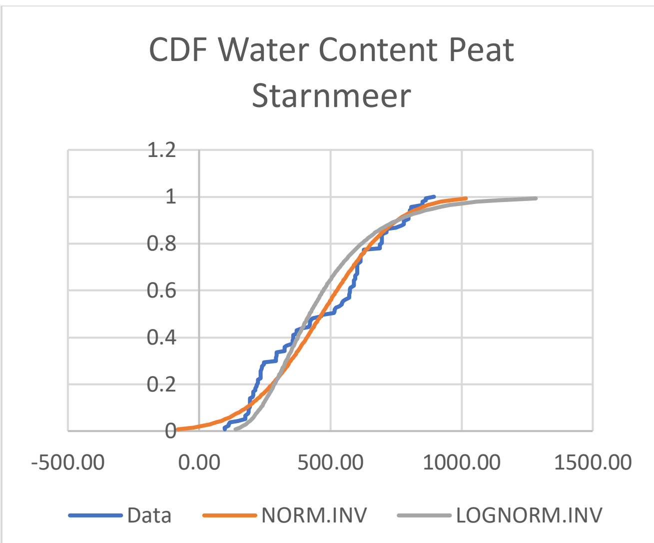
Peat:

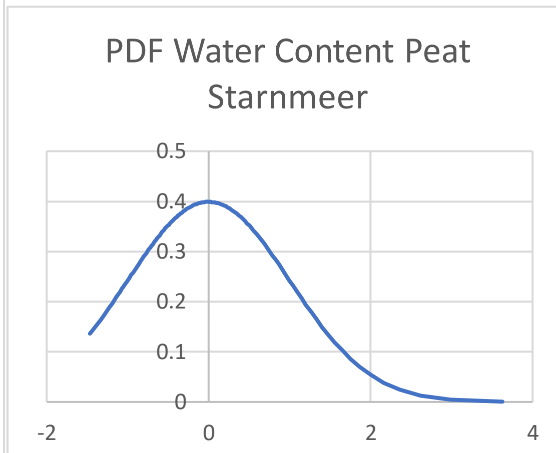

CDF ywet Peat -Starnmeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
5.00
10.00
15.00
Data
NORM.INV
LOGNORM.INV

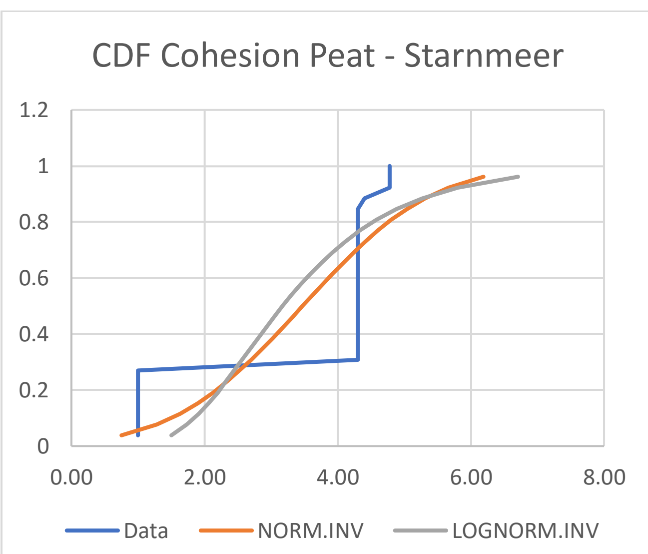

PDF Ywet Peat - Starnmeer
0.5
0.4
0.3
0.2
0.1
0
-4
-2
0
2
4

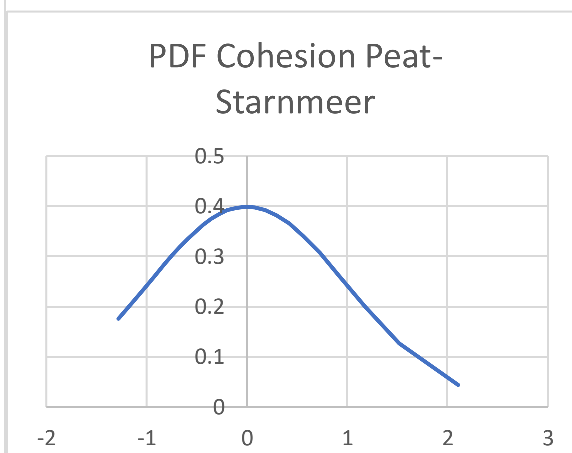

CDF ydry Peat- Starnmeer
1.2
1
0.8
0.6
0.4
0.2
0
-2.00
0.00
2.00
4.00
6.00
8.00
Data
NORM.INV
LOGNORM.INV

