Afbakening van de fosfaatverzadigde gebieden in Vlaanderen op basis van een kritische fosfaatverzadigingsgraad van 35%

(VLM/MESTBANK/TWOL2006/MB2006/5)

Finaal rapport - deel 1a, afbakening

Prof. Dr. ir. Marc Van Meirvenne, Meklit Tariku & Joost Salomez
Onderzoeksgroep ruimtelijke bodeminformatietechnieken (ORBIT)
Vakgroep Bodembeheer en bodemhygiëne

30 juni 2007

Table of content

List of Figures

List of Tables

Summary

The delineation of phosphate saturated areas in Flanders was first conducted in 1997 using block ordinary kriging. For every location to be estimated block ordinary kriging resulted in estimation variances and estimated phosphate saturation degree (PSD) values. The 95% estimation uncertainty was then determined as a difference between the estimated PSD values and twice the standard deviation. When this difference was equal or greater than 40%, that location was determined as phosphate saturated. In 2007, due to the need to change the threshold value from 40 to 35% it was found necessary to update the demarcation of the phosphate saturation areas. This geostatistical analysis was done with sequential Gaussian simulation (SGS) and used additional 708 observations. Simulations were preferred over kriging estimations because of the possibility to obtain more realistic quantification of the local uncertainty through the generation of multiple equi-probable values per location. Both the estimated value and the estimation uncertainty can then be obtained from this set of values as the mean estimated value and the probability of exceeding the defined threshold.

The frequency distributions of the 3644 initial observations, referred to as the "old data set", and the 4352 data, referred to as the "complete data set", were similar. The analysis was conducted for both data sets. After processing the simulated values, the old data set resulted in an area of 54 km^2 where the PSD values exceeded 35% with a 95% certainty. However, most of these locations appeared as scattered individual pixels. Using the complete data set a slightly larger area was obtained: 64 km^2, but these formed more contiguous patches. This result indicated the importance of considering the additional data in the analysis.

During the analysis of 1997, an area of 73 km^2 was defined with 95% certainty to have a PSD of more than 40%. In the current analysis, using the threshold of 35% PSD, 64 km^2 was delineated at 95% certainty as phosphate saturated. These seemingly contradicting result is the consequence of using a different technique with an extended dataset. To breach the gap it was suggested to consider different alternatives. The first alternative was to analyze the complete dataset with SGS but consider as PSD threshold the values of 25% or 30%, instead of 35%. At 95% certainty these alternatives resulted in the delineation of 1812 km^2 and 460 km^2, respectively. The second alternative was to change the level of certainty of estimation from 95% to 80%. For the thresholds of 25%, 30% and 35% PSD this alternative resulted in the delineation of 3827 km^2, 1928 km^2 and 578 km^2, respectively. The third alternative was to reanalyze both the old and the complete datasets using block ordinary kriging to define the 95% probability of exceeding the 35% PSD. Following this alternative, 197 km^2 and 157 km^2 were determined using the old data set and the complete data set, respectively.

Of all approaches, maps were produced.

Samenvatting

De afbakening van de fosfaatverzadigde gebieden in Vlaanderen in 1997 was gebaseerd op blok ordinaire kriging. Hierbij werden locaties waarvan met 95% zekerheid geschat werd dat de fosfaatverzadigingsgraad (FVG) 40 % bereikte verondersteld fosfaatverzadigd te zijn. Voor elke locatie geeft block ordinary kriging niet alleen een geschatte waarde maar ook een schatting van de standaard afwijking. De 95% schattingszekerheid werd dan ook berekend als het verschil tussen de schatting en tweemaal de standaard afwijking op deze schatting. Wanneer het resultaat van deze berekening groter dan 40% was dan werd de locatie ingekleurd als fosfaatverzadigd. In 2007, door de verlaging van de drempelwaarde van 40 tot 35% werd het belangrijk opnieuw de afbakening van de fosfaatverzadigde gebieden te doen. Dit keer werd voorgesteld om blok sequentiële Gaussiaanse simulatie (SGS) te gebruiken met daarenboven 708 bijkomende gegevens. De simulatietechniek is aan te raden omdat in vergelijking met de ordinaire kriging de resultaten meer realistisch zijn. Met SGS wordt voor elke locatie een aantal even waarschijnlijke schattingen berekend waardoor de gemiddelde schattingen en de schattingsonzekerheid kunnen bepaald worden.

De 3644 gegevens die gebruikt werden bij de analyse in 1997, hierna ”de oude data” genoemd, en de 4352 gegevens, “de complete data”, hebben een vergelijkbare frequentieverdeling. Maar de bijkomende 708 datapunten introduceren een duidelijk hogere lokale, kleinschalige variabiliteit in gebieden waar een hogere FVG gemeten werd. Daarom werd voorgesteld om de analyse voor zowel de oude als voor de complete dataset te doen. Op basis van de oude data en met 95% zekerheid bakent de blok SGS procedure slechts 54 km^2 af als fosfaatverzadigd met een verzadigingsgraad boven 35%. De meeste van deze gebieden zijn echter individuele pixels. Met de volledige dataset vergroot deze fosfaatverzadigde oppervlakte tot 64 km^2. Belangrijk hierbij is op te merken dat de betere continuïteit van de resulterende kaart, een betere schatting aangeeft. Dit verschil is het resultaat van de lagere schattingsonzekerheid door het gebruik van de bijkomende gegevens.

In 1997 werd 73 km^2 oppervlakte met een FVG groter dan 40% afgebakend door de analyse van de oude data met de blok ordinaire kriging bij 95% zekerheid. Deze analyse werd opnieuw uitgevoerd, maar waarbij nu slechts 64 km^2 oppervlakte werd ingekleurd als fosfaatverzadigd (max. 35% FVG). Het verschil tussen de 1997 en de huidige resultaten kan verklaard worden door de verschillen tussen de gebruikte methoden en de uitgebreide dataset. De oppervlakte van de resultaten van de analyse in 1997 waarbij de blok ordinaire kriging en de 40% FVG gebruikt werden, is kleiner dan de verzadigde oppervlakte die bepaald wordt door het gebruik van SGS met een 35% FVG. Vandaar dat verschillende alternatieven werden onderzocht. Het eerste alternatief was om met SGS en de complete data de 35% FVG drempelwaarde te wijzigen tot 25% of 30%. Met 95% zekerheid leverde dit een oppervlak van 1812 km^2

en 460 km^2 op met een FVG van boven 25% en 30%, respectievelijk. Het tweede alternatief was om 80% te gebruiken als zekerheidsdrempel in plaats van 95%. Voor een FVG van 25%, 30% en 35% resulteert dit alternatief respectievelijk in een oppervlakte van 3827 km², 1928 km² en 578 km². Het derde alternatief bestond uit het analyseren van de oude en totale data set met blok kriging om fosfaatverzadigde gebieden aan te duiden gebaseerd op een FVG van 35% en met een 95% zekerheid. Dit alternatief bakende respectievelijk 197 km² en 157 km² fosfaatverzadigde gebieden af voor respectievelijk de oude en de volledige dataset.
Van alle benaderingen werden kaarten aangeleverd.

1 Introduction

In Flanders in 1997 the delineation of areas with a Phosphate Saturation Degree (PSD) of 40% was conducted by means of block kriging. Two times the kriging standard deviation was subtracted from the estimations to account for the estimation uncertainties. Only the areas where this result exceeded 40 % were identified as "areas with a 95 % probability to exceed a PSD of 40 %". The New Manure Decree of 2006 imposed a shift of the limit to 35% PSD. For the demarcation of areas with a PSD level of 35% or more the data analysis had to be repeated, including additional measurements and updating the geostatistical technique used.

2 Concepts and theory about Phosphate Saturated Soils (after Lookman, 1995)

2.1. Basic assumptions

In order to prevent risks of excessive phosphorus (P) leaching from agricultural soils, Van der Zee et al. (1990a, b) developed a protocol to test whether a soil is P saturated or not, on a routine lab-procedure basis. Both P and the phosphate sorption capacity (PSC) are determined with a single oxalate extraction. P, iron (Fe) and aluminium (Al) are determined in the extract and the PSC and PSD are calculated as (M represents the sum of oxalate extractable iron + aluminium in mmol kg^{-1}):

$$PSC = \alpha_m \times M \text{ (mmol P } kg^{-1}) \quad 2.1$$

$$PSD = P_{ox}/PSC \times 100 \text{ (\%)} \quad 2.2$$

with the subscript 'ox' denoting 'oxalate extractable'. The parameter α_m represents the fraction of M that can bind P since not all of M is available to fix P. The value of $\alpha_m = 0.5 \pm 0.1$ was withheld for general use in the protocol.

