Abstract - Geochemical patterns in the soils of Zeeland, natural variability versus anthropogenic impact

Chapter 1. Introduction

The close relation between soil, health, and global sustainability confirm that soil plays a critical role as a major interface in our environment and that soil quality can be an important indicator for sustainable environmental management.

This thesis is concerned with obtaining an overview of general (geo)chemical soil quality, within the framework of diffuse anthropogenic pollution and sustainable soil management in the Netherlands. It aims at assessing patterns in the geochemical soil composition and distinguishing the natural variability from the anthropogenic impact. The variability is assessed in geographical space, in which the spatial interaction of soil components takes place, and in attribute space, the space where interaction between soil constituents is exposed. Patterns from both spaces can then be related to processes influencing the soil composition.

Based on these patterns, and related processes, anthropogenic influence can be distinguished from natural variability. This can provide tools and information which can be part of the country wide overview needed as required by the Dutch government in their national environmental policy plan (NEPP 3).

The Province of Zeeland in the south west of the Netherlands was the area chosen for this research. The large rural area of young Holocene deposits and rich human history make it a suitable area for testing the main hypothesis that human influence leaves a distinct and quantifiable pattern of soil pollutants, based on variability within the geographical and attribute space of the soil composition.

Chapter 2. Geology and pedology of Zeeland

Zeeland is located in the marine delta of the rivers Rijn, Maas en Schelde. It was the sea level rise following the end of the Weichselian ice age that brought Zeeland within the
marine realm. The soils of Zeeland can be regarded as relatively homogeneous and young. They were moulded in a continuous interaction of natural processes and human endeavour. Considering the fact that soils of Zeeland are in general developed in marine clay deposits, it can be expected that a main source of geochemical variability is the varying clay content. The natural pattern of creek ridges and pool grounds creates relatively abrupt transitions. As local homogeneity or gradual variation may be further disturbed by extensive human activity, variability in soil composition is expected to be still higher than is directly evident from the soil map. Each (embanked) area with its own history, both sedimentary and human, may have its own pattern of soil variability. Given the fact that 80% of terrestrial Zeeland is used for agriculture, nearly 70% of which is arable land, human processes related to fertilisation and pesticide use are further expected to have influenced soil composition. This will result in elevated concentrations of so called “heavy metals” (Cd, Cu, Pb, and Zn) and persistent organochlorine pesticide residues. Finally, atmospheric inputs should also be considered as contributing to soil geochemistry.

Chapter 3. Natural and anthropogenic patterns of covariance and spatial variability of minor and trace elements in agricultural topsoil

The aim of this chapter was to analyze the sources of variation in soil geochemical composition related to spatial scale and sampling procedure. Apart from an intrinsic interest, research was motivated by practical questions regarding sampling strategies, using single or composite samples, for regional soil surveys. An unbalanced nested sampling design, based on single samples, was implemented at Walcheren/Zuid-Beveland. The samples were digested with aqua regia and analyzed by ICP-MS. The resulting data for 40 elements were examined through principal component analysis in combination with varimax rotation (PCA-V), unbalanced analysis of variance (UANOVA) and a fuzzy c-means clustering of the UANOVA results. The increase in noise when using single samples was further assessed through comparison with results of a previous study in the area based on composite samples. The three factors of the PCA-V are interpreted as being the results of natural, human, and geohydrological processes. The UANOVA analysis distinguishes three spatial variance patterns. Two groups show a gradual increase of variance with distance and contain respectively clay/feldspar related elements and heavy rare earth elements. The third group encompasses all anthropogenically and geohydrologically influenced elements with a larger part of the variance at small scales.