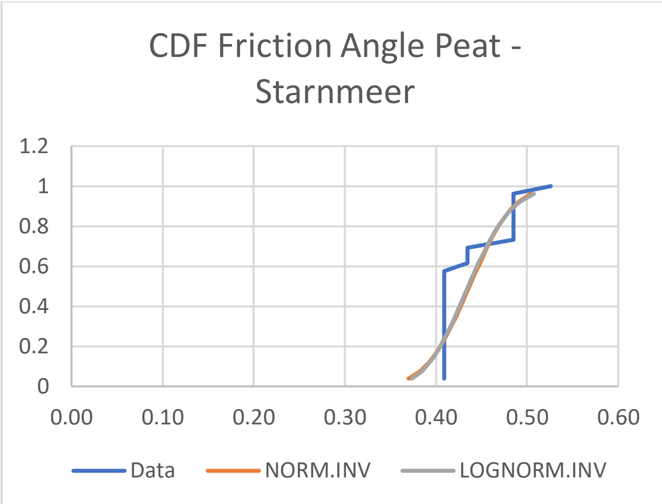

PDF ydry Peat- Starnmeer
0.5
0.4
0.3
0.2
0.1
0
-2
0
2
4
6

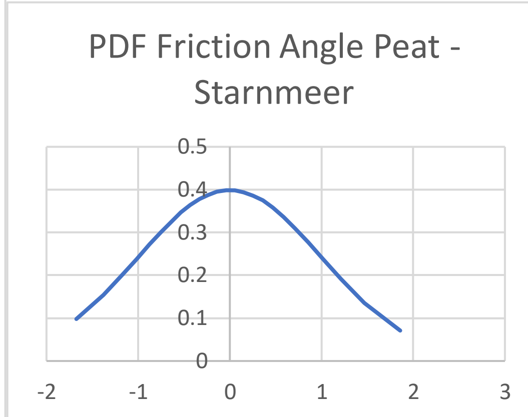

CDF Water Content Peat Starnmeer
1.2
1
0.8
0.6
0.4
0.2
0
-500.00
0.00
500.00
1000.00
1500.00
Data
NORM.INV
LOGNORM.INV


PDF Water Content Peat Starnmeer
0.5
0.4
0.3
0.2
0.1
0
-2
0
2
4


CDF Cohesion Peat - Starnmeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
2.00
4.00
6.00
8.00
Data
NORM.INV
LOGNORM.INV