It is likely that for soils with pH-H_2O above 6.5-7 calcium (Ca) plays an important role in the P sorption processes as well and that accordingly equation 2.1 would yield an underestimation of the PSC.

2.2. Criterion P-saturated soil

To establish the value of the PSD, starting from which a soil is classified 'P-saturated' ($PSD_{critical}$), the following equation is put forward:

$$PSD_{critical} = \frac{K\chi\varepsilon c}{1 + K\chi\varepsilon c} \times 100 \ (\%)$$

with K the Langmuir adsorption constant, χ the ratio adsorbed/precipitated phosphate (= 0.3), ε the ratio of the PSC of the subsoil with respect to the topsoil (=1) and c the P concentration allowed in the solution leached form the field. Taking K = 10.6 and c = 0.1 mg ortho-P L^{-1}, the value of $PSD_{critical}$ = 25% can be calculated.

For a more in-depth discussion on phosphate saturated soils, we refer to part 1b of this onderzoeksrapport.

3 Geostatistical analysis

Geostatistics offers two ways to predict the value of a variable at unsampled locations: **estimation** and **simulation** (Goovaerts, 2000). Both techniques are based on the procedure of kriging interpolation which provides the best linear unbiased prediction of a value in terms of minimum error variance (Chilés and Delfiner, 1999). The difference is that estimation aims at obtaining a single “best” predicted value for every location where every estimate is obtained independently of the neighbouring estimates. Simulation on the other hand focuses on the reproduction of the spatial variability by drawing multiple, equally probable realizations from a random function whereby a simulated value for every location is conditioned by not only the neighboring available data but also with previously simulated values. So simulated maps show a stronger similarity between the histogram and variogram of the sample data. As a result, a map obtained through simulation is more ‘realistic’ than the result of an estimation, at the expense of not being the “best prediction at individual locations”. To compensate for this limitation, simulated maps are repeated many times, similar to the Monte Carlo technique used in process simulations and error propagation studies. The set of simulated maps allows to derive e.g. the predicted mean value, but also to construct the distribution of expected values at every location. The latter offers the possibility of obtaining extra information, not obtainable from geostatistical estimation.

One of the major drawbacks of geostatistical estimation is the non-uniform smoothening of the local variability (Goovaerts, 1999). There is overestimation of lower concentrations and underestimation of higher concentrations depending on the density of data. The smoothening obstructs understanding of the variability and detecting patterns of extreme values, such as strongly polluted zones. It might even result in artificial patterns in cases of uneven distributions of data. In Geostatistical estimation the assessment of the prediction uncertainty is obtained from the kriging variance, which is a parametric measure of the precision of interpolation in terms of the global variance. The kriging variance is not suitable as a local measure of predictive precision; it mainly reflects the sampling configuration in terms of the global variogram.

Geostatistical simulation allows the assessment of the uncertainty of prediction locally and jointly over multiple locations from a multiple of realizations that are conditioned by locally available information. This local measure of prediction uncertainty has a significant importance in improving decision making when the uncertainty is non-parametric because then it will not depend on the assumption of normality of prediction error (Van Meirvenne and Goovaerts, 2001).

A sequential geostatistical simulation technique was used for the analysis of the PSD data since this technique provides a better measure of the local estimation uncertainty than kriging estimations (Goovaerts, 2000). The method that was used in this study was Sequential Gaussian Simulation (SGS). This method has been used for the spatial variability distribution and uncertainty assessment of soil phosphorus in a south Florida wetland (Grunwald et al., 2004). It is also one of the standard methods for delineation of soil pollution (Yu-pin Lin et al., 2001) and analysis of soil properties such as soil strength/compaction (Lapen et al., 2001).

The SGS algorithm follows a number of consecutive steps as shown in Figure 1 as a flow chart:

1. The original n data (α = 1,..,n) are transformed to normal scores according to $y(\mathbf{x}_\alpha) = \phi(z(\mathbf{x}_\alpha))$, with $z(\mathbf{x}_\alpha)$ the original data, $\phi(.)$ a transformation function and $y(\mathbf{x}_\alpha)$ the normal scores with a Gaussian histogram with a mean of zero and a variance of one (Figure 2).

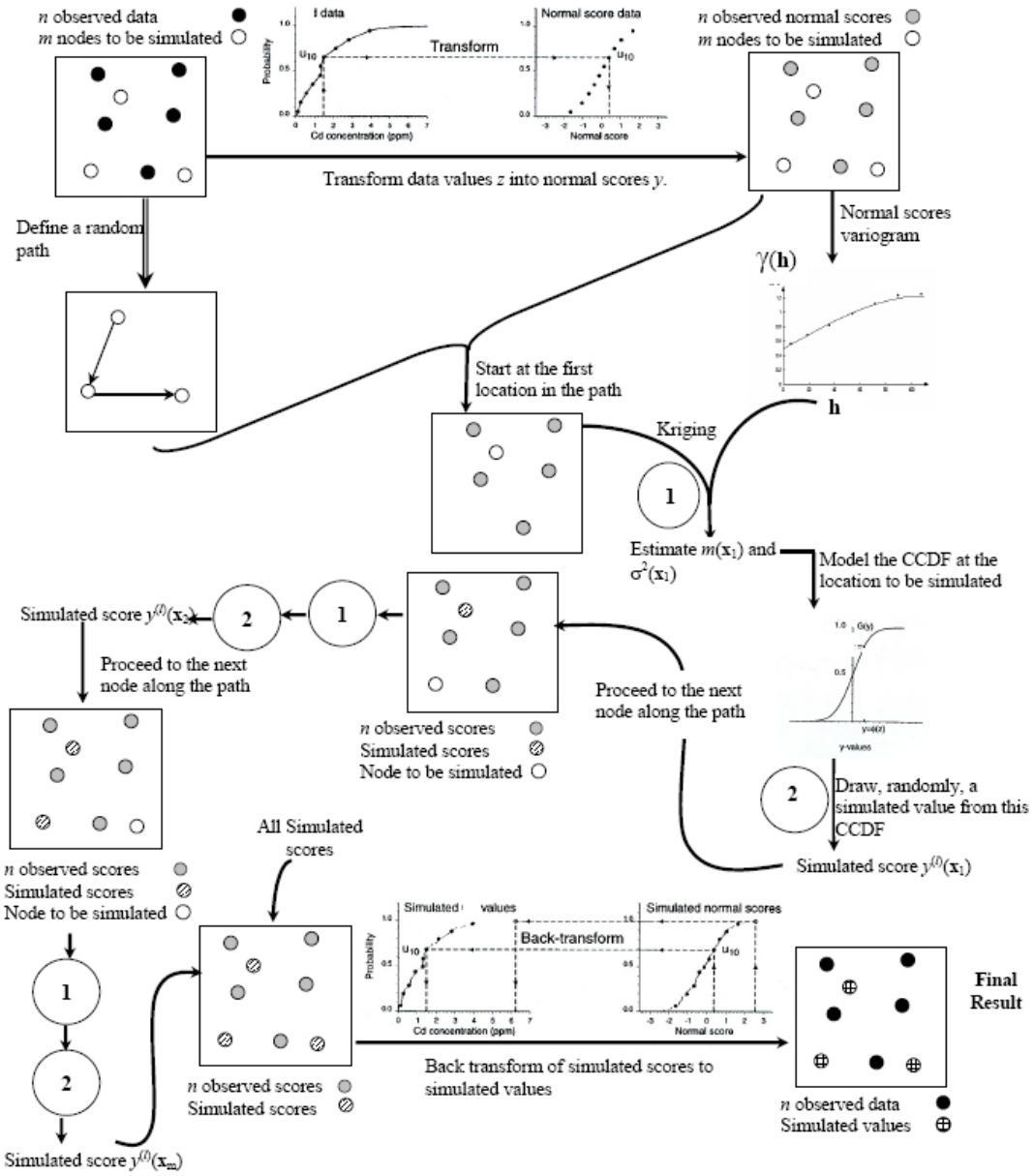

Figure 1. Flow path of Sequential Gaussian Simulation (SGS), (From Fagroud & Van Meirvenne, 2002)

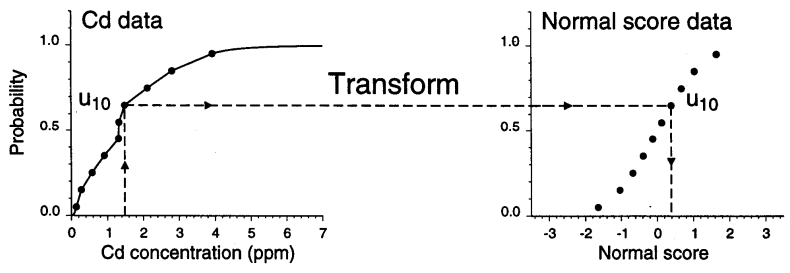

Figure 2. Example of transformation of original z-values (here Cd concentrations) into normal score values.

