Forsvarsudvalget, Grønlandsudvalget 2011-12
FOU alm. del Bilag 171, GRU alm. del Bilag 23
Offentligt

Spatial statistical analysis of contamination level of ^{241}Am and ^{239}Pu
Thule, North-West Greenland

Risø-R-Report

Jens Strodl Andersen, JSA EnviroStat
Risø-R-1791(EN)
October 2011

Risø DTU
National Laboratory for Sustainable Energy

Author: Jens Strodl Andersen, JSA EnviroStat
Title: Spatial statistical analysis of contamination level of ^{241}Am and ^{239}Pu, Thule, North-West Greenland

Risø-R-1791(EN)
October 2011

Abstract
A spatial analysis of data on radioactive pollution on land at Thule, North-West Greenland is presented. The data comprises levels of ^{241}Am and ^{239,240}Pu on land. Maximum observed level of ^{241}Am is 2.8×10^{5} Bq m^{-2}. Highest levels were observed near Narsaarsuk. This area was also sampled most intensively. In Grønnedal the maximum observed level of ^{241}Am is 1.9×10^{4} Bq m^{-2}. Prediction of the overall amount of ^{241}Am and ^{239,240}Pu is based on grid points within the range from the nearest measurement location. The overall amount is therefore highly dependent on the model. Under the optimal spatial model for Narsaarsuk, within the area of prediction, the predicted total amount of ^{241}Am is 45 GBq and the predicted total amount of ^{239,240}Pu is 270 GBq.

ISSN 0106-2840
ISBN 978-87-550-3933-9

Contract no.:

Group's own reg. no.:

Sponsorship:

Cover :

Pages:73
Tables: 24
References: 50

Information Service Department
Risø National Laboratory for Sustainable Energy
Technical University of Denmark
P.O.Box 49
DK-4000 Roskilde
Denmark
Telephone +45 46774005
bibl@risoe.dtu.dk
Fax +45 46774013
www.risoe.dtu.dk

Contents

1 Geographic location of the measurements

The geographic location is Thule, north-west Greenland. The gamma radiation level of americium-241 (Bq/m^2) is measured in the laboratory on soil samples and by a portable spectrometer on site. A total of 878 measurements were taken on 614 locations. On some locations several measurements (up to 5 per location) were taken to explore small scale variability. The coordinates on these locations were shifted according to Figure 1. The first data point is kept; the second is shifted 20cm to the South, the third is shifted 60 cm East from the second point, the fourth 2 m north of the third and finally the fifth is shifted 5m west of the fourth (Figure 1).

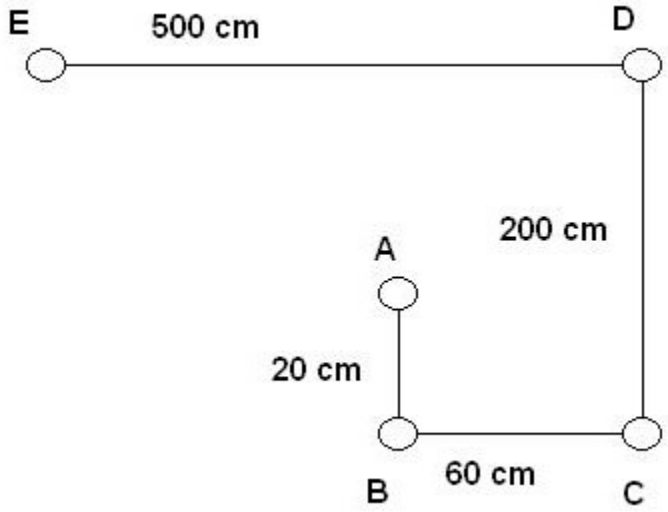


Figure 1 Shift pattern for data points at coincident locations

Coastline data was downloaded from (http://rimmer.ngdc.noaa.gov/) with a 1:250.000 resolution (Figure 2).

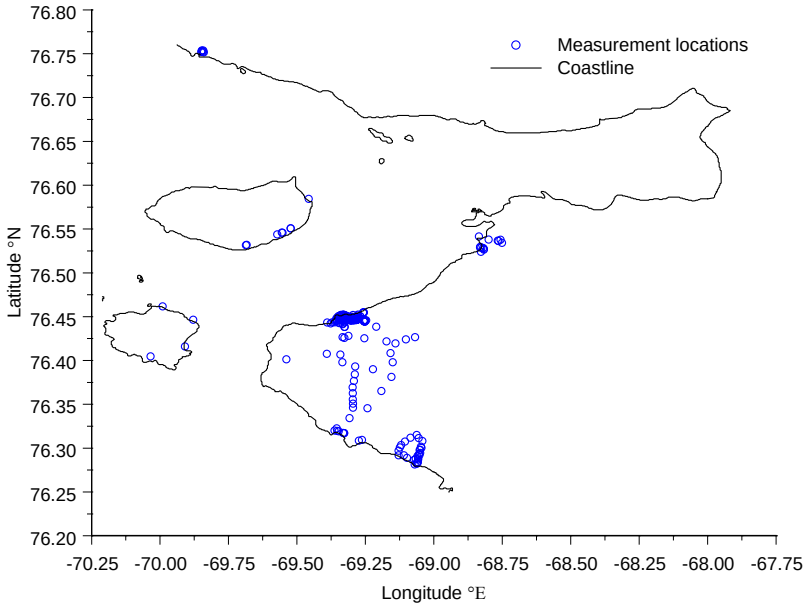


Figure 2 Coastline and location of measurements (Longitude,Latitude)

To be able to work and do spatial statistics with the data the coordinates were transformed to metric coordinates. Thule is located in UTM-zone 19. The UTM coordinates are shown in Figure 3.

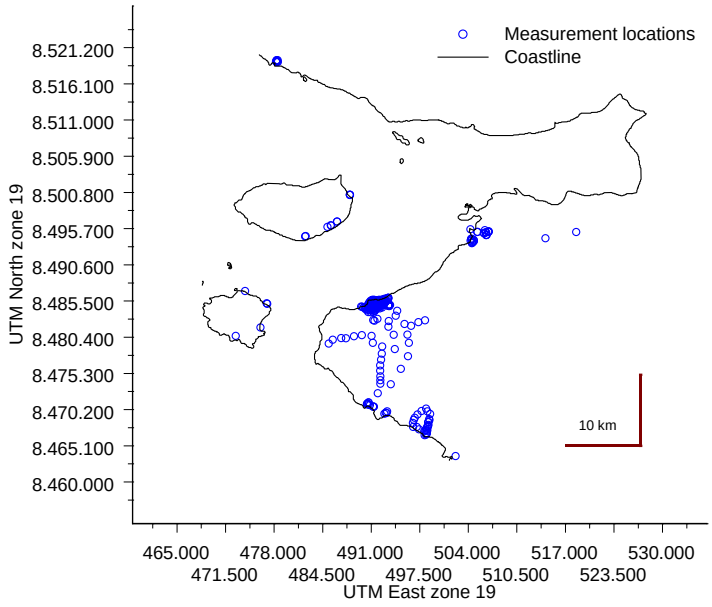


Figure 3 Coastline and location of measurements (UTM-zone 19)

The data are divided into different regions and analysed separately because the regions assumed to be independent, in the sense that the level of americium-241 is independent from region to region (except region 2 being a subset of region 1). The regions are shown in Figure 4.

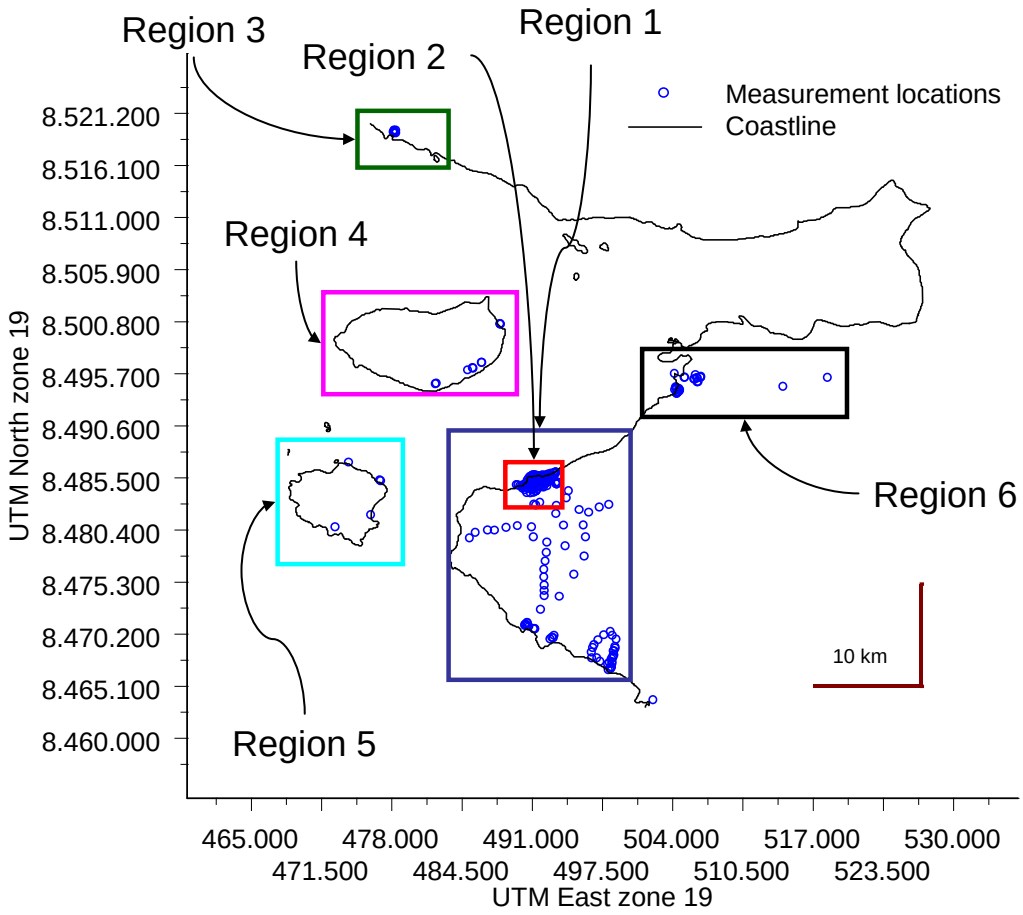


Figure 4 Split in different geographic regions

Detailed spatial statistical analysis is performed for region the area around and south of Narsaarsuk including Grønnedal (region 1), for the area around Narsaarsuk (region 2) and also for a sub-region of approximately 1x1 km of Region 2.

In the following regions only summary statistics is provided because data is to sparse to do spatial statistics:
Region 3, the area around Moriusaq
Region 4, Saunders Ø
Region 5, Wolstenholme Ø
Region 6, the area near the Thule airbase

2 Statistical methods

The data is analysed using the package geoR (http://www.leg.ufpr.br/geoR) that provides functions for geostatistical data analysis using the software R (http://www.r-project.org/). Furthermore the packages splancs and maptools were used for calculating neighbourhood distances and importing ESRI-Shapefiles for the altitude data in region 1. Furthermore code was developed, to do simulations of the predicted surface and calculate the total level of Americium and Plutonium along with 95% confidence intervals.

The basic idea in spatial statistics is that data points close in space are more correlated than data points further apart. The geographic variation/correlation structure is often described by a variogram, parameterised by *nugget*, *sill* and *range*. *Nugget* is the variation between samples taken at coincident or very close locations; *sill* is the variance between points at a distance where they are no longer correlated (that distance is denoted *range*). The parameters are illustrated in Figure 5. In the software used the *partial sill = sill – nugget* is estimated. The *sill* is then found by adding *nugget* and *partial sill*.

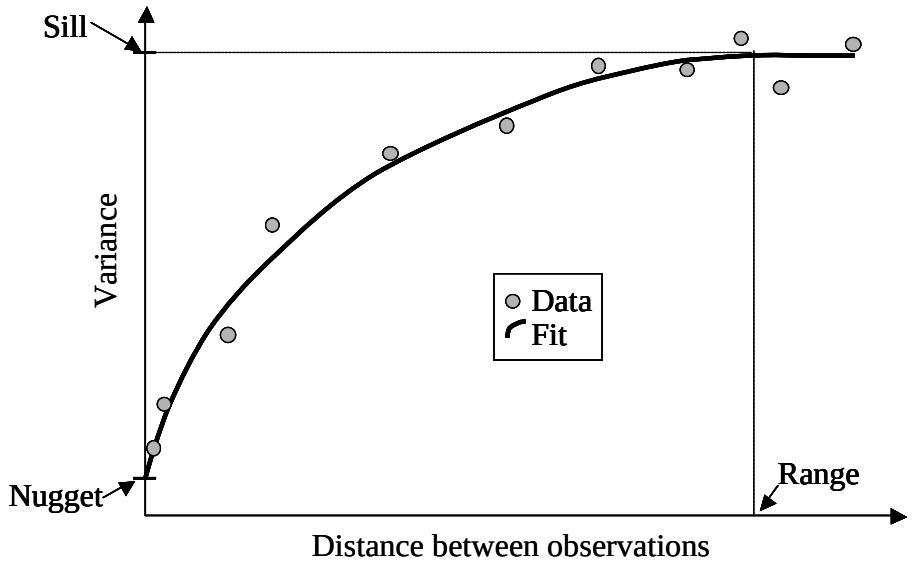


Figure 5 Theoretical variogram

2.1 Likelihood based parameter estimation

The correlation structure is estimated using model based spatial statistics. The main function estimate the parameters of the Gaussian random field model, specified as:

$$Y(x) = mu(x) + S(x) + e$$

where

- *x* defines a spatial location.
- *Y* is the variable observed ($\log_{10}(^{241}\mathrm{Am})$).
- *mu(x) = X %*% beta* is the mean component of the model (a mean or a trend).