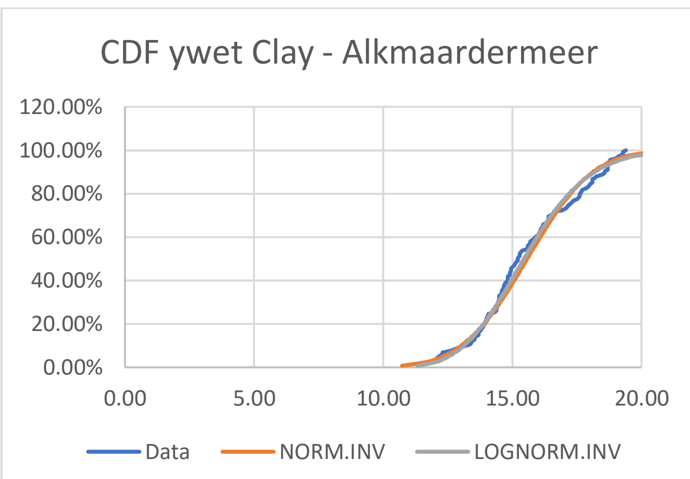

PDF Cohesion Peat- Starnmeer
0.5
0.4
0.3
0.2
0.1
0
-2
-1
0
1
2
3

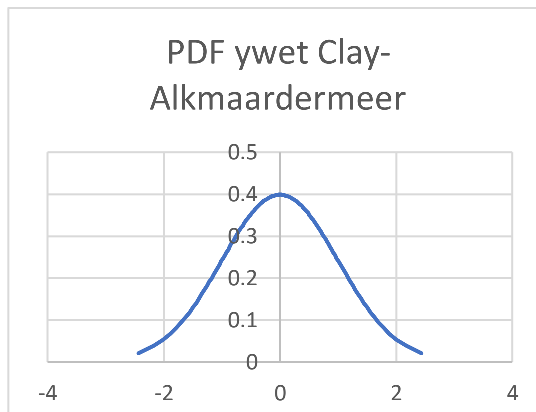


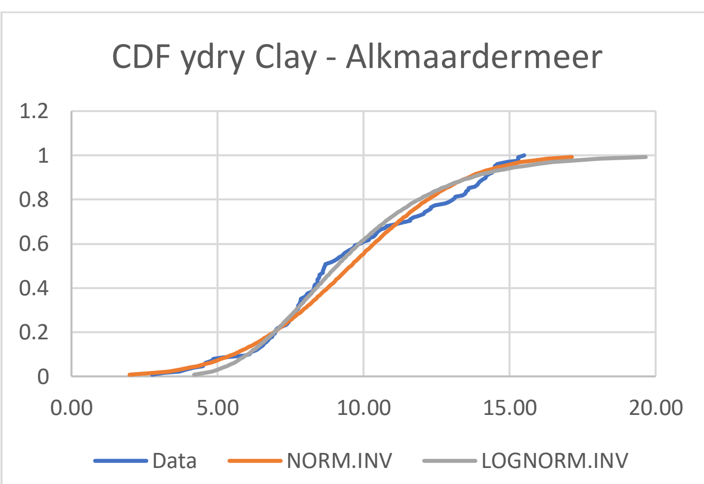


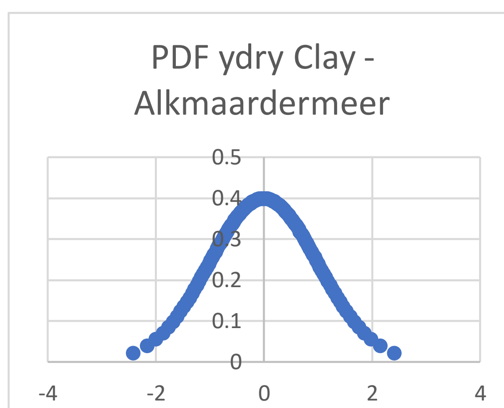
Alkmaardermeer:

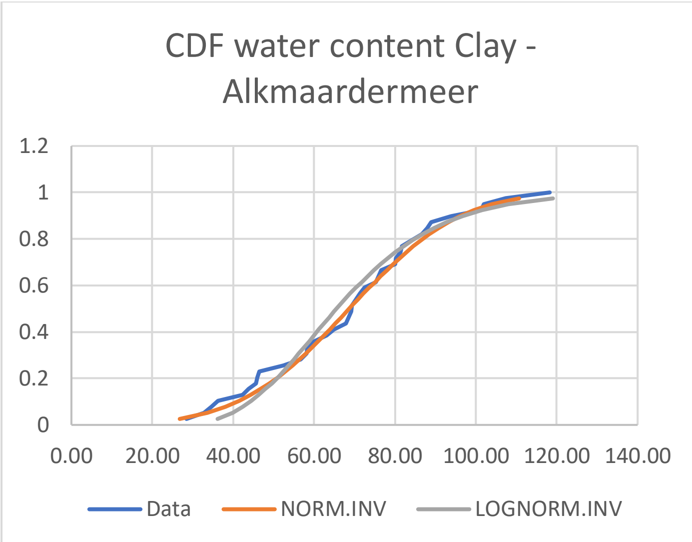
Clay:

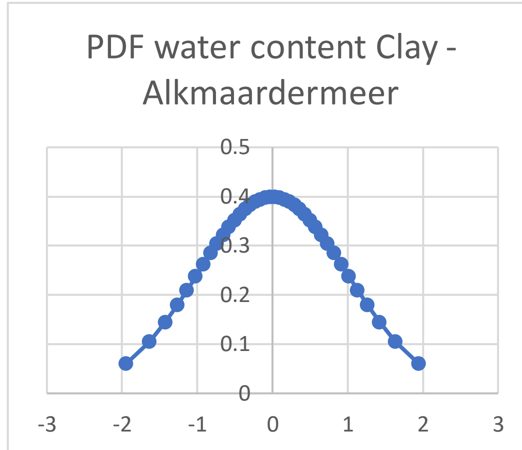


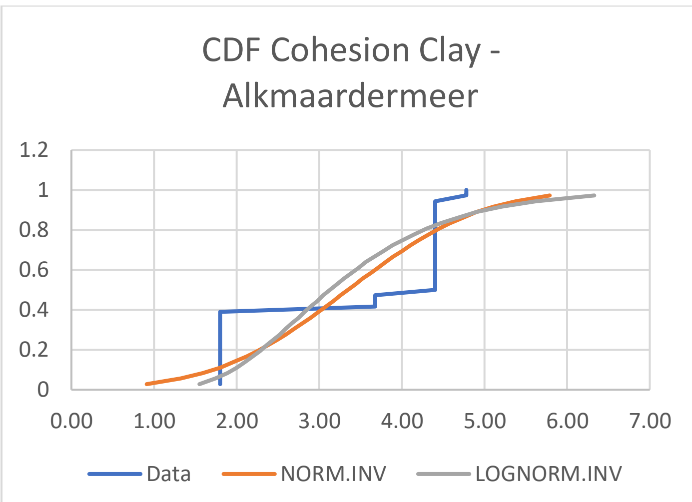


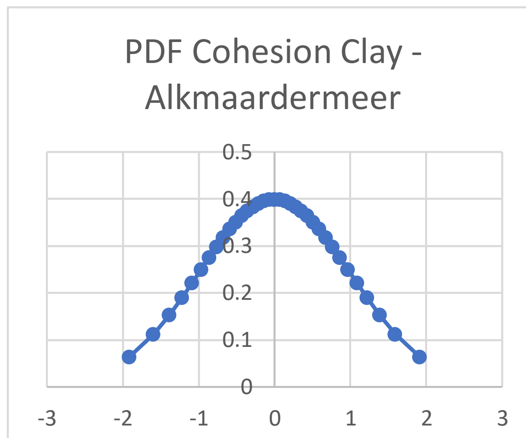


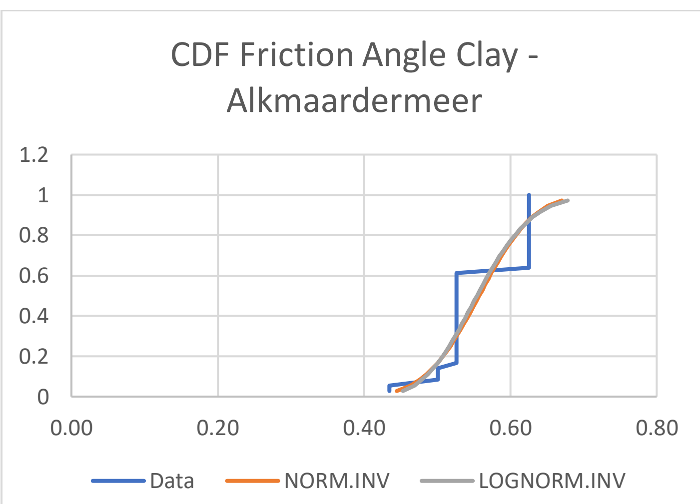


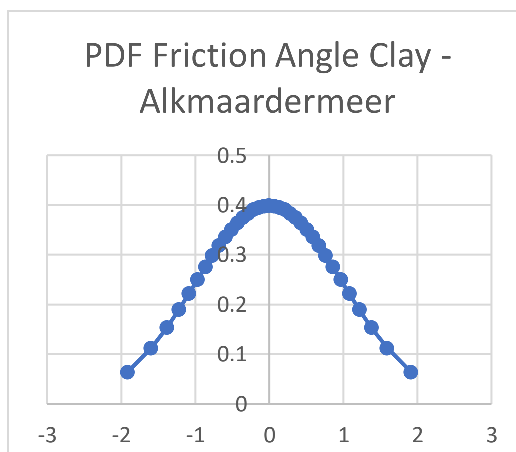

CDF water content Clay - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00 20.00 40.00 60.00 80.00 100.00 120.00 140.00
Data NORM.INV LOGNORM.INV