For the transformation first the cumulative frequency of the original data is made by setting the data in increasing order. In case there is a preferential sampling in the data, declustering weights can be used at this stage. The following step is matching the corresponding percentiles of this cumulative frequency to a standard normal distribution with mean zero and variance (Figure 2).

2. The variogram of the normal scores is computed using the formula:

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} \{y(\mathbf{x}_\alpha) - y(\mathbf{x}_\alpha + \mathbf{h})\}^2$$

Where: $\gamma(\mathbf{h})$ is the variogram

$N(\mathbf{h})$ is the number of couples

$y(\mathbf{x}_\alpha)$ and $y(\mathbf{x}_\alpha+\mathbf{h})$ are normal scores separated by a distance vector $\mathbf{h}$

3. L simulations are performed using the normal scores in the following procedure:
   - Define a random path whereby each of the m unsampled locations are simulated once;
   - At each unsampled locations $\mathbf{x}_0$ estimate the parameters, mean and variance, of the conditional cumulative distribution function (ccdf) with ordinary kriging using the normal score variogram. (step 1 in Fig. 1); the conditioning information consists of n neighbouring data of both original normal score data $y(\mathbf{x}_\alpha)$ and previously simulated values $y^{(l)}(\mathbf{x}_\alpha)$ where l being the realization number (l= 1,…,L);
   - Draw randomly a simulated value $y^{(l)}(\mathbf{x}_\alpha)$ from the ccdf and put it into the data (step 2 in Figure 1);
   - Proceed to the next estimation location along the random path and repeat the previous two steps and do this until all m location are simulated;
   - Proceed with next simulation by repeating the previous steps until all L realizations are generated.

4. Finally the simulated values are back transformed to their original values by applying the inverse of the normal score transform $\phi(.)$ (Figure 3).

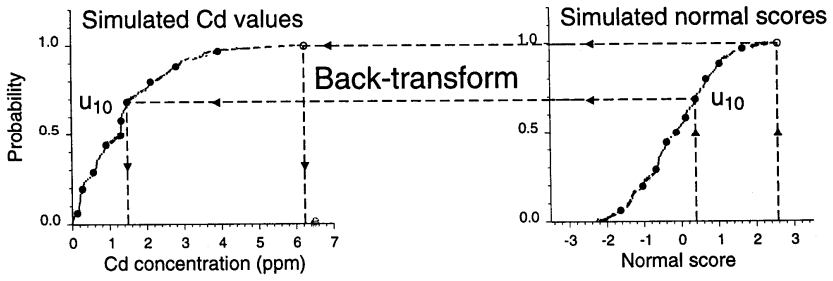

Figure 3. Back transformation of simulated normal scores to the original Z-values

Each of these *L* simulations are a realization of the unknown spatial distribution of the variable *Z*. Differences between them provide a measure of the spatial uncertainty about *Z*. The estimated values as well as the probability of exceeding a threshold can be obtained from post-processing the set of *L* realizations on a pixel basis.

In this document maps were made for 80% and 95% probability of exceeding a critical threshold value z_c with the following legends and interpretations:

1. For 95% probability

$\text{Prob}\{ z(\mathbf{x}_o) > z_c \} < 0.95$:	Grey; large uncertainty
$\text{Prob}\{ z(\mathbf{x}_o) > z_c \} \geq 0.95$:	Red; almost certainly $z(\mathbf{x}_o) > z_c$

2. For 80% certainty

$\text{Prob}\{ z(\mathbf{x}_o) > z_c \} < 0.20$:	Green; almost certainly $z(\mathbf{x}_o) < z_c$
$0.20 < \text{Prob}\{ z(\mathbf{x}_0) > z_c \} < 0.80$:	Grey; too large uncertainty to be certain about the estimated values
$\text{Prob}\{ z(\mathbf{x}_o) > z_c \} \geq 0.95$:	Red; almost certainly $z(\mathbf{x}_o) > z_c$

4 Exploratory data analysis

In addition to the 3.644 observations made in 1997, 708 data which were obtained through 'tegen expertises' were used for the analysis of PSD. For the sake of clarity, in this report, the data that were formerly used in 1997 were referred to as "old data", the additional set as "new" data and the data that contained both as the "all" data. The general statistical distribution of these three datasets is summarized in Table 1 and Figure 4.

Table 1. General statistics of the old, new and all data

	Old data	New data	All data
Number of observations	3644	708	4352
Mean PSD (%)	31.9	40.6	33.3
Standard deviation	14.6	16.6	15.3
Maximum (%)	137.3	104.6	137.3
Median (%)	29.8	37.0	31.1
Minimum (%)	2.0	3.9	2.0

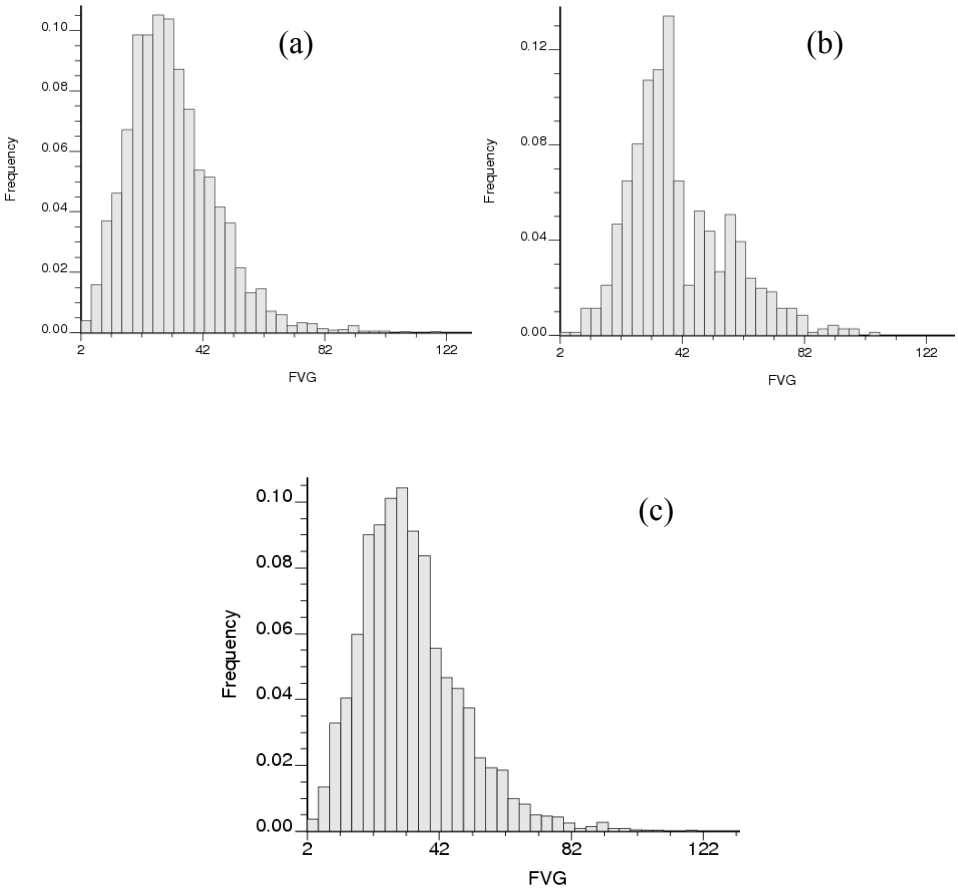

Figure 4. Histogram of the old (a), the new (b), and all data (c).

Based on the statistical parameters presented in Table 1, it is clear that the new dataset has a higher mean and a median. This is a surprise since the additional data were taken from fields of which the farmers expected that they were wrongly delineated as phosphate saturated in 1997. However, one can in general conclude that the old and all data have similar statistical parameters as can be seen from the histograms in Figure 4.

To investigate the spatial distribution, location maps were produced by posting the data values with respect to their sampling locations. The locations maps in Figure 5 indicate a significant clustering of the new observations in specific locations. This clustering of observation provides a possibility of modelling the spatial correlation that exists between measurements within a short distance. The other important remark that can be drawn from the location maps is that the spatial variability of the data is quite large. Since PSD is depending on land management practices specific for every field there is an inherently large local variability in the data.