- *S(x)* is a stationary Gaussian process with variance *sigma^2* (partial sill) and a correlation function *rho(h)* parameterised by *phi* (the range parameter) as: *rho(h) = 1 - 1.5 * (h/phi) + 0.5(h/phi)^3 if h < phi , 0 otherwise* (h is the distance between point pairs) and the anisotropy parameters *phi_R* and *phi_A* (anisotropy ratio and angle, respectively) .

- *e* is the error term with variance parameter *tau^2* (nugget variance).

A parameter λ allows for the Box-Cox transformation of the response variable *Y*. If used (i.e. if $\lambda \neq 1$) *Y(x)* above is replaced by *g(Y(x))* such that

$$g(Y(x)) = ((Y^{\lambda}(x)) - 1)/\lambda .$$

Two particular cases are $\lambda = 1$ which indicates no transformation and $\lambda = 0$ indicating the log-transformation.

The mean component of the model is assumed to be a constant $mu(x) = beta_0$ or a first order polynomial on the coordinates $mu(x) = beta_0 + beta_1 \bullet x + beta_2 \bullet y$.

The geometric anisotropy is described by the Anisotropy angle *phi_A*, defined here as the azimuth angle of the direction with greater spatial continuity, i.e. the angle between the y-axis and the direction with the maximum range. Anisotropy ratio *phi_R*, defined here as the ratio between the ranges of the directions with greater and smaller continuity, i.e. the ratio between maximum and minimum ranges. Therefore, its value is always greater or equal to one.

For all datasets, a full model with all parameters (including trend and anisotropy) are estimated simultaneously in the maximum likelihood estimation. To avoid overfitting, the model is reduced based on Akaike's Information Criteria *AIC* and Bayesian Information Criteria, *BIC*.
Akaike Information Criteria is defined as AIC=-2 ln(L) + 2 p where L is the maximised likelihood and p is the number of parameters in the model.
Bayesian Information Criteria is defined as BIC=-2ln(L) + p log(n), where n is the number of data, L is the maximised likelihood and p is the number of parameters in the model.

AIC is aimed at finding the best approximating model to the unknown true data generating process. AIC tries to select the model that most adequately describes an unknown, high dimensional reality. This means that reality is never in the set of candidate models that are being considered. BIC is designed to find the most probable model given the data BIC tries to find the true model among the set of candidates. Numerically BIC differs from AIC only in the second term which now depends on sample size n.

An individual AIC value and individual BIC value is meaningless, but differences in AIC values and BIC values are meaningful. For both information criteria a smaller value is better.
There is no consensus about which criteria is the optimal to use. When comparing the Bayesian Information Criteria and the Akaike's Information Criteria, the penalty for additional parameters is higher in BIC than AIC. Here, both criteria are reported.

In general the two criteria agree on the selected model and in case of disagreement both models is reported and discussed and one is selected.

2.2 Validation

The estimated model/correlation structure is validated by looking at the residuals of the model. The residuals should be normal distributed and there should be no spatial correlation. Furthermore the profile-likelihood of each of the spatial parameters is presented.

2.3 Spatial prediction

Spatial prediction onto a grid can be done once the spatial correlation structure is estimated. Only grid points on land are included. To estimate the concentration on a location that has not been sampled a weighted average of all points within the *range* is calculated (Figure 6). The weight allocated to each point depends on the distance and the estimated correlation structure.

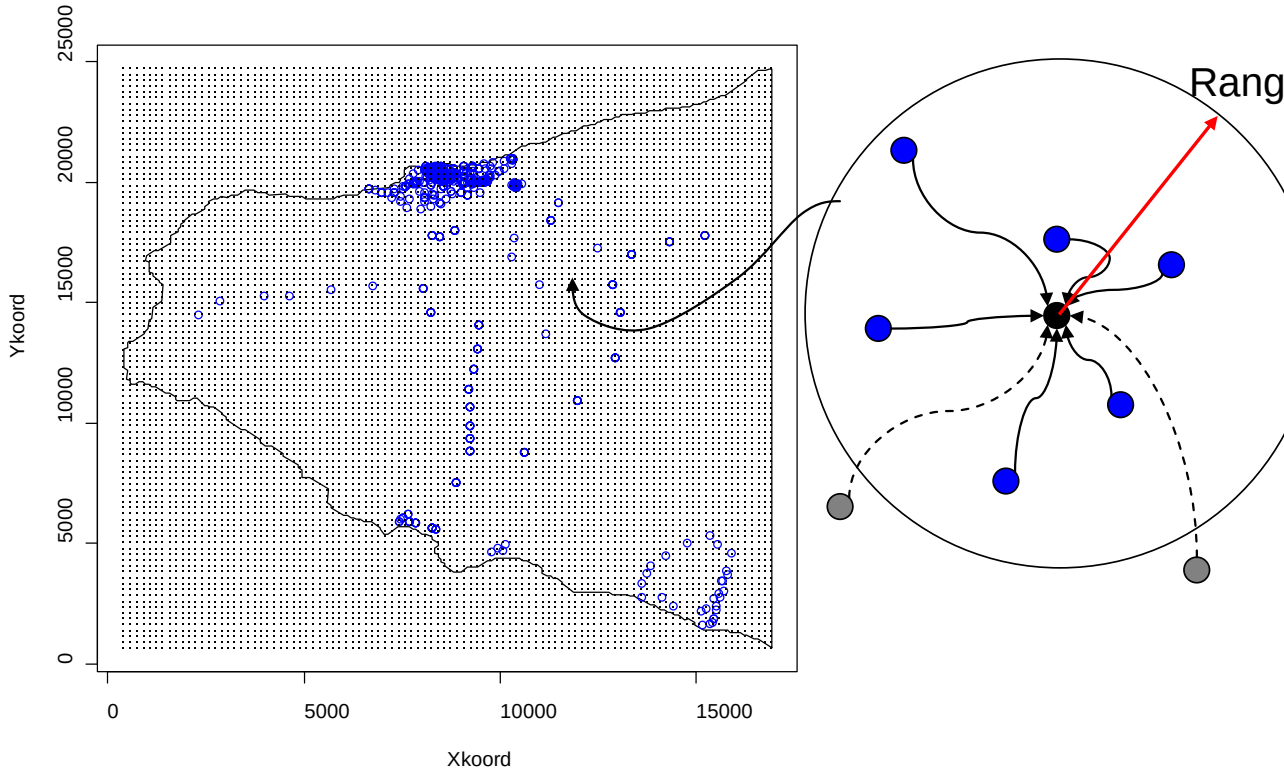


Figure 6 Illustration of spatial prediction on a grid. The two grey locations (outside the range*) have no weight in the prediction.*

If significant anisotropy or trend is present these parameters are also included in the weighting and prediction.

2.4 Uncertainty of spatial prediction

Confidence intervals for the estimated spatial parameters are calculated as profile-likelihood intervals, so the are not necessary symmetric. The 90% and 95% confidence intervals are calculated as the value of the maximised log-likelihood - $0.5\times\chi^2_{0.90}(1)$ or $0.5\times\chi^2_{0.95}(1)$, respectively. The intersections of the profile likelihood curves are shown in the profile likelihood plots for the spatial parameters: nugget, partial sill and range.

The amount and uncertainty of ^{241}Am(Bqm^{-2}) at each grid point (the prediction locations) is based on the selected model. Based on the model a predicted concentration and a prediction variance is calculated. The error is assumed normal distributed so standard statistical uncertainty measures can be calculated for the predicted values at each grid point. The following 3 is presented

- The relative standard deviation
- The upper 95% confidence limit
- The probability of exceeding a threshold ($\log_{10}(^{241}\text{Am})=2$ and $\log_{10}(^{241}\text{Am})=3$)

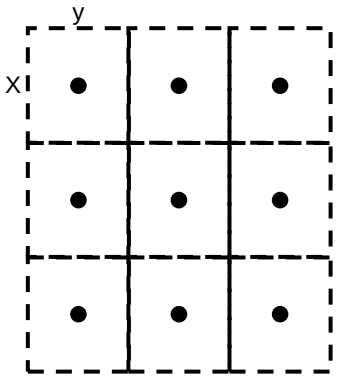
2.5 Prediction of the total amont of Plutonium ^{239,240}Pu

Based on the predicted contamination levels of ^{241}Am it is possible to predict the contamination level of Plutonium (^{239,240}Pu). The relationship between ^{241}Am and ^{239,240}Pu is shown in the table below

Level of ^{241}Am	<50Bqm^{-2}	>=50Bqm^{-2} and <=1kBqm^{-2}	>1kBqm^{-2}
Calculated level of ^{239,240}Pu	3×^{241}Am	4×^{241}Am	6×^{241}Am

Table 1 Calculation of the Plutonium level based on the Americium level

Each grid point (the prediction locations) represents a square.



For each grid point a predicted level and a predicted variance is calculated. The level of americium-241 in grid square *i* is calculated as the product of the area of the *i*'th square ($A_i=x\bullet y$) and the predicted level of Am_i in the *i*'th grid point: $A_i \times Am_i$. The sum over the entire area is equivalent to the sum over the entire grid $\sum(A_i \times Am_i)$. Since all grid points represents an area of the same size the sum can be calculated as $A \times \sum Am_i$.

The level ^{239,240}Pu is calculated by multiplying the predicted level of Am_i in the *i*'th grid point by the factors from the table above and then calculating the sum of ^{239,240}Pu for the entire area as described for ^{241}Am.

Prediction of the level of ^{241}Am and ^{239,240}Pu will only be performed in grid points located within the estimated *range* form the nearest observed location. This approach has been chosen to avoid a bias in the prediction. In the spatial statistical model applied, grid points with a larger distance to the nearest observed location than the estimated *range* will be set to the mean component of the model. This is based on the assumption that the best guess of the level in such a location is the mean. However, if sampling in the area of interest has been targeted at specific locations where the level is high, this will results in an upwards biased predicted level in the locations outside the estimated *range*. Leaving out the prediction in some remote grid points does not imply that the level is 0, just that there is no information to base the prediction on.

The uncertainty of the total predicted level of ^{241}Am and ^{239,240}Pu is estimated by simulation. Based on the estimated spatial model a predicted level and a predicted variance are calculated at each grid point. A random sample from each gridpoint is drawn from a $N(\mu_i,\sigma_i^2)$, where i denote the i^{th} grid point. The total level of ^{241}Am

and ^{239,240}Pu is then calculated as described above. This is repeated 10000 times and the 2.5% and 97.5% quantile is used as an empirical 95% confidence interval for the predicted level of ^{241}Am and ^{239,240}Pu.
The simulations are performed for all three models: The basic, the trend model and the model including anisotropy.

For the selected model in region 2 and the sub-region 2 further simulations are performed that also take into account the uncertainty of the estimated spatial parameters. This is done by sampling from the 95% confidence intervals of the nugget, partial sill and range. For each sample, the predicted level of americium and the related prediction variance is calculated and simulation is performed. The confidence intervals for the total level of ^{241}Am and ^{239,240}Pu will be wider than the confidence intervals based on the simulations from the selected model.

The following chapters will present the spatial statistical analysis, divided into the described regions

3 Analysis of data in Region 1

This section describes the analysis of data in region 1 (see Figure 4). All spatial statistics is performed on $\log_{10}$ transformed data.