In general regional variance predominates over local variance and the extra effort of compositing is of limited value for a regional survey. However, environmental priority metals like Cd, Cu, Pb, Sb, Sn and Zn, were found to have larger local variance components. For these elements compositing is relatively more beneficial.
Chapter 4. Sampling and analyses for a regional environmental soil geochemical survey

When assessing soil composition for individual concentrations of elements it is often difficult to distinguish anthropogenic effects from natural variation. Establishing a bivariate or multivariate geochemical baseline, that incorporates a major part of natural variability, is an accepted method to derive background concentrations. Divergence from this background might then be caused by human disturbance. To quantitatively assess which components are affected by human activities, a consistent regional dataset is needed where special attention is paid to quantifying data quality. This chapter documents the construction of such a dataset for the province of Zeeland.

The sampling design was aimed at sampling present-day soil and pristine soil concentrations by sampling an influenced topsoil layer and an assumed pre-anthropogenic deeper subsoil layer at a density of 1 sample per 1-10 km². The analytical methods used (ICP-MS with aqua regia digestion and XRF on pressed powder tablets) and the broad group of elements measured (major elements like Al₂O₃ and Fe₂O₃, anthropogenically affected elements like Cd, Pb, Sb, trace elements like the REE) resulted in a data set of 87, partly overlapping, parameters. Each group of elements is thought to provide specific information on natural variation, human influence, or their interaction, as will the comparison between top and subsoil.

Precision was estimated by analysing about 15% of the ICP-MS samples and about 4% of the XRF samples in duplicate. The reproducibility is in general better for the XRF (0.5-10% relative to the mean) than for the ICP-MS (5-13% relative). The analytical variability is generally smaller than the local sampling variance as found in chapter 3, and therefore adequate for the desired regional evaluation of soil concentrations. The number of outliers found in the data also appears quite acceptable.

In this study the levelling of the data from each survey, based on Between Survey Duplicates (BSD), proved to be a good method to eliminate subsurvey biases. The fact that BSD levelling uses a set of samples that has the full range of element composition and concentration present, and is thus representative of the total survey, makes this method preferable to corrections based solely on analytical standards.

Based on the precision and accuracy a selection was made of 31 parameters that were measured in all surveys, and hence were sampled for virtually all locations and soil depths. This selection is thus optimised for multivariate statistical analysis by maximizing data quality and minimizing the number of missing values. This optimal dataset is limited compared to all the data analysed, but still comprises a sufficient subset of major, minor, and trace elements for environmental assessment of the soils of Zeeland. It is intended to form the core for all multivariate statistical and geochemical analysis, from which cross reference to the complete data set can still be made.

Summary statistics show that for most of the parameters neither a normal nor a lognormal distribution should be assumed. Although most major elements tend to a more or less
normal distribution, while most minor and trace elements are positively skewed, deviations from the standard type distributions are apparent. Suspected contaminants like Cd, Sb, As, Pb and Zn are more skewed than other trace elements.

Chapter 5. Enrichment and natural variability versus anthropogenic impact

Chapter 5 adopts a structured approach, using several statistical techniques that increase in complexity and the amount of auxiliary data needed. This approach provides detailed insight in the soil geochemistry of Zeeland. It first shows which components are most likely influenced by anthropogenic processes. Then, this anthropogenic imprint can be quantified based on their relation with non-influenced soil components.

Cumulative frequency plots are a good aid to show at a glance which elements have anomalous concentrations. From these plots percentiles can be derived thereby dividing anomalous data from background. The use of the value corresponding to the $P_{90}$, an often used standard percentile in Dutch soil surveys, as a baseline or threshold for assumed to be influenced elements might be a responsible choice, but only if the corresponding cumulative distribution plots are studied. In this study most (trace) elements are chiefly associated with natural processes and the $P_{90}$ is an obvious underestimation, the $P_{99}$ is more appropriate. Also for Pb the $P_{90}$ is an underestimation, while for some other probably anthropogenically influenced elements (Cd, Cu and Sb) the $P_{90}$ is an overestimation of the threshold value. This suggests that the arbitrary choice for the $P_{90}$ in many Dutch soil surveys should be reconsidered. Notwithstanding the advantage of its simplicity, the method only shows if a components distribution deviates from normality with a distinct tail towards higher values. It does not show the nature of the deviation, anthropogenic or natural, nor does it provide a clear overview of the extent of the enrichments, due to the subjectivity and sometimes ambiguity in the choice of the threshold percentile. However, at a reconnaissance stage the method can be very useful.