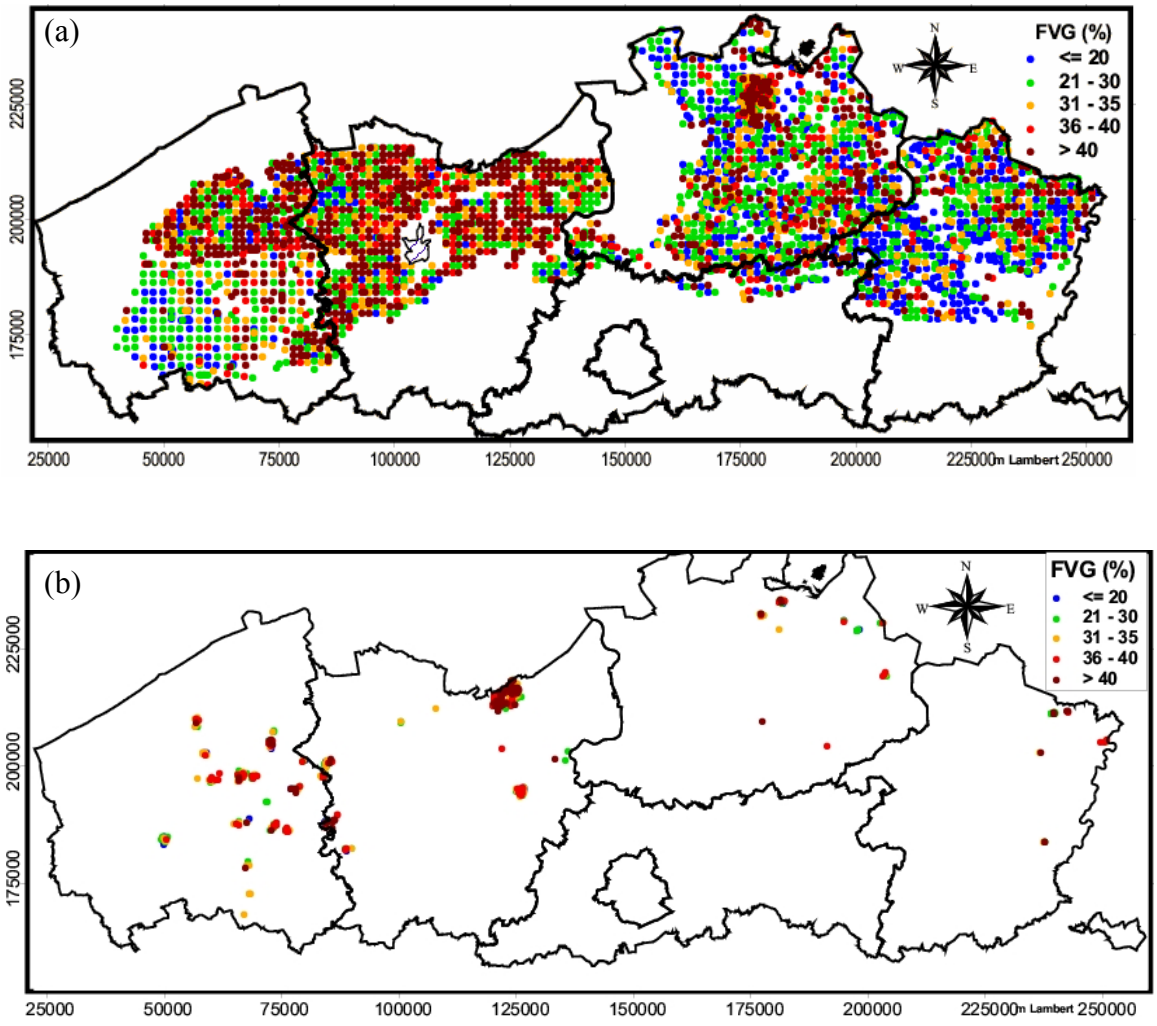

Figure 5. Location maps of the old (a) and the new (b) datasets

5 Result and discussion

5.1. Sequential Gaussian Simulation

Both the 3644 of the old and the 4352 of the complete data set were analysed using SGS. The variograms of these data were modelled individually after normal score transformations of the respective data (Figure 6). The variograms were modelled with double spherical structure with parameters specified in Table 2.

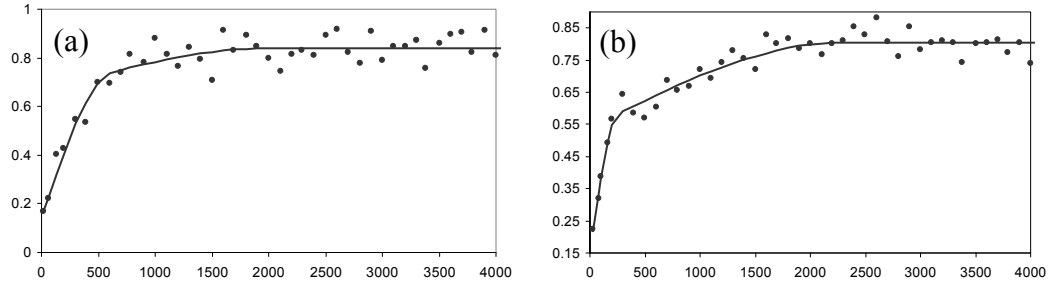

Figure 6. Variograms of the normal score values of the old (a) and all (b) data (X-axis: lag h in m, Y-axis: $\gamma(\mathbf{h})$)

Table 2. Parameters of the variograms of normal score values of old (a) and all (b) data

Variogram Parameters	Old data	All data
Nugget	0.16	0.15
$Sill_1$	0.57	0.39
$Range_1$ (m)	777	255
$Sill_2$	0.11	0.27
$Range_2$ (m)	2002	2311

Using the variogram parameters and SGS for every estimation location 500 equi-probable realizations were generated. After post-simulation processing the realizations, maps that show the predicted PSD values for both datasets were obtained, Figure 7.