3.1 Geographical location

In region 1 a total of 805 measurements were taken on 561 locations, 625 soil samples and 180 CAP measurements. The locations are shown in Figure 7

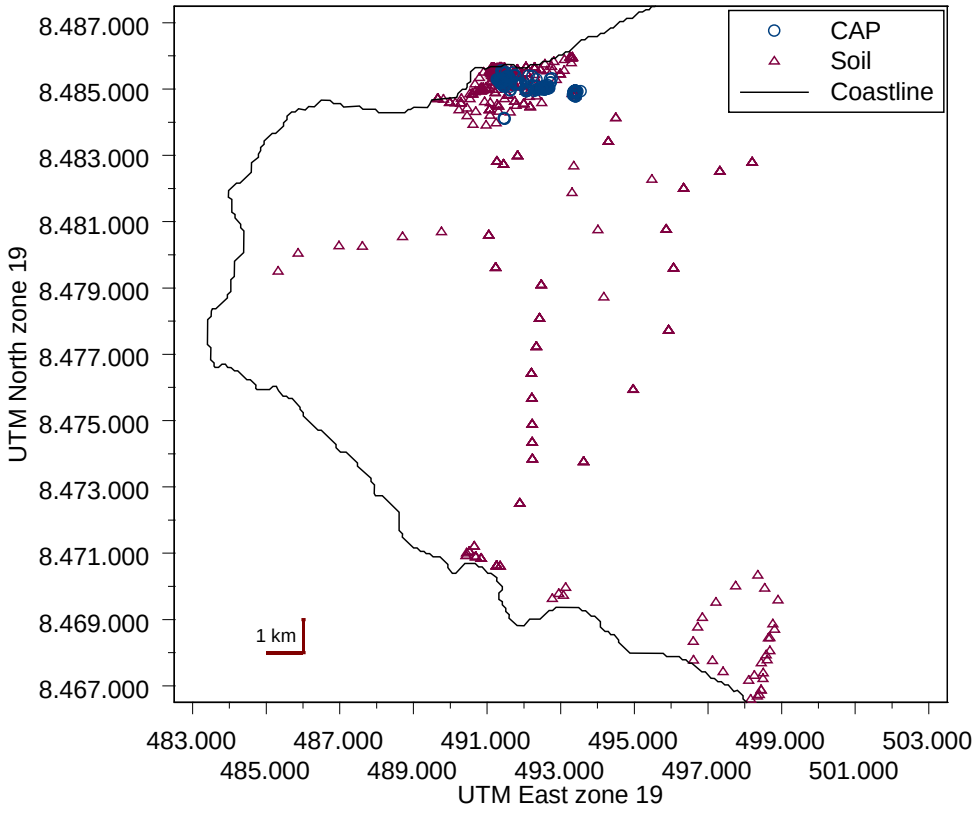


Figure 7 Data points in Region 1.

The locations in region 1 is bound by the following coordinates

East (UTM zone 19)	Maximum	500000
	Minimum	483000
North (UTM zone 19)	Maximum	8490000
	Minimum	8465000

The minimum point (483000, 8465000) is subtracted from the coordinates in the spatial statistical presentation.

3.2 Descriptive statistics

A summary of the measurement is shown inTable 2.

Region 1	Minimum	25% quantile	Mean	Median	75% quantile	Maximum
^{241}Am (Bqm^{-2})	1	29	7.0×10^{3}	2.1×10^{2}	7.3×10^{3}	2.8×10^{5}
$Log_{10}(^{241}Am)$	0.04*	1.4	2.6	2.3	3.9	5.4
Distances (m)	0.16	747	4643	1741	7497	20308

Table 2 Summary statistics, region 1

*0.1 was added to the ^{241}Am concentrations to obtain numerical stability

A summary plot of the data in section 1 is shown in Figure 8

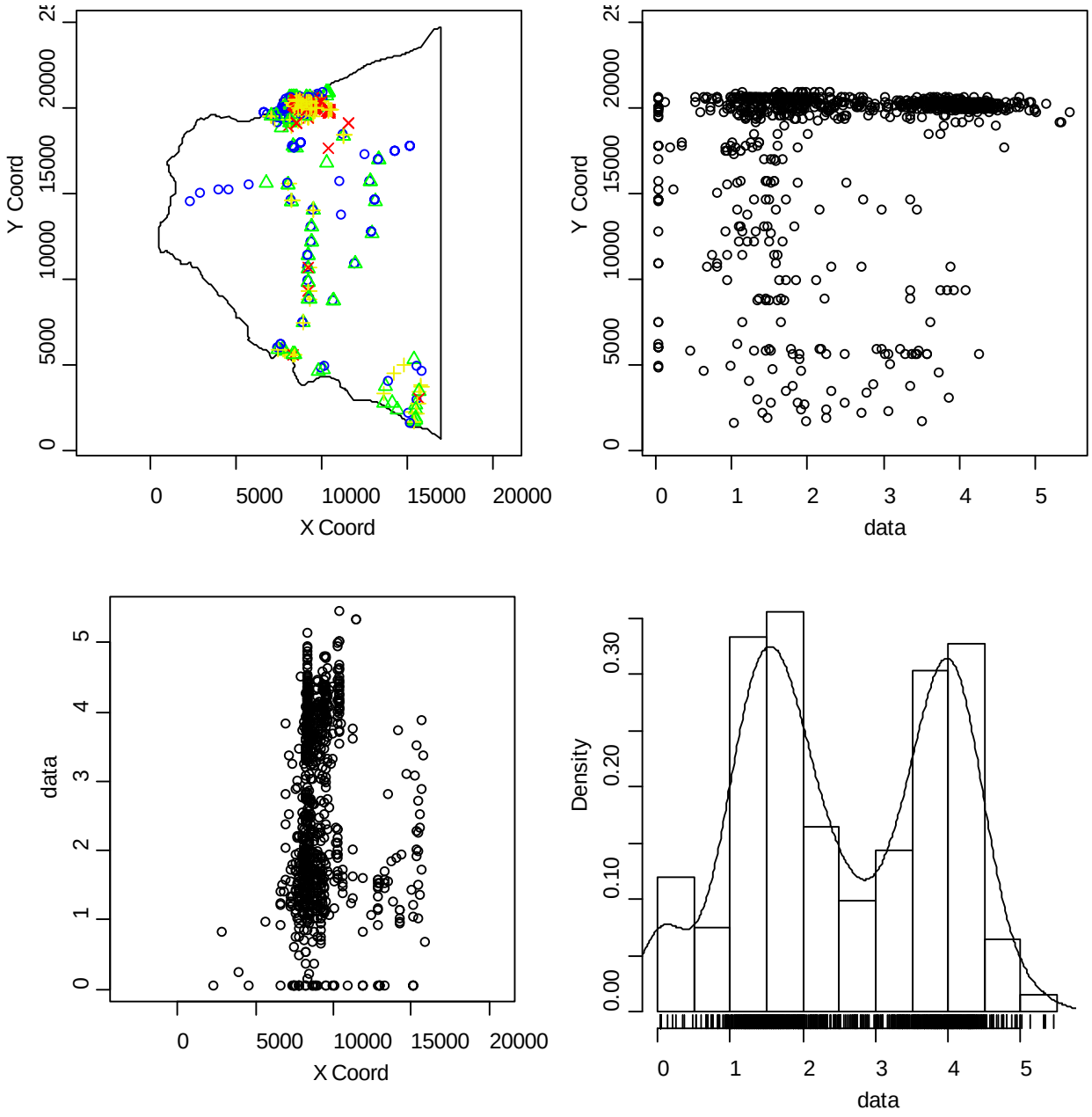


Figure 8 Summary plot of the $\log_{10}(^{241}Am)$ data in region 1. Upper left: Blue circle is the 0-25% quantiles, green triangle is the 25-50% quantiles, yellow plus is the 50-75% quantiles and red cross is the 75-100% quantiles.

In Figure 8 it can be seen that high concentrations also exist outside the coastal area. The histogram shows a bi-modular distribution. The bi-modal data distribution might reflect the sampling strategy, where the intension was to sample inside as well as outside hot-spot areas. The bi-modal distribution does not affect the validity of a spatial statistical analysis.

3.3 Spatial correlation

Before doing the maximum likelihood estimation an empirical variogram is calculated. The calculated values and the fitted spherical variogram function is shown in Figure 9

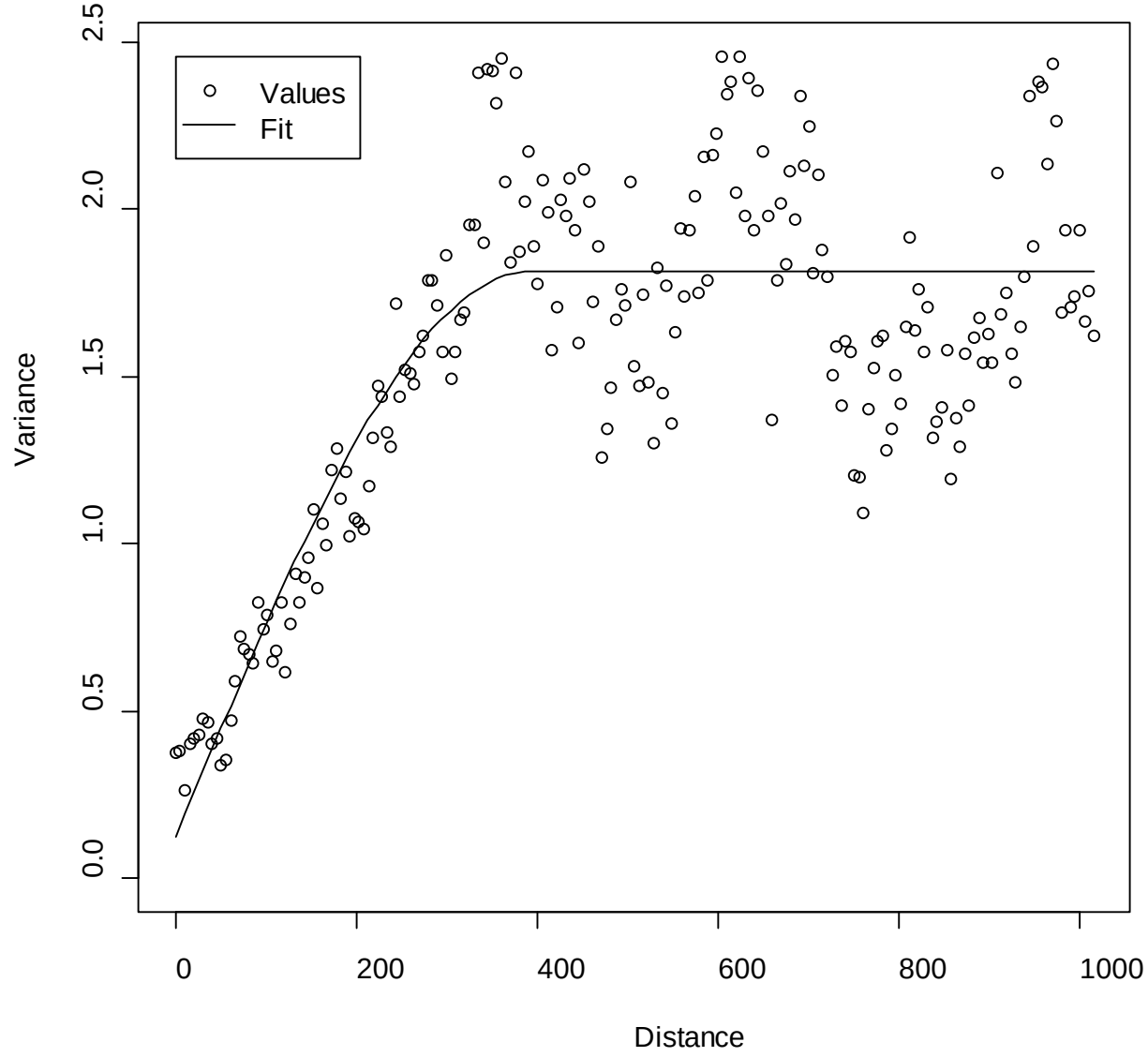


Figure 9 Variogram for $log_{10}(^{241}Am)$ in Region 1. Distances are divided with bins of approximately 5 m.

A very clear spatial correlation is seen. The parameters in the fitted spherical variogram function is shown in Table 3

Variogram parameter	Nugget	Sill	Partial Sill	Range (m)
Estimated value	0.13	1.81	1.69	391

Table 3 Parameter values from nonlinear weighted regression of the variogram, region 1.

So, preliminary it is assumed that data is correlated within a range of approximately 400 m.

3.4 ML-estimation and prediction of contamination level

The spatial maximum likelihood (ML) estimation of the three models resulted in the correlation structure shown in Table 4

Model	Spatial parameters				Box-cox transformation	Trend	Anisotropy		Information criteria	
	Nugget	Sill	Partial Sill	Range (m)	Lambda	Beta0, BetaX, BetaY	Angle	Ratio	AIC	BIC
Basic model	**0.437**	**1.44**	**1.01**	**380**	**0.876**	**0.899**			**2171**	**2194**
+Anisotropy	0.412	1,38	0.965	250	0.879	0.979	2.59	1.98	2166	2199
+trend	0.437	1,40	0.965	374	0.875	-0.556 8.7×10^{-5} 4.0×10^{-5}			2169	2207

Table 4 Parameter values from the maximum likelihood estimation in region 1.

Based on the AIC a model with anisotropy is the most optimal (smallest AIC). Based on the BIC the basic model without trend or anisotropy is optimal (smallest BIC). The two model selection criteria disagree.

In region 1 large areas are sparsely sampled, therefore it is decided to use the basic model without an overall spatial trend and without anisotropy (marked in bold in the Table 4 above).