PDF water content Clay - Alkmaardermeer
0.5
0.4
0.3
0.2
0.1
0
-3 -2 -1 0 1 2 3

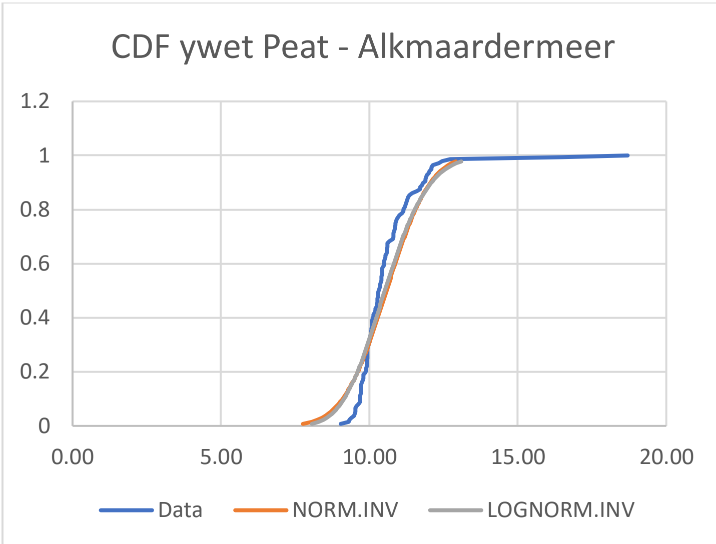

CDF Cohesion Clay - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00 1.00 2.00 3.00 4.00 5.00 6.00 7.00
Data NORM.INV LOGNORM.INV

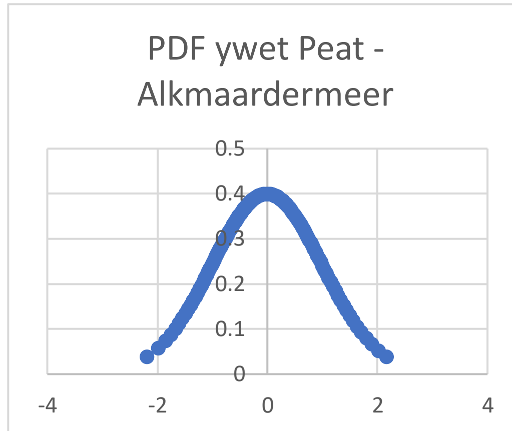

PDF Cohesion Clay - Alkmaardermeer
0.5
0.4
0.3
0.2
0.1
0
-3 -2 -1 0 1 2 3

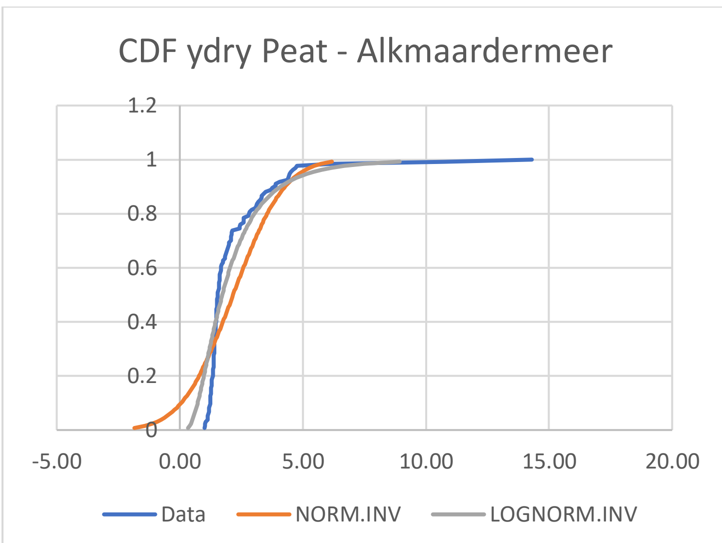

CDF Friction Angle Clay - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00 0.20 0.40 0.60 0.80
Data NORM.INV LOGNORM.INV