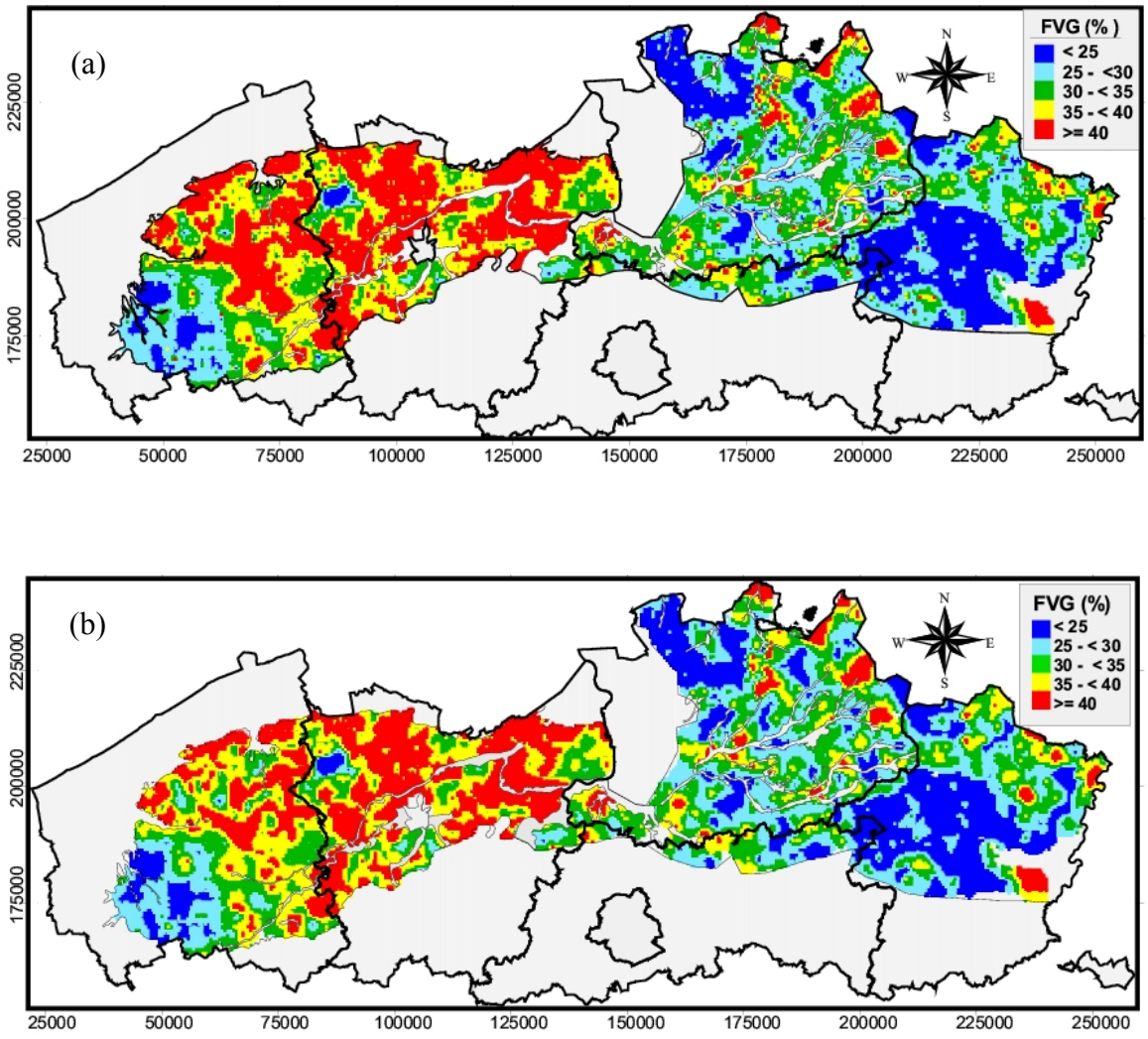

Figure 7. The estimated PSD using old data (a) and all data (b) with SGS

The two maps in Figure 7 display some differences mainly in areas where the additional 708 data were located. Since some of these additional data have lower values in their neighbourhood the estimated values in those locations were reduced. For instance, in West-Flanders around Roeselare and Lichtervelde a smaller area is estimated to have a PSD above 40% with the all dataset. A summary of the comparison of the two maps is given in Table 3.

Table 3. Area for intervals of estimated PSD values as estimated using the old and all data sets with SGS

PSD (%)	Area (km^2)	
	Old data	All data
< 25	1348	1415
25 – 29	1668	1618
30 – 34	1688	1733
35 – 39	1378	1477
≥ 40	1464	1319

The 95% probability of exceeding the 35% PSD was computed following the SGS procedure. The result for both the old and all data is presented as probability maps in Figure 8.

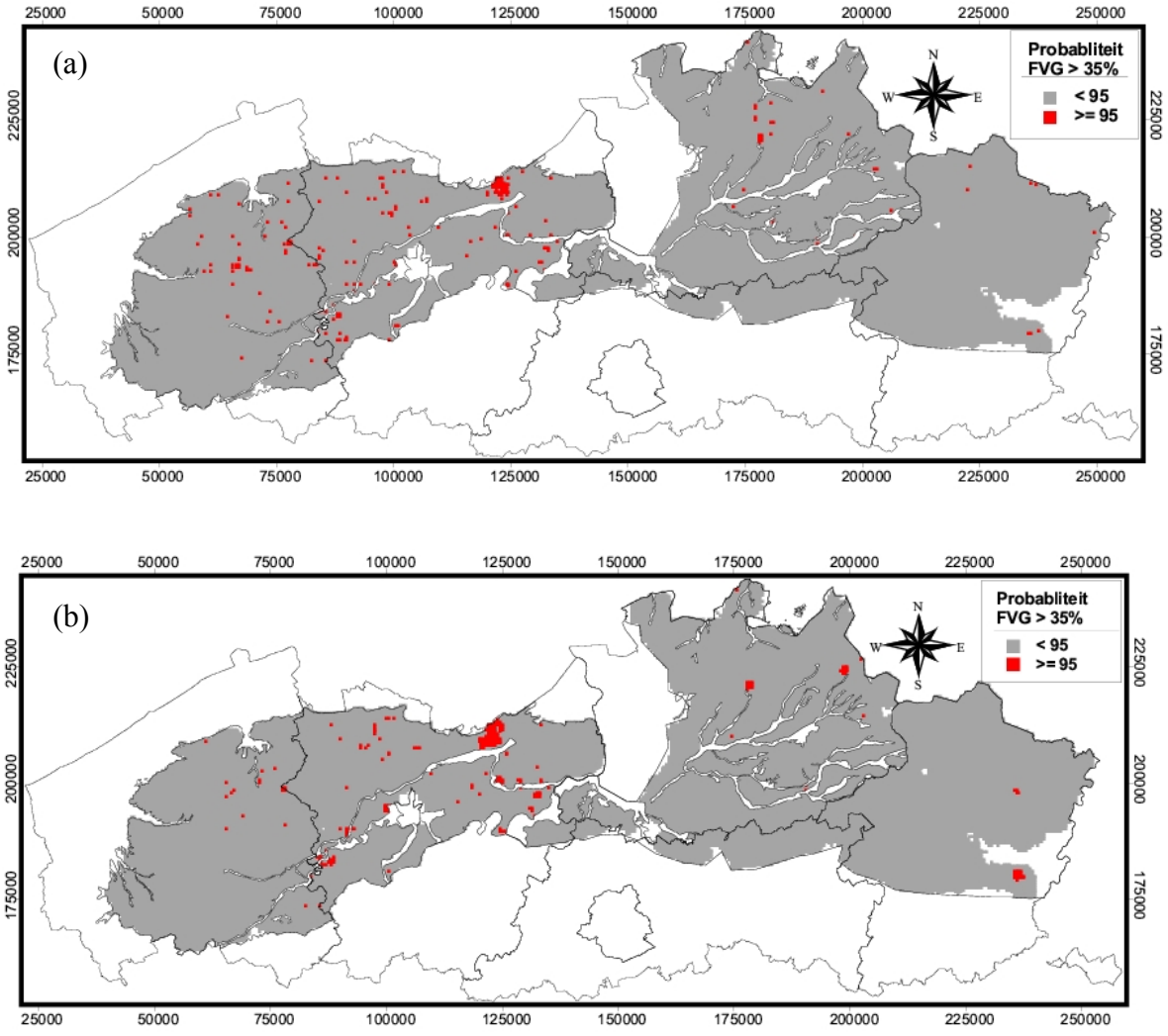

Figure 8. The 95% probability of exceeding 35% PSD using old data (a) and all data (b) with SGS

Using only the old data, 54 km^2 was delineated to have a 95% probability of exceeding 35% PSD. Most of the delineated zones are rather individual as scattered pixels (Figure 8a). Using the complete data resulted in a larger area, 64 km^2, of more contiguous patches (Figure 8b) which is due to reduced prediction uncertainties due to the additional measurements. This result indicates the importance of considering the additional data in the analysis.

5.2. Alternative-1: 25% and 30% PSD

The 500 realizations generated following the SGS procedure using all data were processed to compute the 95% probability of exceeding the 25% and 30% PSD thresholds (Figure 9).