Based on the ML-estimation data is correlated within a distance of 380 m. The correlation is not depending on direction. This implies that locations further away than 380 m from a measurement location can not be predicted based on the measurements. The partial sill is approximately 2 times higher than the nugget.

The 90% and 95% confidence intervals for the spatial parameters are shown in

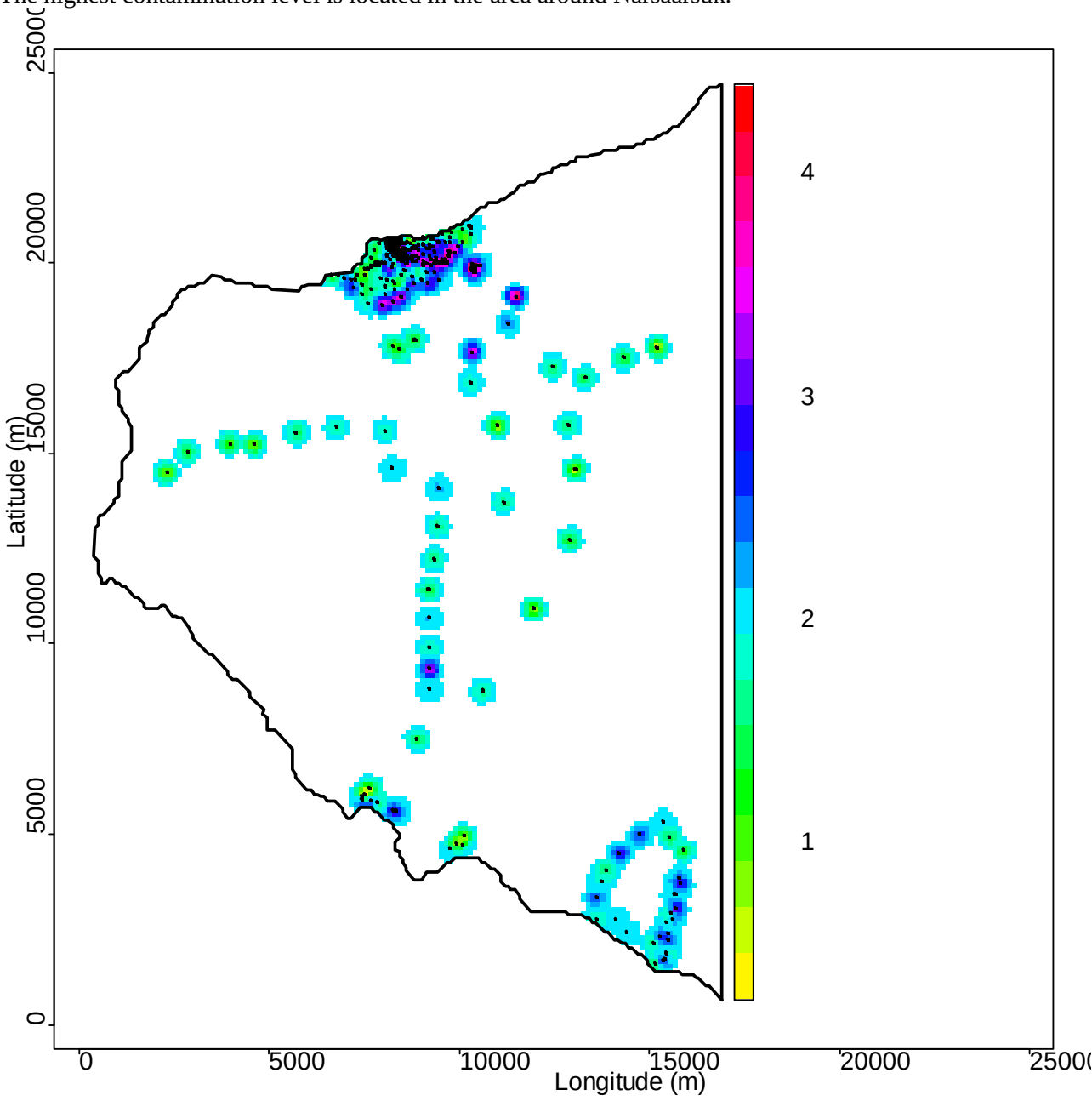
	ML-estimate and 95% profile likelihood confidence intervals		
	Nugget	Partial sill	Range
Estimate	0.44	1.0	380
90% CI	[0.35;0.52]	[0.98;1.0]	[340;>1000]
95%CI	[0.37;0.54]	[0.99;1.0]	[320;>1000]

Table 5 Maximum likelihood estimates and 90% and 95% profile likelihood confidence intervals for the spatial parameters in region 1

The spatial prediction on a 100m×100m grid of the $\log_{10}$(Am) concentrations in region 1 is shown in Figure 10. Locations further than 380 m from a measurement location are not predicted. The altitude is overlaid the predicted concentrations shown by height curves.

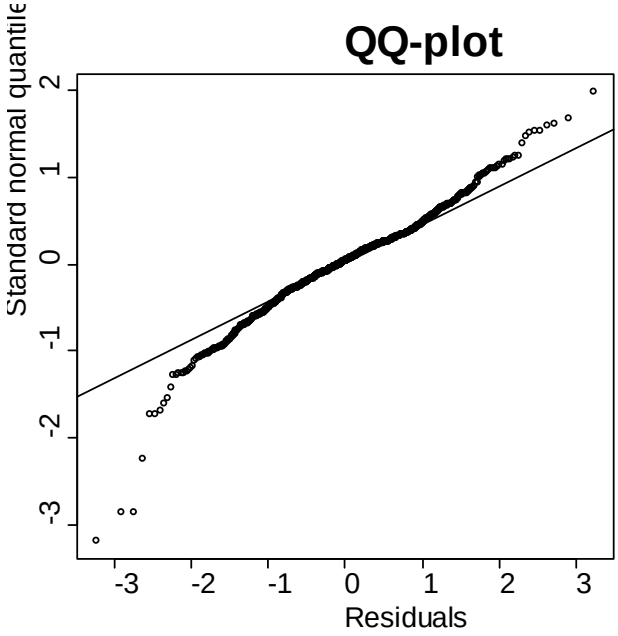
Based on the estimated range, it is evident that further sampling is needed to estimate the contamination level in the entire area.

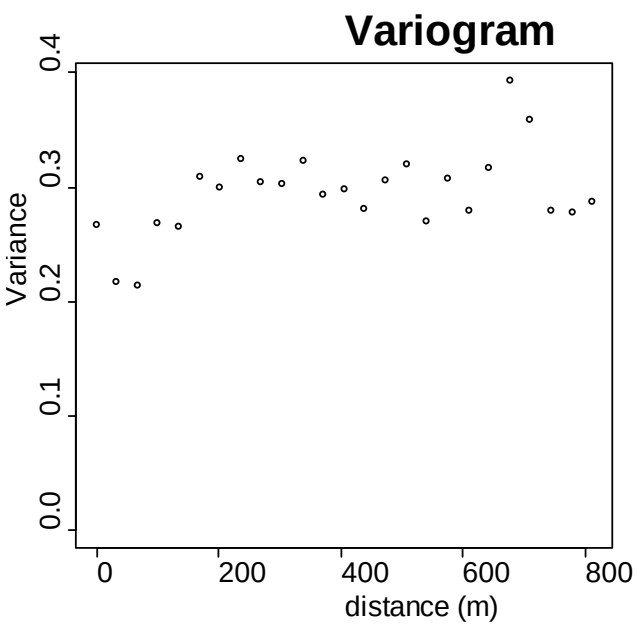
The highest contamination level is located in the area around Narsaarsuk.

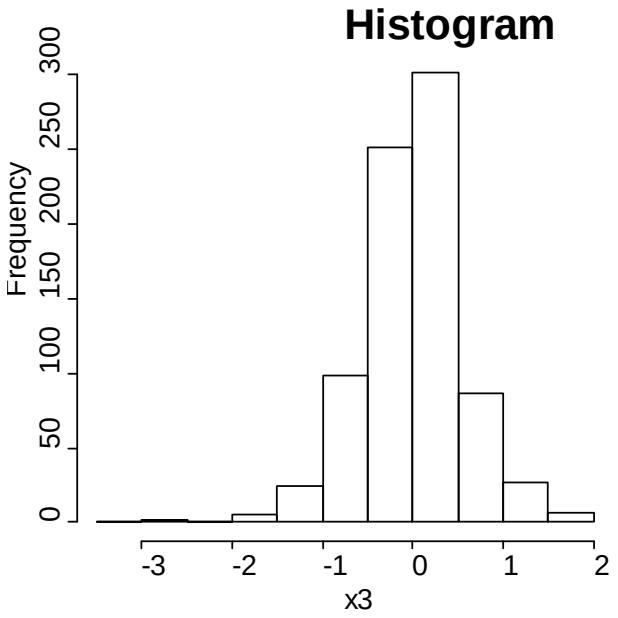


3.5 Validation of spatial maximum likelihood estimation

In Figure 11 the residuals from the maximum likelihood estimation is evaluated. The QQ-plot, Histogram and the Box-Whisker plot show approximately normal distributed data. The variogram show a small spatial dependency with data points very close having slightly less variance. However, with an estimated partial sill of 1 the small correlation in the residuals is not considered of major importance.

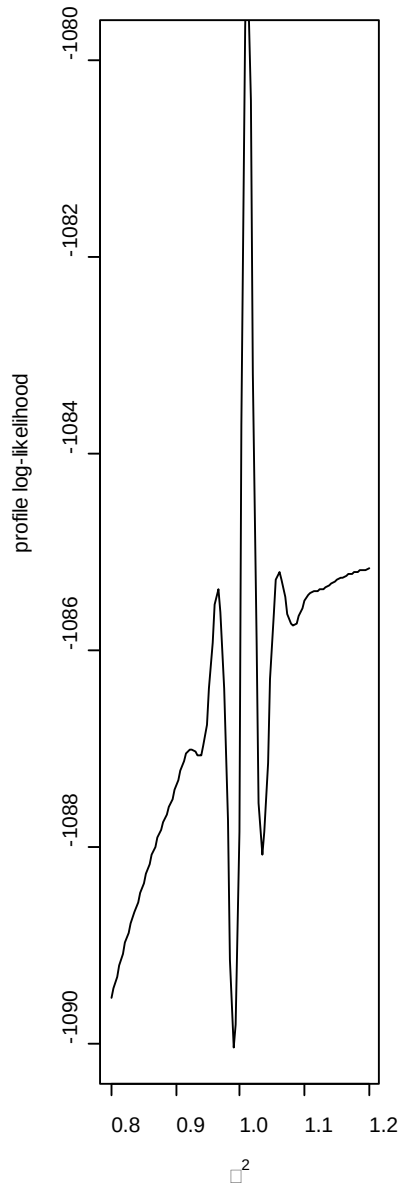






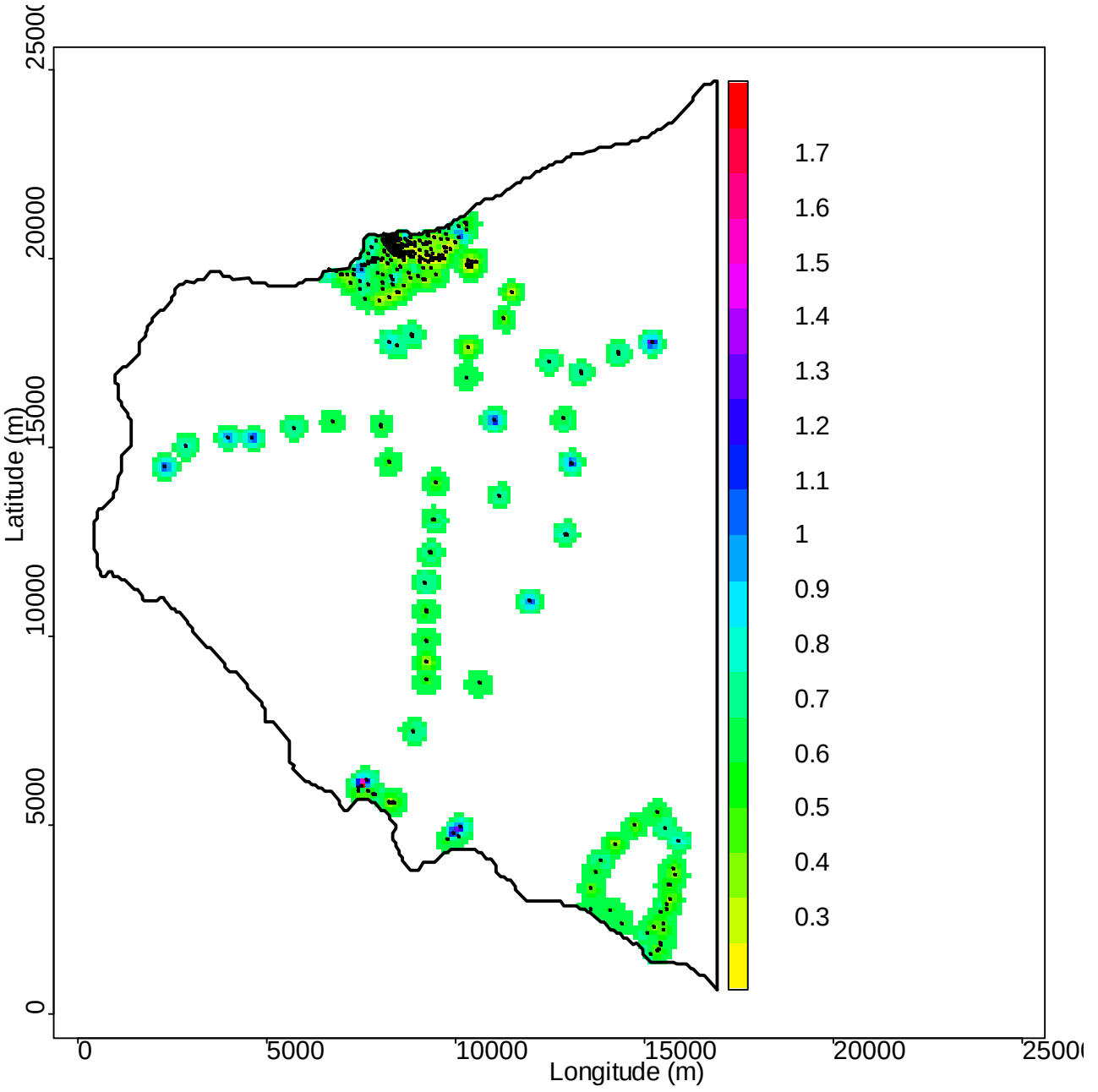
Profile-likelihood curves

In Figure 12 the profile likelihood is shown for the final basic model of Region 1. Only the nugget has a nice bell-shaped curve. The dotted horizontal lines indicate approximate 90% and 95% confidence intervals for each parameter.