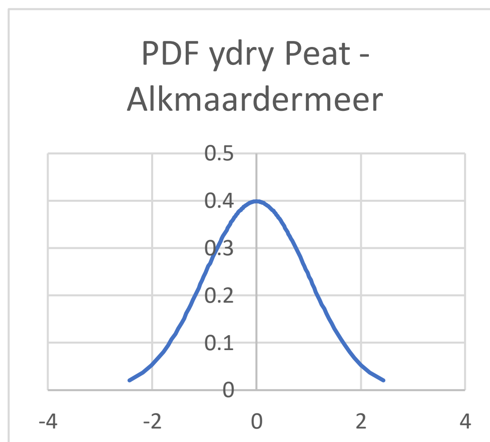

PDF Friction Angle Clay - Alkmaardermeer
0.5
0.4
0.3
0.2
0.1
0
-3 -2 -1 0 1 2 3

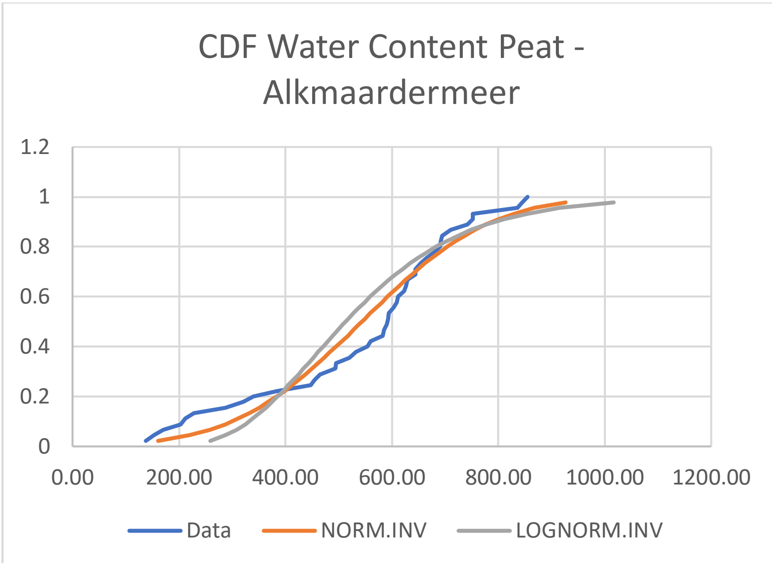
Peat:

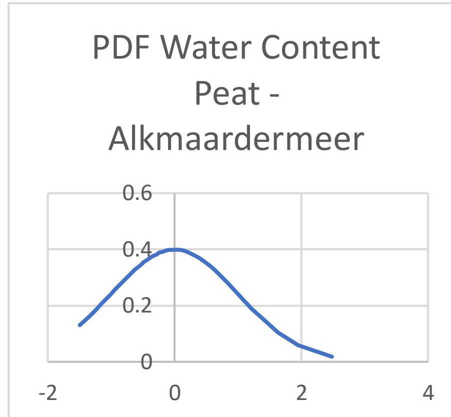

CDF ywet Peat - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
5.00
10.00
15.00
20.00
Data
NORM.INV
LOGNORM.INV