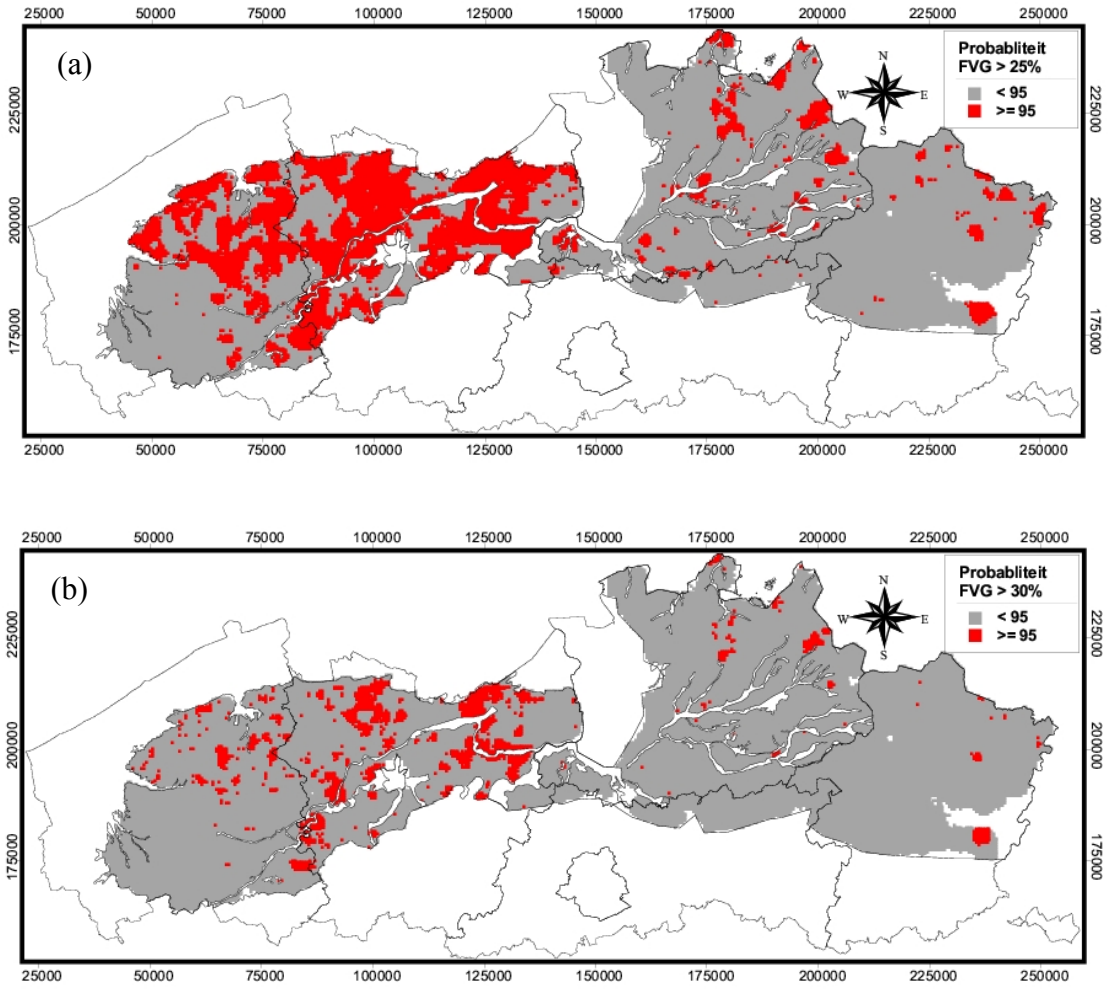

Figure 9. The 95% probability for PSD > 25% (a) and PSD > 30% (b), using all data and SGS

The area covered with 95% probability for PSD > 25% is 1812 km^2 and for PSD > 30% it is 460 km^2.

5.3. Alternative-2: probability 80 %

The second alternative was to lower the probability level from 95% to 80% and compute the probability of exceeding the 25%, 30% and 35% PSD thresholds using all data (Figure 10). The classical probability level of 95 % used in biometry is very strict in conditions where spatial variability is investigated on the basis of a limited number of point observations. In literature the 80% probability level has been used for delineation of contaminated sites. Garcia and Froidevaux (1997) for example classified an area as safe when probability < 20%, as polluted with probability > 80% and as uncertain or unclassified when the probability was between 20 and 80%.

By changing the probability level from 95% to 80%, the area that was delineated as having a PSD of more than 35% increased importantly (Table 4).

Table 4. Area with PSD > 25%, 30% and 35% as determined at 80% probability with all data and SGS

PSD (%)	**Area (km^2)** determined at 80% probability	**Area (km^2)** determined at 95% probability
> 25	3827	1812
> 30	1928	460
> 35	578	64

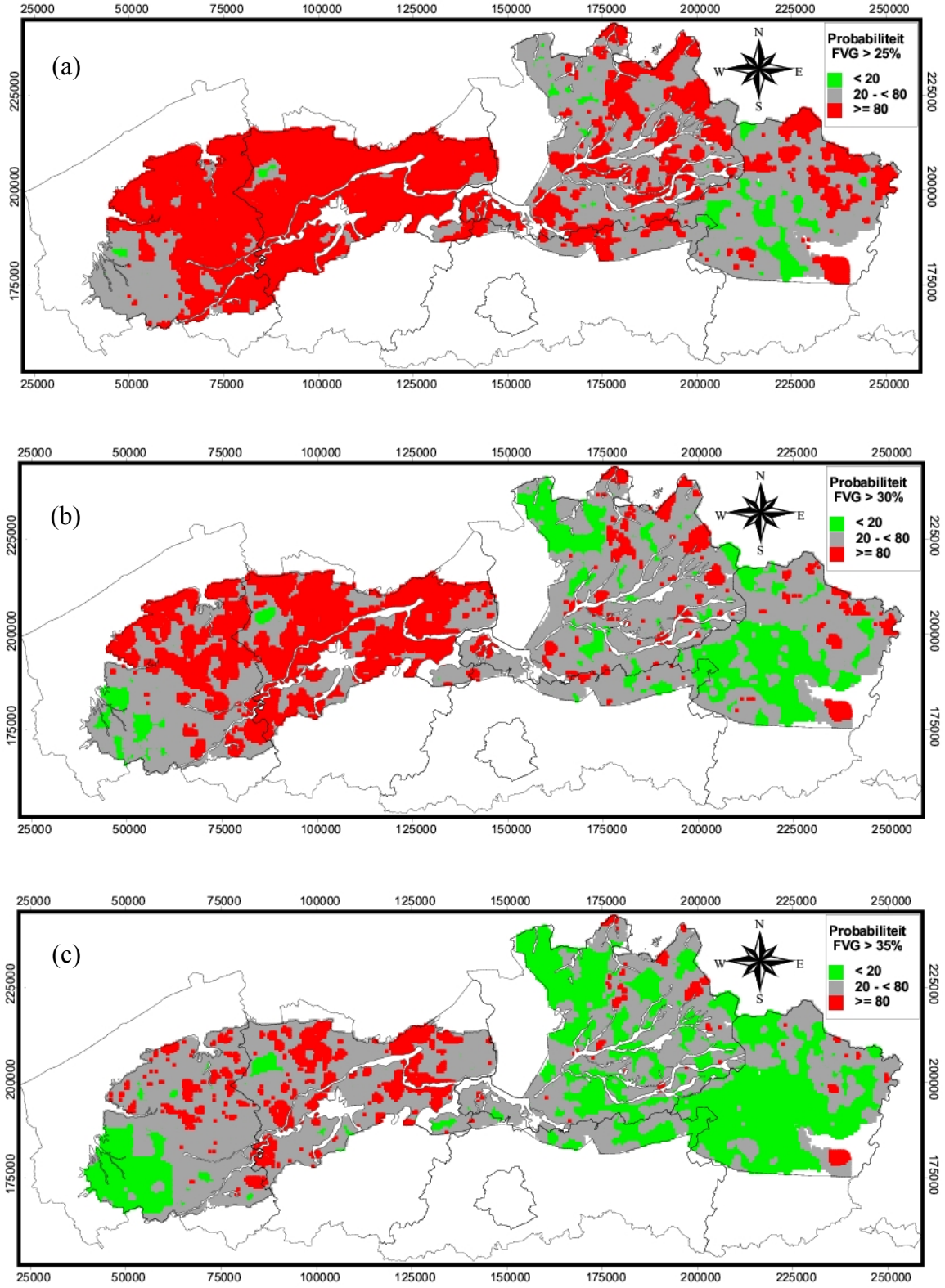

Figure 10. The 80% probability of exceeding the 25% (a), 30% (b) and the 35% (c) PSD thresholds, using all data and SGS

5.4. Alternative-3: ordinary block kriging

In 1997 the delineation of phosphate saturated regions in Flanders was done with ordinary block kriging. To compare the performance of this technique with SGS both the old and the all dataset were analyzed with the ordinary block kriging.

First the variograms of the two data sets were modelled (Figure 11). The data were used in their original values without any transformation.

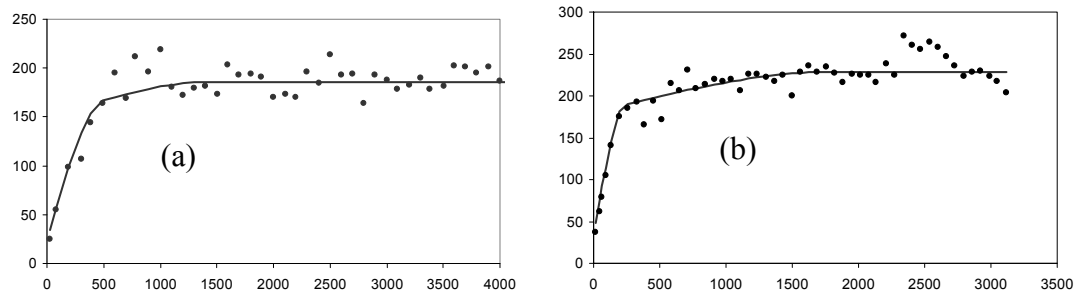

Figure 11. Variograms of the old (a) and the all (b) data using untransformed values (X-axis: lag h in m, Y-axis: $\gamma(\mathbf{h})$)

Both variograms were modelled using a double spherical structure (Table 5).