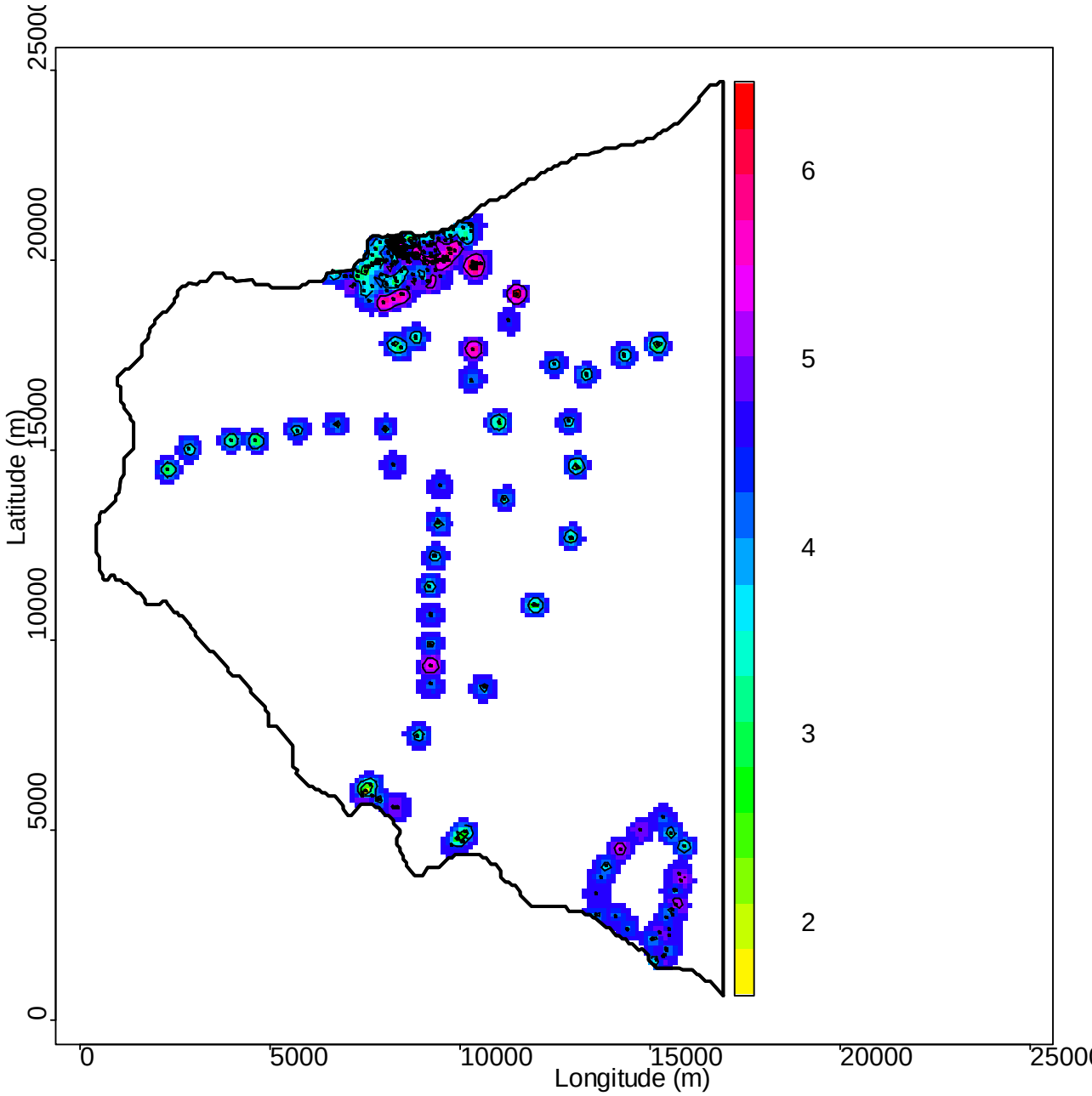


3.6 Uncertainty of predicted contamination level

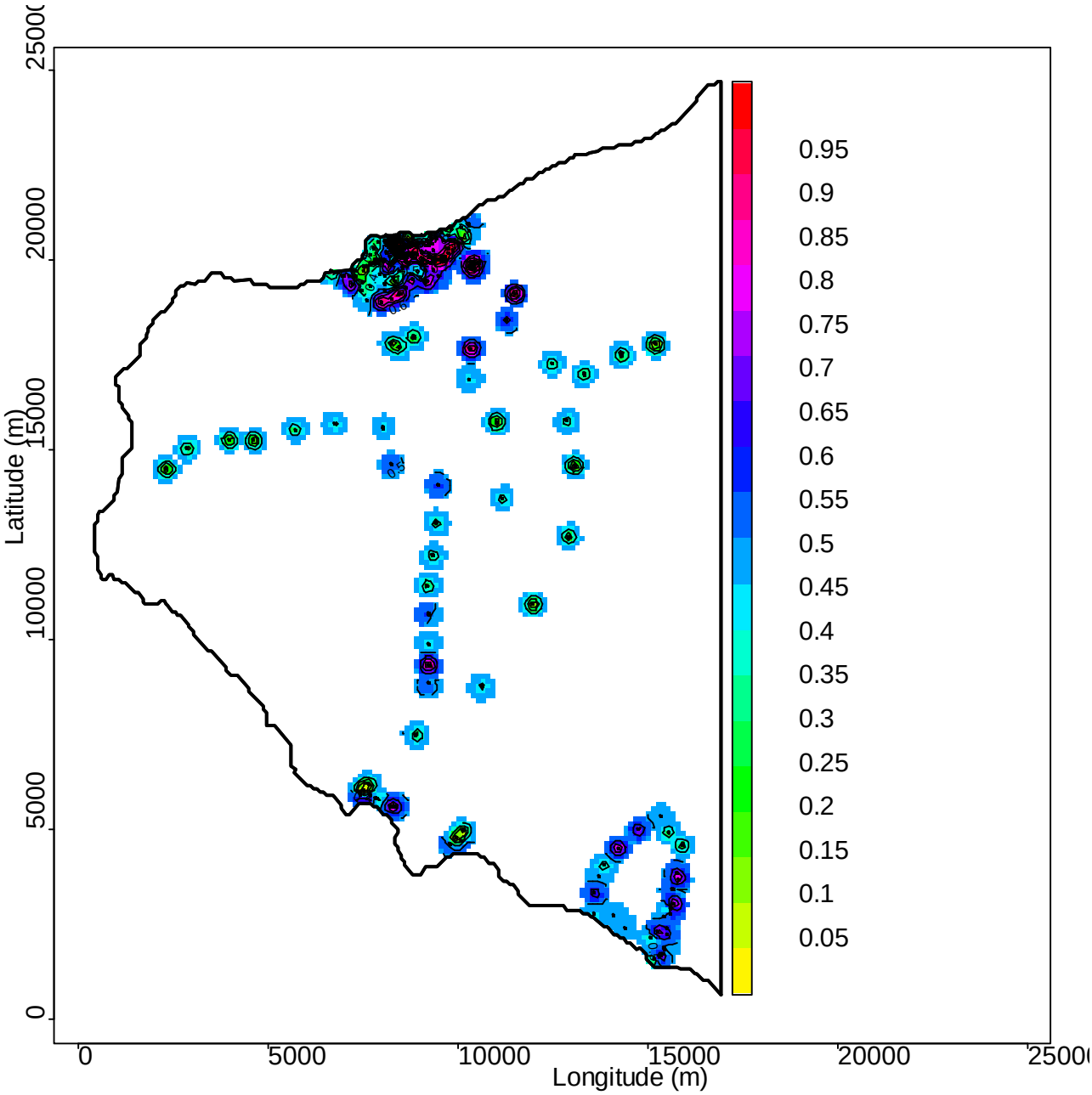
The relative standard error (standard error/predicted value) of the predicted concentrations are shown in Figure 13. The relative error is in general small in the more densely sampled areas. It does of course also depend on the local variation.

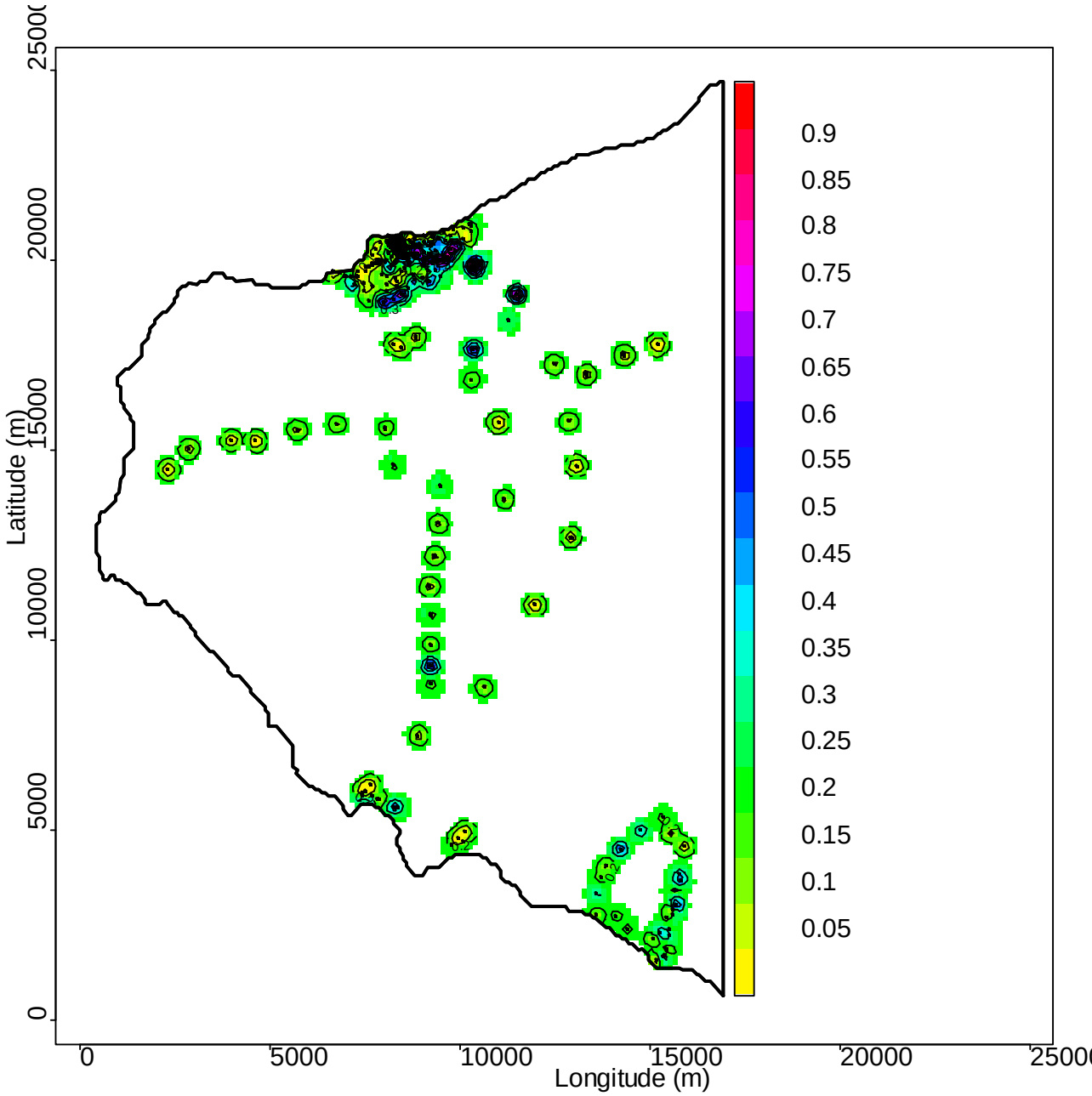


Based on the standard error of each predicted value the upper 95% confidence limit for the predicted values is shown in Figure 14. Notice that it is a point-wise 95% confidence interval, so it is not a 95% confidence interval for the surface as such. It can however be used to evaluate the upper limit for each grid-point



Based on normal probability assumptions the probability that the concentration exceeds a certain threshold can be calculated. In Figure 15 and Figure 16 the probability of the $\log_{10}$ concentration exceeding 2 and 3 respectively is shown.