The spatial distribution of elements with low threshold percentiles ($P_{80}$ to $P_{90}$) and suspected of human influence, shows regional variability that differs distinctly from the more or less similar patterns that emerge for elements related to natural processes like variability in clay/Al$_2$O$_3$. This indicates that these probably influenced elements indeed have a single process underlying their regional variability.

Calculation of median ratios, with the upper continental crust (UCC), median subsoil concentrations, and legal limits, is an easy method to determine which elements are enriched compared to the reference used. Ratios with the UCC indicate that As, Cd, Cr, Pb, and Zr are relatively enriched. For As it is assumed that this enrichment is partially caused by the low lying deltaic environment of Zeeland, Cr and Zr could be related to heavy mineral concentrations, while the other elements might be enriched due to fertiliser usage and/or traffic exhaust. Depleted elements are Ba, Ca, K, Mg, Mn, Na, and Sr, which is assumed to be a natural regional feature. On average, the Zeeland soils also appear somewhat less
aluminous and ferromagnesious than the UCC. The ratios based on the subsoil show that
the elements As, Cd, Cu, P, Pb, S, Sb, Sn, and Zn are enriched. These are all elements for
which enrichments can usually be related to anthropogenic activities. The ratios with the
official legal limits for Dutch soil, often used as reference values, show ratios of about 0.5-
0.8, with the higher ratios (Cr, Ni) probably caused by the better analytical performance of
the XRF method used as compared to the prescribed aqua-regia extraction. Even the $P_{99}$ of
the ratios, based on median clay values, are still below one for all elements.

The implications of the ratio results are basically twofold: 1) the area of Zeeland covered
by this study is clean as regarding the governmental regulations and 2) the legal threshold
limits are an overestimation of soil background concentrations, already including (minor)
anthropogenic enrichments. They are thus less suitable as a reference value for a study into
natural and anthropogenic geochemical patterns, and the ratios with the local reference
(subsoil data) are preferred for identifying the influenced elements and determining the
extent of their enrichment.

The correlation with $\text{Al}_2\text{O}_3$ is easily calculated and proved to be very useful to indicate
elements which are possibly altered by human processes. The elements $\text{P}_2\text{O}_5$, Cd, Cu,
Pb, S, Sb, and Sn show relatively low correlations with $\text{Al}_2\text{O}_3$ in the topsoil compared
to other elements and are also enriched in the topsoil. CaO, MnO, Na$_2$O, and Sr show
relatively low, or negative, correlations with $\text{Al}_2\text{O}_3$ but are not enriched in the topsoil. The
correlations confirm the similarities observed in the spatial distribution of the elements.

Overall higher correlations with $\text{Al}_2\text{O}_3$ in the reference subsoil confirm the hypothesis
that the subsoil can be regarded as a pristine, non-influenced, reference. A graphical
representation of the relation between the various elements and $\text{Al}_2\text{O}_3$ for both topsoil
and subsoil clearly depicts the relative addition in the topsoil. Using a linear regres-
sion model based on the subsoil data the extent of the enrichment was quantified for
each topsoil sample. When expressed as the ratio between observed concentration and
calculated baseline concentration, the elements in order of decreasing enrichment are
Cd $>$ Cu $\gg$ Sn $\approx$ Sb $>$ Pb $>$ Zn $>$ As. These enrichments are in line with data found
for a similar area in the north of France, except for Cu, which is relatively higher in the
Zeeland topsoils. Spatial distribution patterns of the enrichment ratios show that for most
influenced elements regional variability is high and that hardly any spatial or mutual re-
gional relations appear. The overall regional perspective provided by the contamination
index shows that some areas in the north-east and centre of Zeeland are slightly more en-
riched on average.