Table 5. Parameters of the old and the new data variograms

PSD	Old data	All data
Nugget	23	30
$Sill_1$	125	150
$Range_1$ (m)	500	240
$Sill_2$	38	49
$Range_2$ (m)	1450	1840

Using the parameters of these variograms, block OK estimations of PSD were produced both for the old and all data for blocks of 500 m by 500 m (Figure 12).

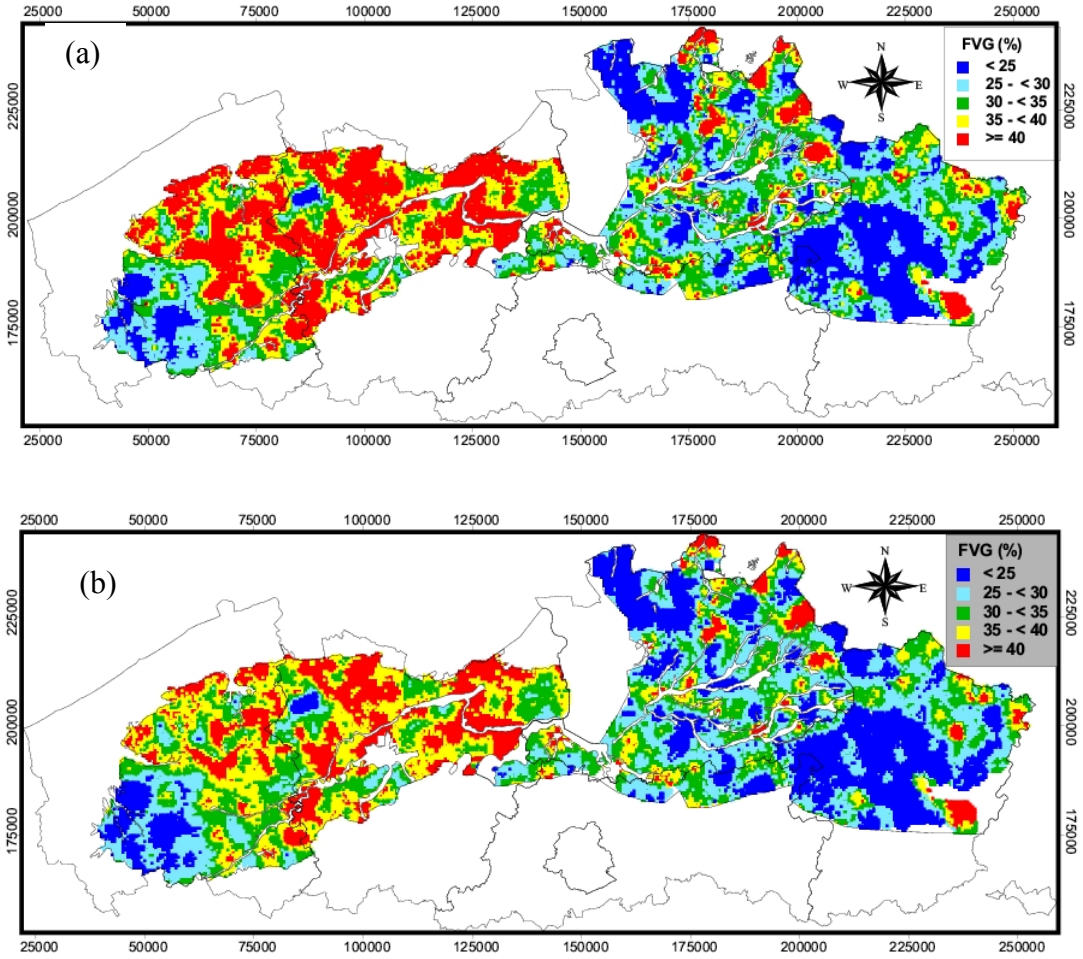

Figure 12. Block estimations of PSD using the old (a) and all (b) data (block size: 500 m)

For both datasets the block estimated values of PSD using block OK and SGS were comparable. The areas covered per interval of estimated values are presented in Table 6.

Table 6. Area in intervals of PSD as estimated using the old and all data with block ordinary kriging

PSD (%)	Area (km^2)	
	Old data	All data
< 25	1491	1528
25 – 29	2008	1687
30 – 34	1293	1691
35 – 39	1333	1450
≥ 40	1437	1263

As compared to the old data, the all dataset resulted in larger areas with values between 30 and 39 % PSD but a smaller number of pixels with a PSD of more than 40%.

Another difference with the 1997 procedure is the critical z-value (from the standard normal table) that was used to calculate the 95% probability interval. In statistical inference one can choose between a one- or two-tailed t-test according to the tested hypothesis. A one-tailed test is used to check the significance of a value being larger, or smaller, than a given threshold. This is the case with PSD. Then the critical value for testing the hypothesis at a 95% confidence interval is 1.64 instead of the 1.96 used in 1997. Hence in the formula to calculate the 95% probability to exceed a critical threshold we have to deduct from the estimated value 1.64* kriging standard deviation instead of 2* kriging standard deviation.

The 95% probability of exceeding the threshold of 35% PSD as calculated from the estimated block values and the associated block kriging variances is shown in Figure 13, using the 1.64 z-value.

At 95% certainty the old data resulted in a 197 km^2 area with PSD > 35% (Figure 13(a)), with all data this reduced to 157 km^2. But the spatial pattern was much less scattered with all data compared to the old data.

For comparative reasons the 95% probability of exceeding the 30% and 40% PSD were also conducted using the old data and applying the ordinary block kriging procedure (Figure 14).

For the threshold values of 30% and 40% PSD, 578 and 69 km^2 were delineated to be phosphate saturated at 95% probability, respectively. In 1997 the old data were analyzed with block ordinary kriging using a z-value of 1.96 as test criterion which resulted in a delineation of 73 km^2 with a PSD $\geq$ 40% at 95 % probability (actually 97.5 %). Using the same method (but with a t-value of 1.64) we have obtained an area of 69 km^2 which confirms that our procedure was comparable.