Latitude (m)
0
5000
10000
15000
20000
25000
Longitude (m)
0
5000
10000
15000
20000
25000
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05

3.7 Conclusion region 1

A clear spatial correlation structure is present in region 1. The basic spatial model with nugget, partial sill and range as spatial parameters was selected.
Data are correlated up to a distance of approximately 380m (estimated *range*). The partial sill was estimated to be 1 and the nugget approximately 0.5, so the partial sill is 2 times the nugget.
Due to the size of the area and the sampling performed, large areas are further than 380m away from the nearest observation. This means that the information in the sampled points is of no use in these areas, resulting in the average concentration being the best guess. This most likely doesn't reflect reality so interpretation in these areas should not be done based on the presented data. Further sampling is needed to conclude anything in these areas.

In region 1 the maximum observed level of ^{241}Am was 2.8×10^{5} Bqm^{-2}. Most high levels were observed near Narsaarsuk. This area was also sampled most intensively. However, in Grønnedal the maximum observed level of ^{241}Am was 1.9×10^{4} Bqm^{-2}. Based on the topography of the area and the location of the hot-spots a hypothesis could be wind, snow and water transport particles to pools where the level therefore increase.

The area is so sparse sampled that it is of no use to try to predict the overall amount of ^{241}Am and ^{239,240}Pu. This is only done for region 2.

4 Analysis of data in Region 2

This section describes the analysis of data in region 2 (see Figure 4)

4.1 Geographical location

The locations in region 2 is bound by the following coordinates

East (UTM zone 19)	Maximum	494000
	Minimum	489500
North (UTM zone 19)	Maximum	8486000
	Minimum	8483000

The minimum point (489500, 8483000) is subtracted from the coordinates in the spatial statistical presentation. A total of 615 measurements were taken on 439 locations, 435 soil samples and 180 CAP measurements. The location is shown in Figure 17

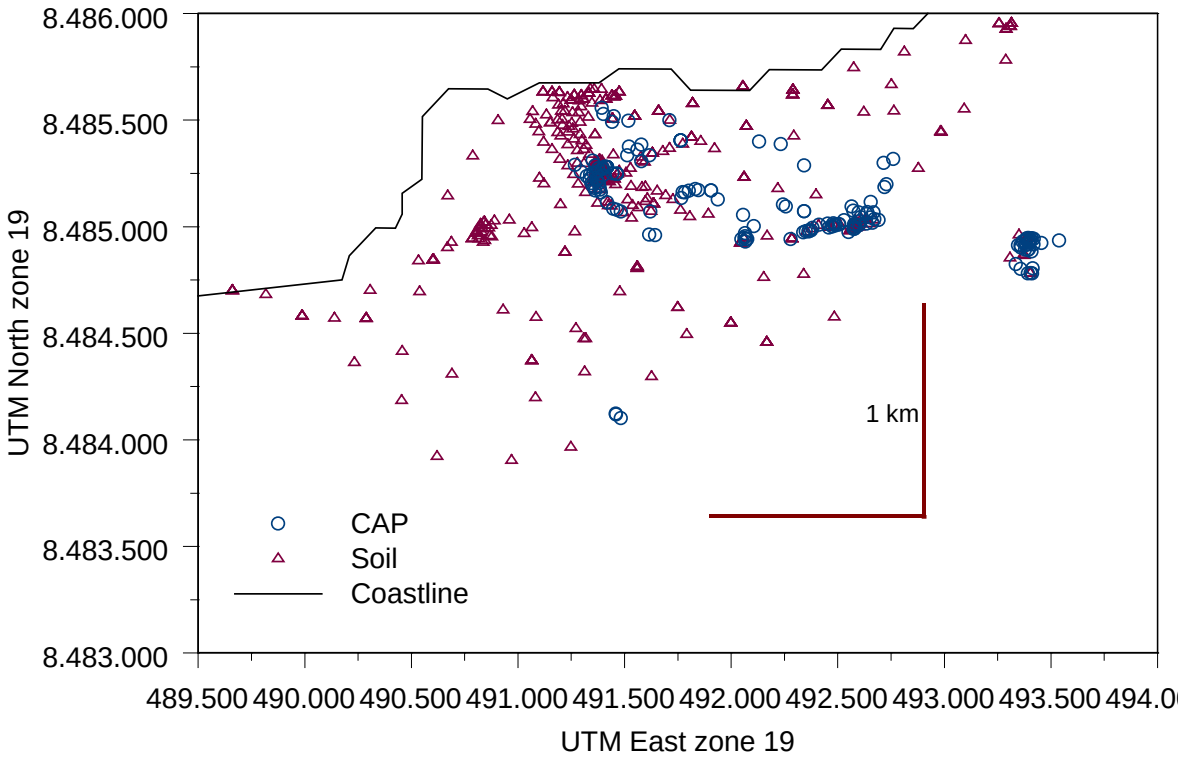


Figure 17 Location of sampling points in region 2

4.2 Descriptive statistics

A summary of the measurement is shown in Table 6

	Minimum	25% quantile	Mean	Median	75% quantile	Maximum
^{241}Am (Bqm^{-2})	1.1	48	8.1×10^3	1.2×10^3	1.0×10^4	2.8×10^5
$Log_{10}(^{241}Am)$	0.04*	1.7	2.8	3.1	4.0	5.4
Distances (m)	0.16	435	987	888	1370	3890

Table 6 Summary statistics, region 2

*0.1 was added to the concentrations to obtain numerical stability

All spatial statistics is performed on $\log_{10}$ transformed data. A summary plot is shown in Figure 18

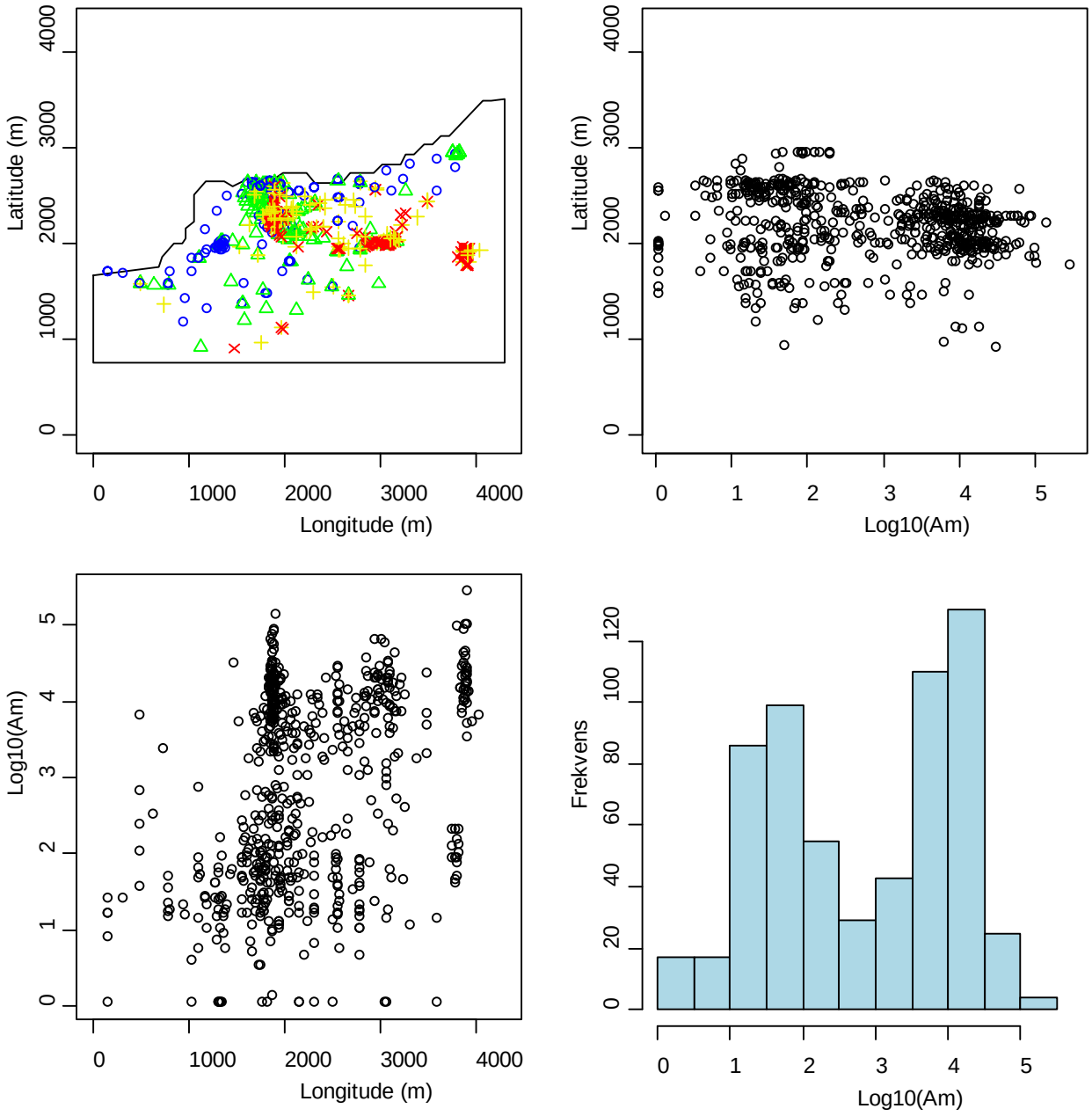


Figure 18 Summary plot of the $\log_{10}(^{241}Am)$ data in region 2. In the upper left: Blue circle is the 0-25% quantiles, green triangle is the 25-50% quantiles, yellow plus is the 50-75% quantiles and red cross is the 75-100% quantiles.

From Figure 18 it can be seen that high concentrations exist throughout the area. The histogram shows a bi-modular distribution. The bi-modal data distribution might reflect the sampling strategy, where the intension was to sample inside as well as outside hot-spot areas. The bi-modal distribution does not affect the validity of a spatial statistical analysis.

4.3 Spatial correlation

Before doing the maximum likelihood estimation a variogram is calculated. The calculated values and the fitted spherical variogram function is shown in Figure 19