Using the subsoil regression as a baseline has many advantages above the current Dutch
practice of calculating a fixed legal threshold value, though corrected for clay and organic
matter content. The natural random variability in baseline concentrations, expressed as the
standard error of the regression, is a property of the model. This contrasts with the legal
threshold value, which is just a rigid value with no range given. Also the regression model
provides a local baseline, specifically suited for the parent soil material of Zeeland, which
obviously should be preferred over a general model that tries to include all possible soil
types. Compared to all other techniques presented, this subsoil regression baseline method
gives the most profound insight into 1) the geological/pedological pattern of geochemical
variability, 2) random natural variation, and 3) the anthropogenic imprint on the Zeeland soil. These advantages definitely outweigh the greater complexity of the procedure.

**Chapter 6. Regional diffuse geochemical patterns and processes**

In this chapter I focus on the extent to which the human contributions actually affect the regional soil composition compared to the natural variability. Using the same data as in chapter 5, the processes or factors determining the regional variability, both in attribute and geographical space are assessed. The regional expression of the interaction of human and anthropogenic processes is further evaluated, to see if regions may be distinguished where specific human processes have had an effect on the soil composition.

The general covariance structure of the data was studied using robust principal component analysis (PCA) with varimax rotation to reveal the processes within attribute space. Groups of covarying geochemical components, may indicate certain processes. To assess which processes explain features in geographical space, clustering techniques were used to identify groups of sample points with similar properties. Due to the expectation that each pattern is a combination of different processes and the general multivariate nature of geochemical data, fuzzy c-means clustering (FCMC) was chosen.

The consecutive use of principal component analysis and fuzzy c-means clustering provided a profound insight into the processes affecting the regional geochemical variability in the Zeeland soils and into the extent of the variability caused by human processes.

The PCA reveals 4 factors that explain 76.6% of the overall variance, these factors are named for convenience “clay”, “calcite”, “anthropogenic” and “silt” respectively. The largest part of the variability (68%) is akin to geogenic clay, calcite, and silt related processes while 8.9% can be explained by anthropogenic processes. The elements related to the latter were the so called “heavy metals” Cd, Cu, Sb, Sn, and to a lesser extent Pb, Zn, and As. The regional variability of the clay, calcite, and silt factor resembles the regional variability of elements associated with this factor. The regional variability of the anthropogenic factor can be largely compared with the contamination index as obtained in chapter 5. When the elements of the anthropogenic factor were examined more closely, no separate subprocesses could be discerned other than unique behaviour for each element.

Considering that the soils of Zeeland are clearly enriched for some elements it is to be expected that this enrichment will also be revealed when looking at the processes within attribute space. The PCA confirmed this, but, the natural clay and calcite related processes still comprise the largest part of the soil variability. The anthropogenic influence on the soil only explains a small part of the variability emphasising that these anthropogenic processes only affect a small number of inorganic elements and are really an imprint on the overall natural variability.

The grouping of locations with FCMC, using the elements with the highest loadings of the PCA, resulted in an optimal 4 cluster model mainly dividing the area in calcareous heavy
clays, non-calcareous heavy clays, moderately calcareous silty clays, and non calcareous silts. No clusters exist where samples are discriminated solely on the basis of anthropogenic enrichment. This suggests that the extent of these human processes, clearly imprinted on the Zeeland soils, are not related to certain regions but rather are diffusely distributed over the province.

Chapter 7. Assessment of regional DDT concentrations in the soils of Zeeland

This chapter aims at providing a regional assessment of current DDT residue concentrations, variability, and degradation for the province of Zeeland. Based on a merging of several separate soil information systems (SIS) data, and comparing the concentrations with legal limits and data from other areas allowed the extent of contamination to be assessed, providing insight into whether the concentrations are just part of the ubiquitous background or exceed them. The geographical distribution of the DDT residue levels might indicate patterns, if any, that possibly relate to distinct processes. In addition, the data integration exercise may help assess the value of the combined SIS for regional environmental policy and may lead to recommendations for the (local) authorities for further research and monitoring.