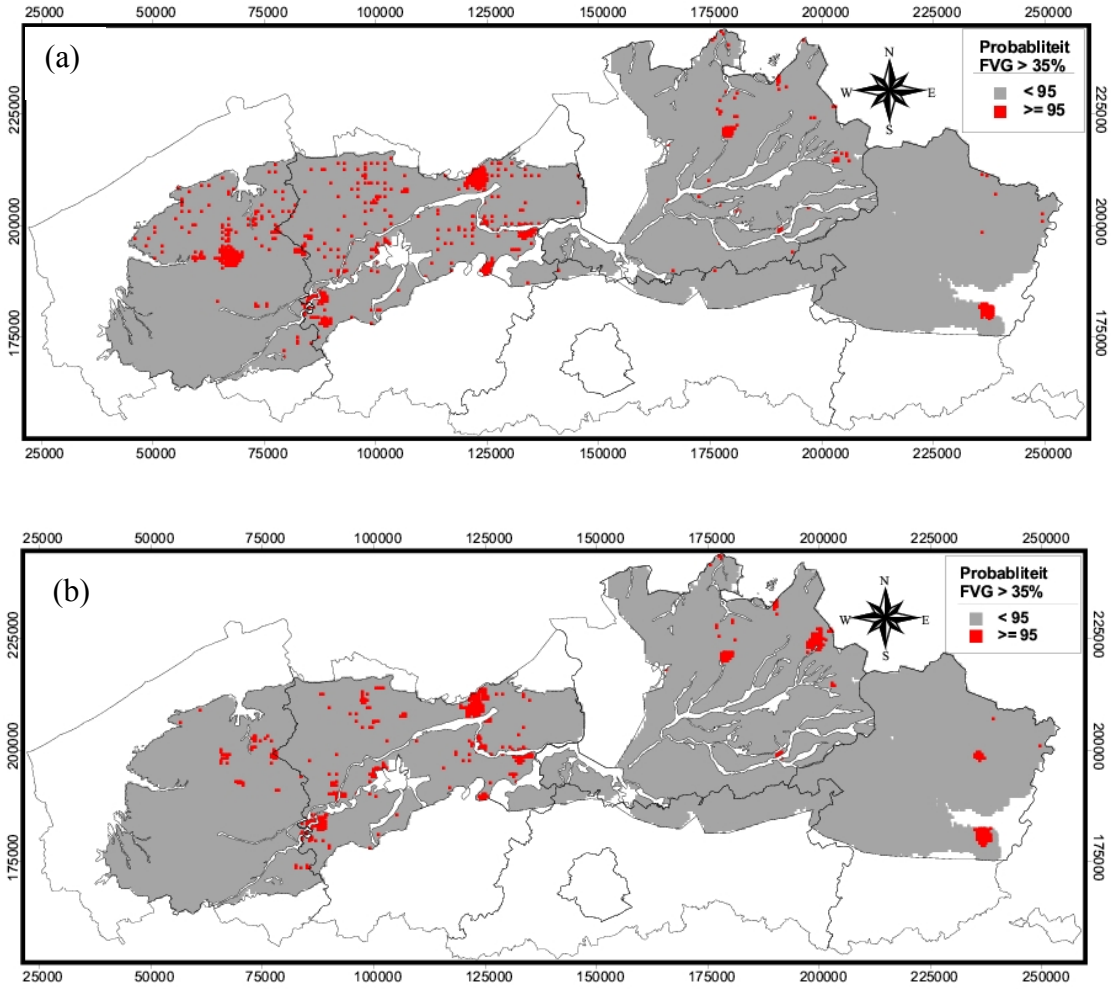

Figure 13. The 95% probability for PSD > 35% using block ordinary kriging with the old (a) and all (b) data

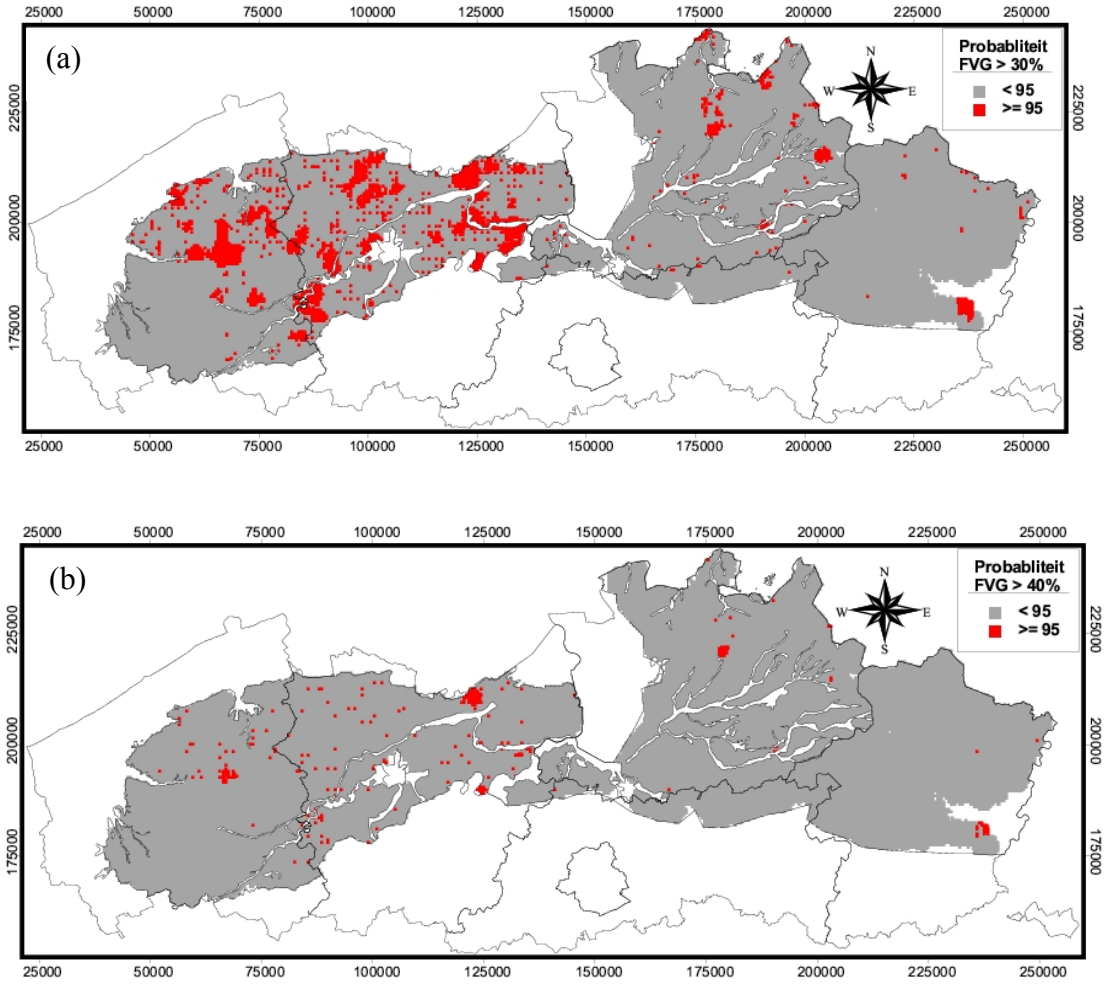

Figure 14. The 95% probability for PSD > 30% (a) and PSD > 40% (b) using the old data and ordinary block kriging

6 Discussion and conclusions

In 1997 the geostatistical method employed was ordinary block kriging using the kriging variance to calculate a critical 95 % probability level of exceeding a PSD of 40 %.

Compared to this procedure the following modifications were implemented:

1. Include 708 extra observations of PSD increasing the data density and thus reducing the uncertainty associated with interpolations.
2. Use a simulation rather than an estimation procedure. This has the advantage that a large number of equi-probable values can be generated allowing the numerical reconstruction of a distribution of possible outcomes at every location. Such a distribution allows to directly derive the probability of

exceeding a critical threshold taking account of the local, rather than the global, uncertainty.

3. Use a critical PSD of 35 %.

Alternative considerations were:

1. To reduce the probability level from 95 % to 80 % in combination with a PSD of 35 % to have a less stringent decision rule.

2. To reduce the PSD to 30 or 25 % and maintain the probability level of 95 %.

The considerations of these two alternatives resulted in an important increase in the size of the area delineated as potentially phosphate saturated (Table 7).

Table 7. An overview of the area delineated using SGS according to several alternatives

PSD (%)	Area (km^2) 95% probability	Area (km^2) 80% probability
> 25	1812	
> 30	460	
> 35	**64**	578

7 List of References

Chilés JP. and Delfinder P. 1999. Geostatistics. John Wiley & Sons.

Fagroud M and Van Meirvenne M. 2002. Accounting for soil spatial autocorrelation in the design of experimental trials on water-use efficiency. Soil Sci. Soc. Am. J., 66**,** 1134-1142.

Garcia M and Froidevaux R. 1997. Application of geostatistics to 3-D modelling of contaminated sites: A case study. In: Soares et al. (eds) geoENV I-Geostatistics for Environmental Applications. Kluwer Academic Publishers. 309 – 325.

Grunwald S, Reddy KR, Newman S and DeBusk WF. 2004. Spatial variability, distribution and uncertainty assessment of soil phosphorus in a south Florida wetland. Environmetrics 15 (8): 811-825.

Goovaerts P. (1999). Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 89,1-45.

Goovaerts P. 2000. Estimation or simulation of soil properties? An optimization problem with conflicting citeria. Geoderma 97, 165-186.

Lapen DR, Topp GC, Hayhoe HN, Gregorich EG and Curneo WE. 2001. Stochastics simulation of soil strength / compaction and assessment of corn yield risk using threshold probability patterns. Geoderma. 104: 325 - 343

Lookman R. (1995). Phosphate chemistry in excessively fertilised soils. Doctoraatsproefschrift, Katholieke Universiteit Leuven, 144 pp.

Van der Zee S.E.A.T.M, Van Riemsdijk W.H. & De Haan F.A.M. (1990a). Het protokol fosfaatverzadigde gronden. Deel I. Toelichting. Vakgroep Bodemkunde en Plantenvoeding, Landbouwuniversiteit Wageningen

Van der Zee S.E.A.T.M, Van Riemsdijk W.H. & De Haan F.A.M. (1990a). Het protokol fosfaatverzadigde gronden. Deel II. Technische uitwerking.Vakgroep Bodemkunde en Plantenvoeding, Landbouwuniversiteit Wageningen

Van Meirvenne M. & Goovaerts P. (2001). Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold. Geoderma, 102, 75-100.

Yu-Pin Lin, Tsun-Kuo Chang and Tung-po Tung. 2001. Characterization of soil Lead by comparing Sequential Gaussian Simulation, Simulated annealing simulation and Kriging methods. Environmental Geology. 41: 189-199.