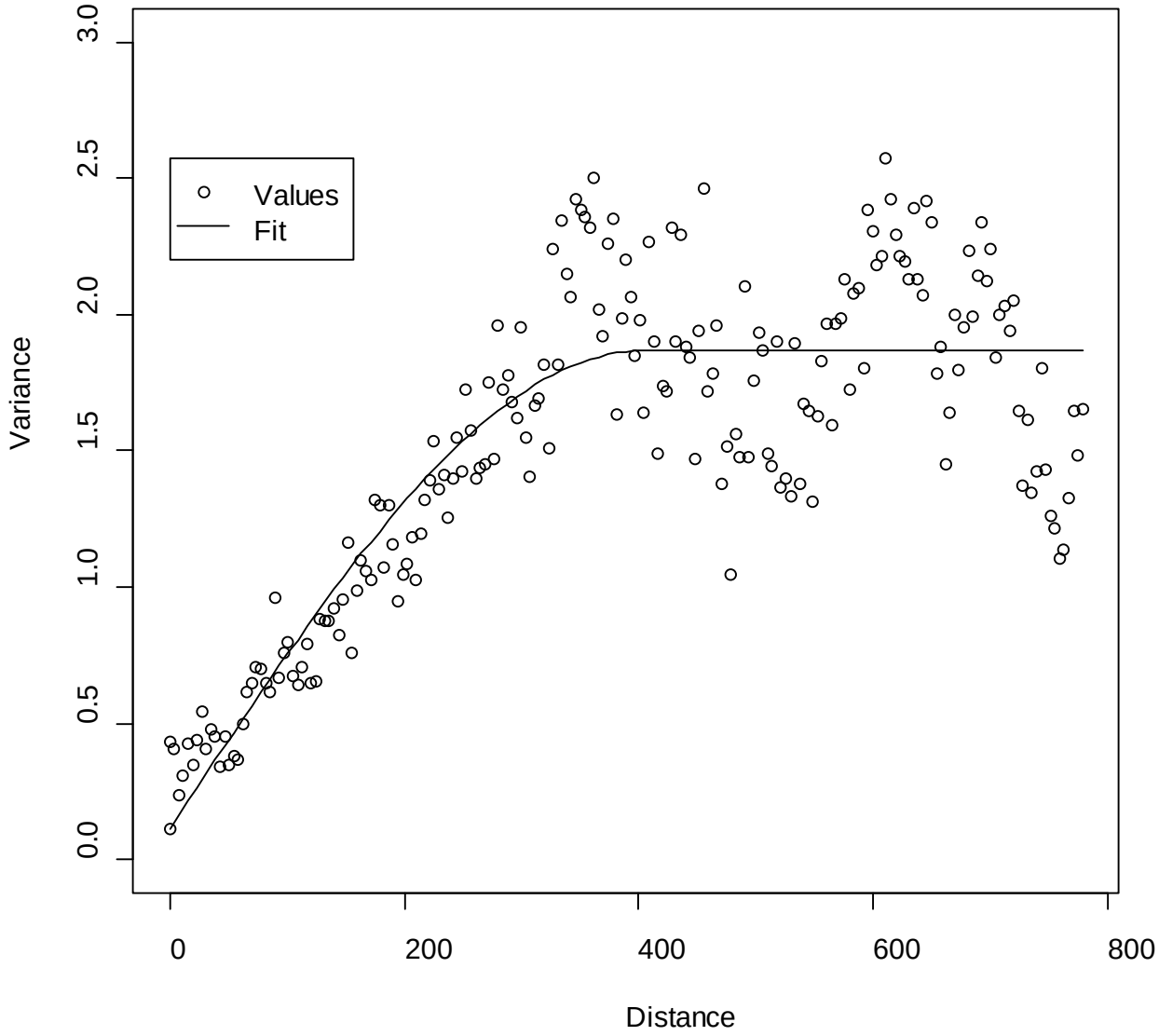


Figure 19 Variogram for log_{10}(Am) in Region 2

A very clear spatial correlation is seen. The parameters in the fitted spherical variogram function is shown in Table 7

Variogram parameter	Nugget	Sill	Partial Sill	Range (m)
Estimated value	0.11	1.87	1.75	402

Table 7 Parameter values from nonlinear weighted regression of the variogram, region 2

So, preliminary it is assumed that data is correlated within a range of approximately 400 m

4.4 Prediction of contamination level

As for region 1 the basic model and more complex models were investigated; 1) Including a 1st order trend and/or 2) including anisotropy. The spatial maximum likelihood estimation resulted in the correlation structure shown in Table 8.

Model	Spatial parameters				Box-cox transformation	Trend	Anisotropy		Information criteria	
	Nugget	Sill	Partial Sill	Range (m)	Lambda	Beta0, BetaX, BetaY	Angle	Ratio	AIC	BIC
Basic model	0.556	1.83	1.27	280	1.07	1.45			1645	1667
+Anisotropy	0.548	2.80	2.25	584	1.06	1.42	2.61	1.80	1641	1672
+trend	**0.556**	**1.49**	**0.938**	**198**	**1.08**	**1.70** 7.6×10^{-4} -9.0×10^{-4}			**1628**	**1659**

Table 8 Parameter values from the maximum likelihood estimation in region 2.

Based on the AIC and BIC a model with a first order trend is the most optimal. The two model selection criteria agree. Therefore it is decided to use the model with an overall spatial trend.

Based on the ML-estimation data is correlated within a distance of 198m. The correlation is not depending on direction. This implies that locations further away than 198 m from a measurement location can not be predicted based on the measurements. The partial sill is approximately 2 times higher than the nugget.

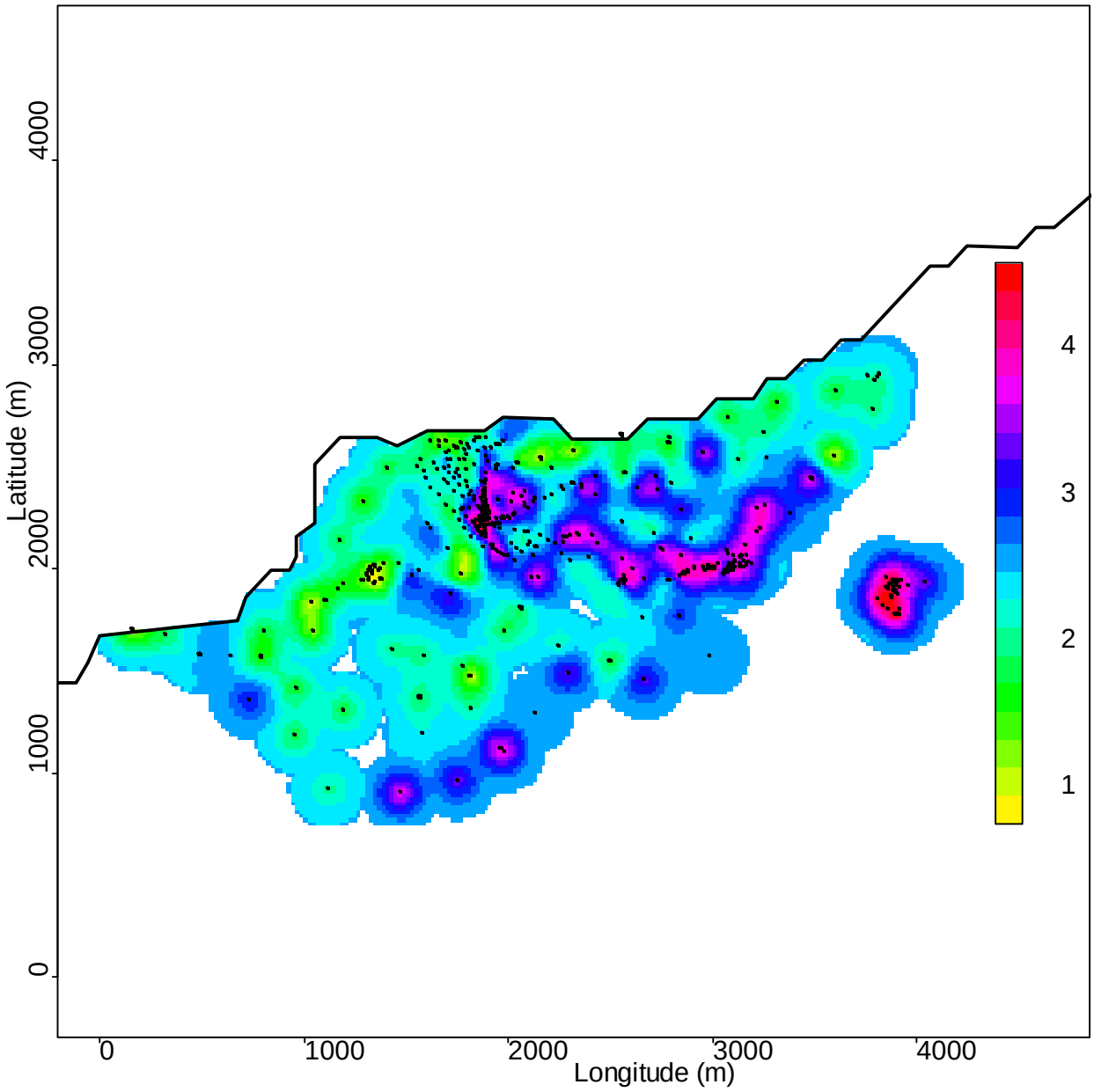
The 90% and 95% profile likelihood confidence intervals for the spatial parameters (nugget, partial sill and range) of the selected model are shown in Table 9

	ML-estimate and 95% profile likelihood confidence intervals		
	Nugget	Partial sill	Range
Estimate	0.56	0.94	198
90% CI	[0.45;0.69]	[0.76;0.95]	[109;292]
95%CI	[0.44;0.72]	[0.64;1.0]	[103;302]

Table 9 Maximum likelihood estimates and 90% and 95% profile likelihood confidence intervals for the spatial parameters in region 2.

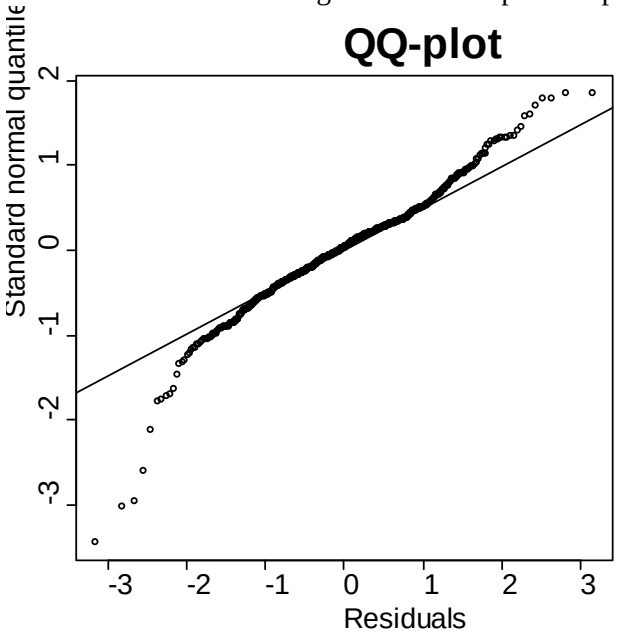
The spatial prediction on a 20m×20m grid of the $\log_{10}$(Am) concentrations in region 2 is shown in Figure 20 with height curves overlaid. Locations further than 198 m from a measurement location is not predicted.

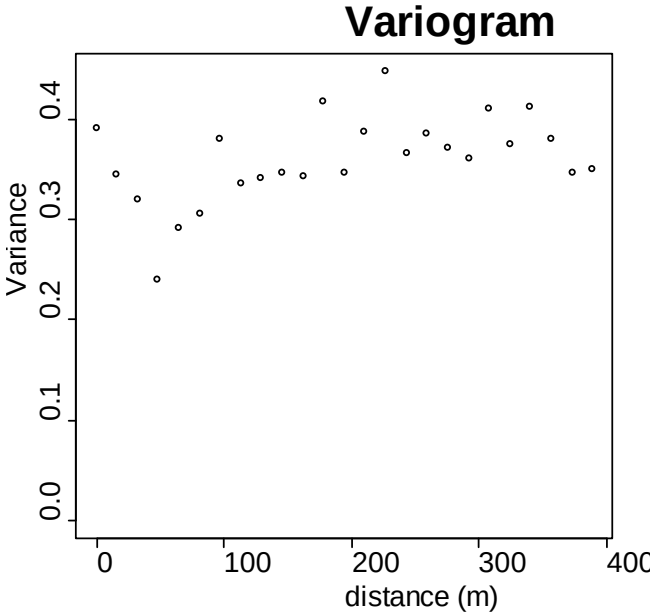
The 1st order trend parameters (Beta0, BetaX and BetaY) show that the Americium level increase in south-east direction

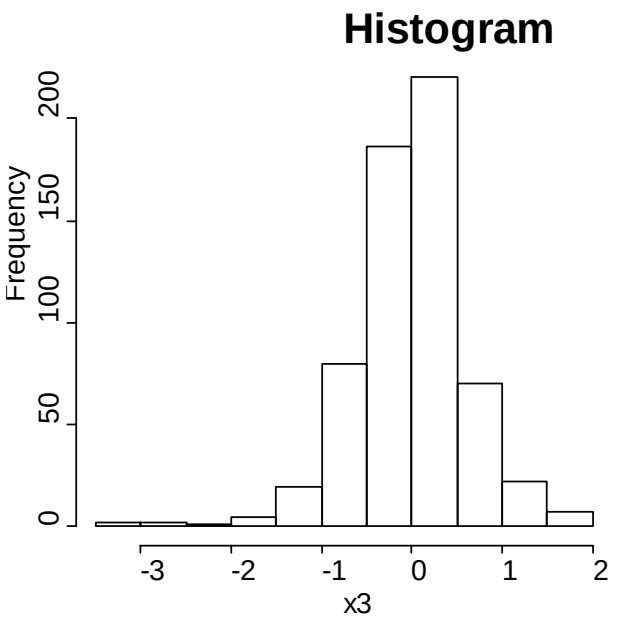

Latitude (m)
0
1000
2000
3000
4000
Longitude (m)
0
1000
2000
3000
4000
1
2
3
4

4.5 Validation of spatial maximum likelihood estimation

In Figure 21 the residuals from the maximum likelihood estimation is evaluated. The QQ-plot, Histogram and the Box-Whisker plot show approximately normal distributed data. The variogram show no spatial dependency.