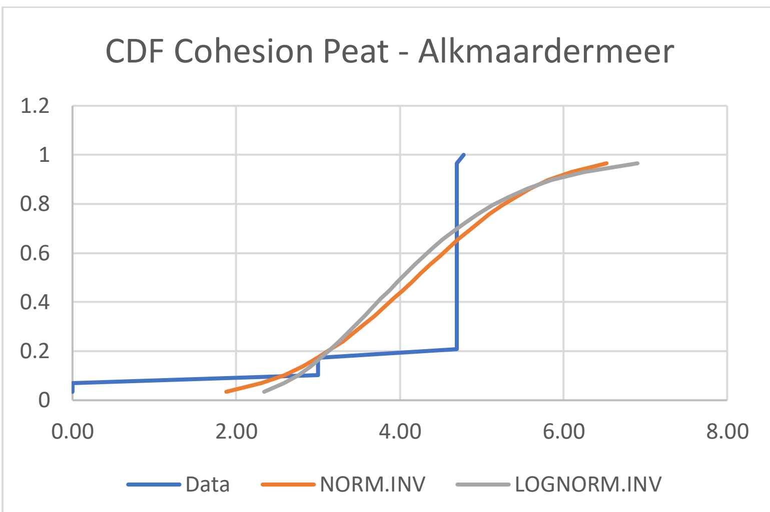

PDF ywet Peat - Alkmaardermeer
0.5
0.4
0.3
0.2
0.1
0
-4
-2
0
2
4

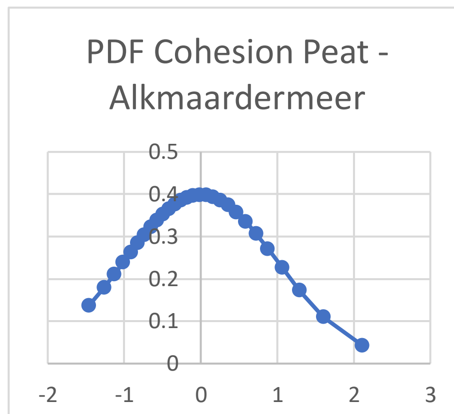

CDF ydry Peat - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
-5.00
0.00
5.00
10.00
15.00
20.00
Data
NORM.INV
LOGNORM.INV

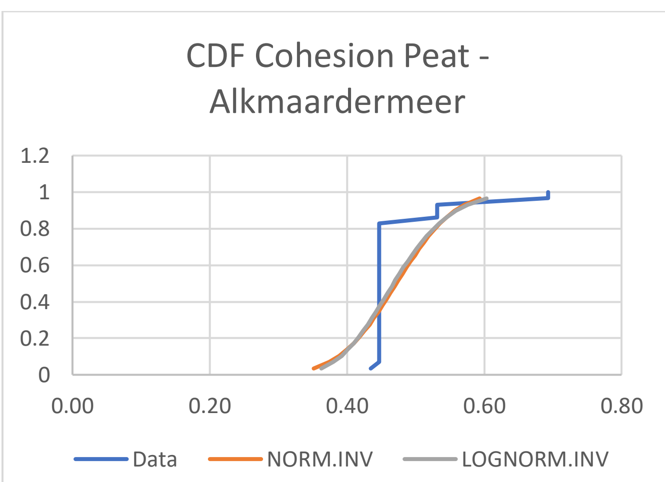

PDF ydry Peat - Alkmaardermeer
0.5
0.4
0.3
0.2
0.1
0
-4
-2
0
2
4

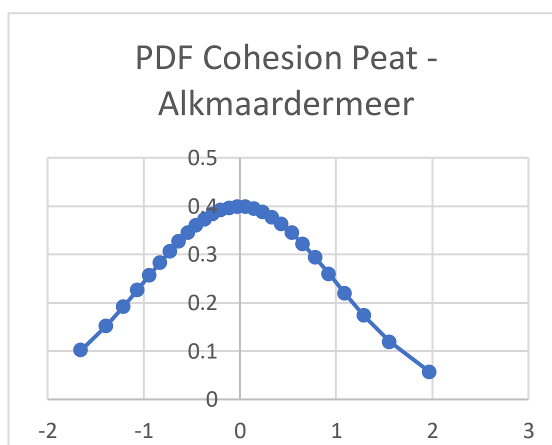

CDF Water Content Peat - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
200.00
400.00
600.00
800.00
1000.00
1200.00
Data
NORM.INV
LOGNORM.INV


PDF Water Content Peat - Alkmaardermeer
0.6
0.4
0.2
0
-2
0
2
4


CDF Cohesion Peat - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
2.00
4.00
6.00
8.00
Data
NORM.INV
LOGNORM.INV


PDF Cohesion Peat - Alkmaardermeer
0.5
0.4
0.3
0.2
0.1
0
-2
-1
0
1
2
3


CDF Cohesion Peat - Alkmaardermeer
1.2
1
0.8
0.6
0.4
0.2
0
0.00
0.20
0.40
0.60
0.80
Data
NORM.INV
LOGNORM.INV


PDF Cohesion Peat - Alkmaardermeer
0.5
0.4
0.3
0.2
0.1
0
-2
-1
0
1
2
3