The combination of the different SIS derived datasets resulted in a database that can be used for a regional assessment. However, the differences between the surveys, although in general based on the same sampling and analytical standards, are evident. It is hard to determine if these differences denote differences in true DDT soil values or that they are largely caused by the slight variations in surveying and laboratory methods.

The \( \Sigma \text{DDx} \) concentrations (sum of the concentrations of all DDT isomers and metabolites) as found in Zeeland are comparable to those in similar agricultural areas, and based on these values there is no reason to believe that the current level of DDT contamination is more severe than can be expected for an agricultural area. Compared to pristine areas the level in Zeeland is about a factor 2 to 10 higher, which indicates that Zeeland can be regarded as enriched compared to a general ubiquitous \( \Sigma \text{DDx} \) background. Despite that the values are within a normal range for agricultural areas, the standard soil normative threshold value \( S \) is exceeded for roughly half the cases and the soil specific value \( s \) in even more cases, which may result in litigations.

The relatively large contribution of local scale variability to the overall variability of \( \Sigma \text{DDx} \) concentrations implies that a regional overview has limited value since most of the observed variance can be regarded as noise caused by small scale features. This is both indicated by the analysis of variance on replicate samples and by the nugget effect of the calculated semivariance.

Given both small scale variability and between survey bias, a final regional overview will contain large uncertainties. Despite that large scale features appear visible, like the relatively higher concentrations in the south of Zuid-Beveland which are presumably caused by
the presence of orchards, these should be judged with caution. Extreme caution is needed when the values are compared with (soil specific) legal limits. Due to the uncertainty caused by the various sources of variability, such comparison might have very limited value.

The degradation level of DDT residues indicates that the Zeeland soils will remain a source for \( \Sigma \text{DDx} \) longer than expected. Since it is not assumed that new applications of DDT will occur, the monitoring of DDT residue levels could be cautionary. The most important issue which should be addressed is the reduction of variability in the monitoring data. Probably the largest reduction can be obtained in more standardisation of surveying, reducing the between survey bias. The small scale variance may be assessed by increasing the number of subsamples and sampling on smaller scales but this might be less cost-effective than reducing the between survey bias. However, policy based on these data should always incorporate the uncertainties as a result of the remaining variability.

**Chapter 8. Synthesis**

This study showed that with the use of state of the art analytical techniques and well established methods from geochemistry, much insight can be obtained from a single survey into the natural soil composition and how it is influenced by human processes. This both provides methods to obtain a reference for natural soil composition and to quantify the anthropogenic imprint. It is shown in this study that the framework of legal limits neither provides such reference nor is it able to detect subtle anthropogenic enrichment. This, in a way, is remarkable considering that this framework is also used for the building material decree and related to soil pollution risk maps that should prevent spreading of contaminants. I therefore suggest that references for soil composition should be derived from (local) baselines rather then legal limits. When obtaining such a baseline, quantification of the variability due to sampling and analytical procedures, and methods to reduce or compensate bias between sub-surveys are as important as the baseline itself. If multi-element techniques are used, like most laboratories do, focus should not only be on environmental priority components but also on components which might provide information about natural geochemical soil composition and variability.

Soil pollution risks in Zeeland appear to be, as defined in the Dutch normative framework, mainly determined by DDT residues. Unfortunately, the low average concentrations in soil, complex analytical determination, and large small-scale variability result in large uncertainties towards the regional variability of this component. Consequently, to obtain a more trustworthy georeferenced overview of DDT residue levels, more effort aimed at reduction of variance and bias should be put in the assessment of this component.

The question remains what “soil quality” should mean for the Dutch situation and which indicators to use. In this chapter it is argued that one of the goals to reach sustainable soil management should be a reconsideration of what is understood by “soil quality”. Only then it is possible to define which soil quality indicators should be assessed. The reference data and methods provided in this study could be a starting point for obtaining such indicators, finally leading towards sustainable soil management.