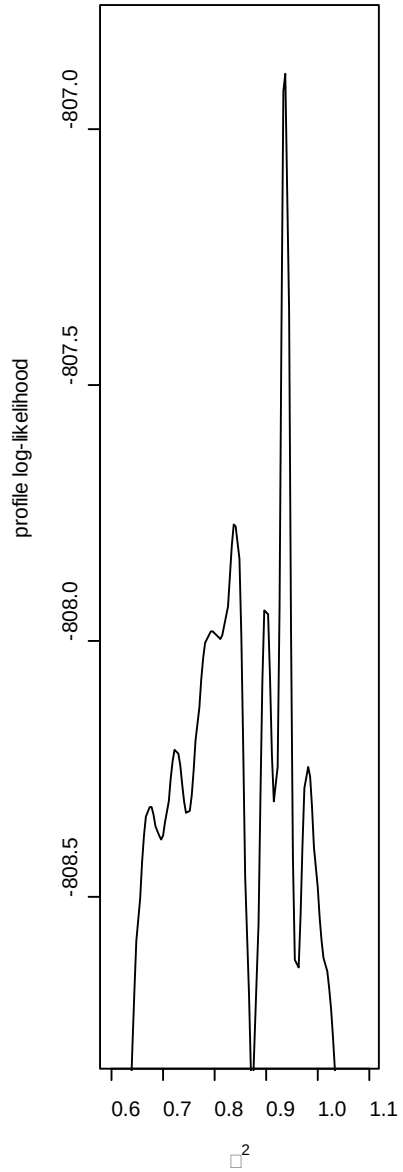






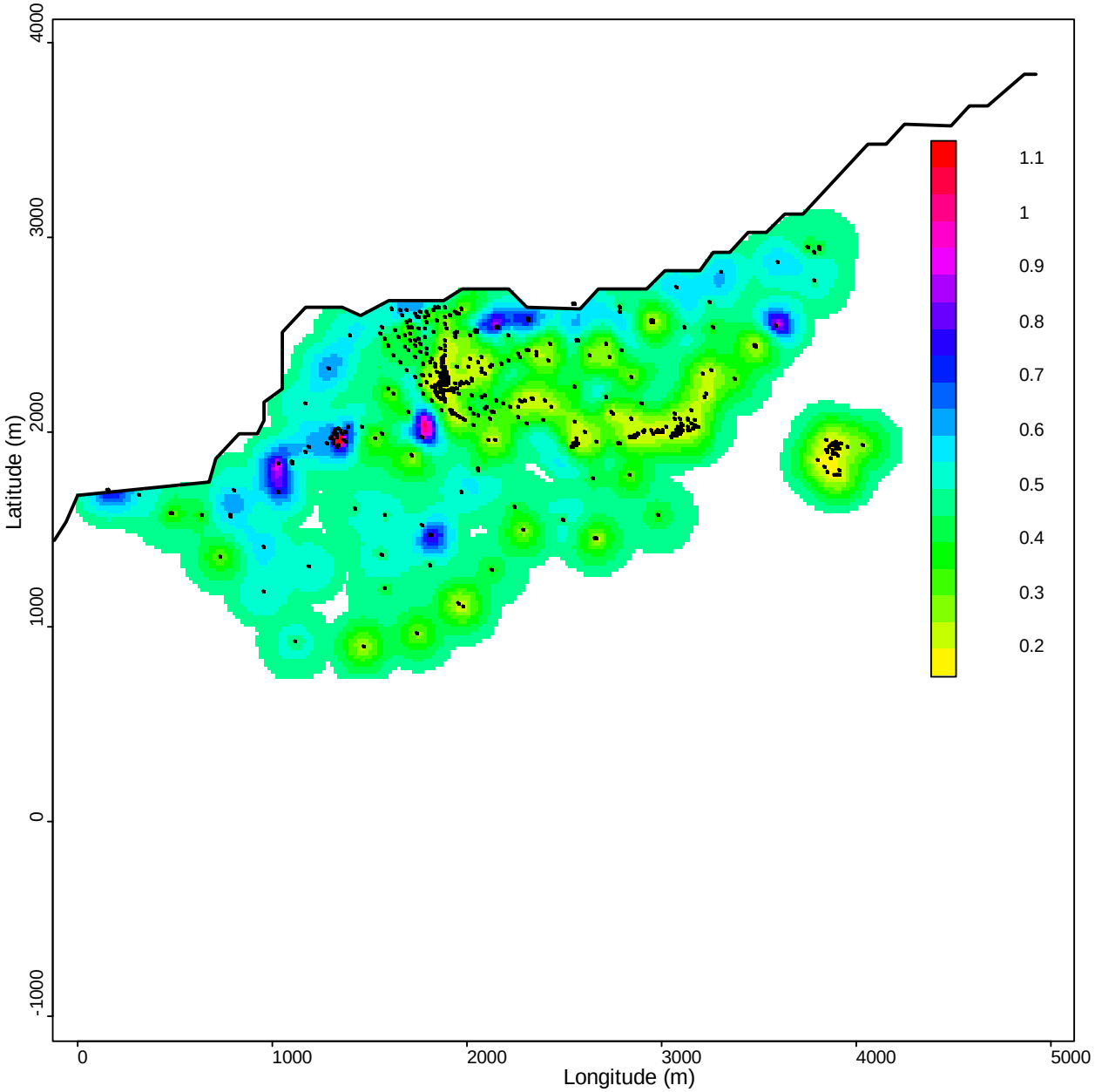
Profile-likelihood curves

In Figure 22 the profile likelihood is shown for the model with 1st order trend for region 2. The dotted horizontal lines indicate approximate 90% and 95% confidence intervals for each parameter.

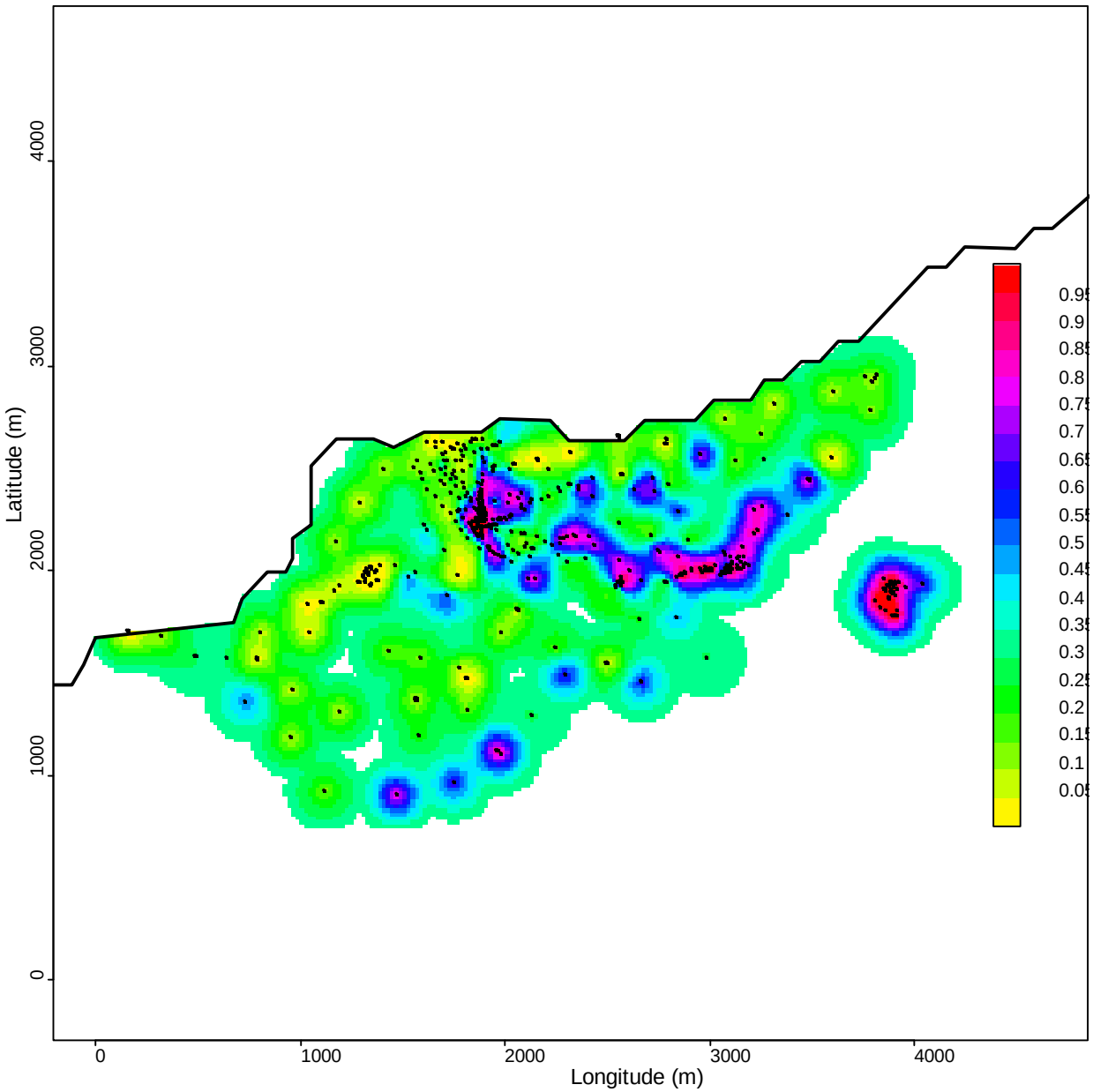


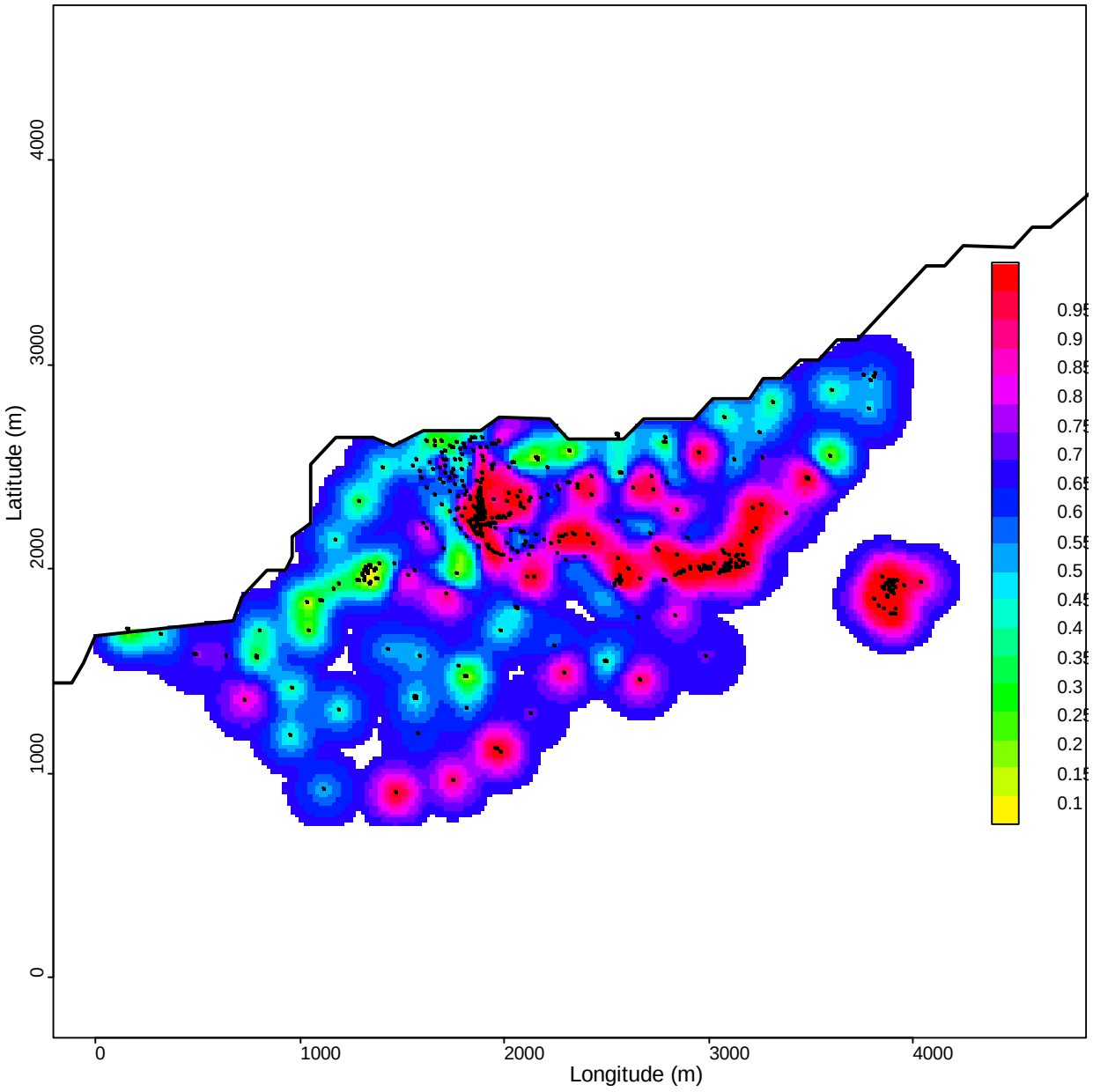
4.6 Uncertainty of predicted contamination level

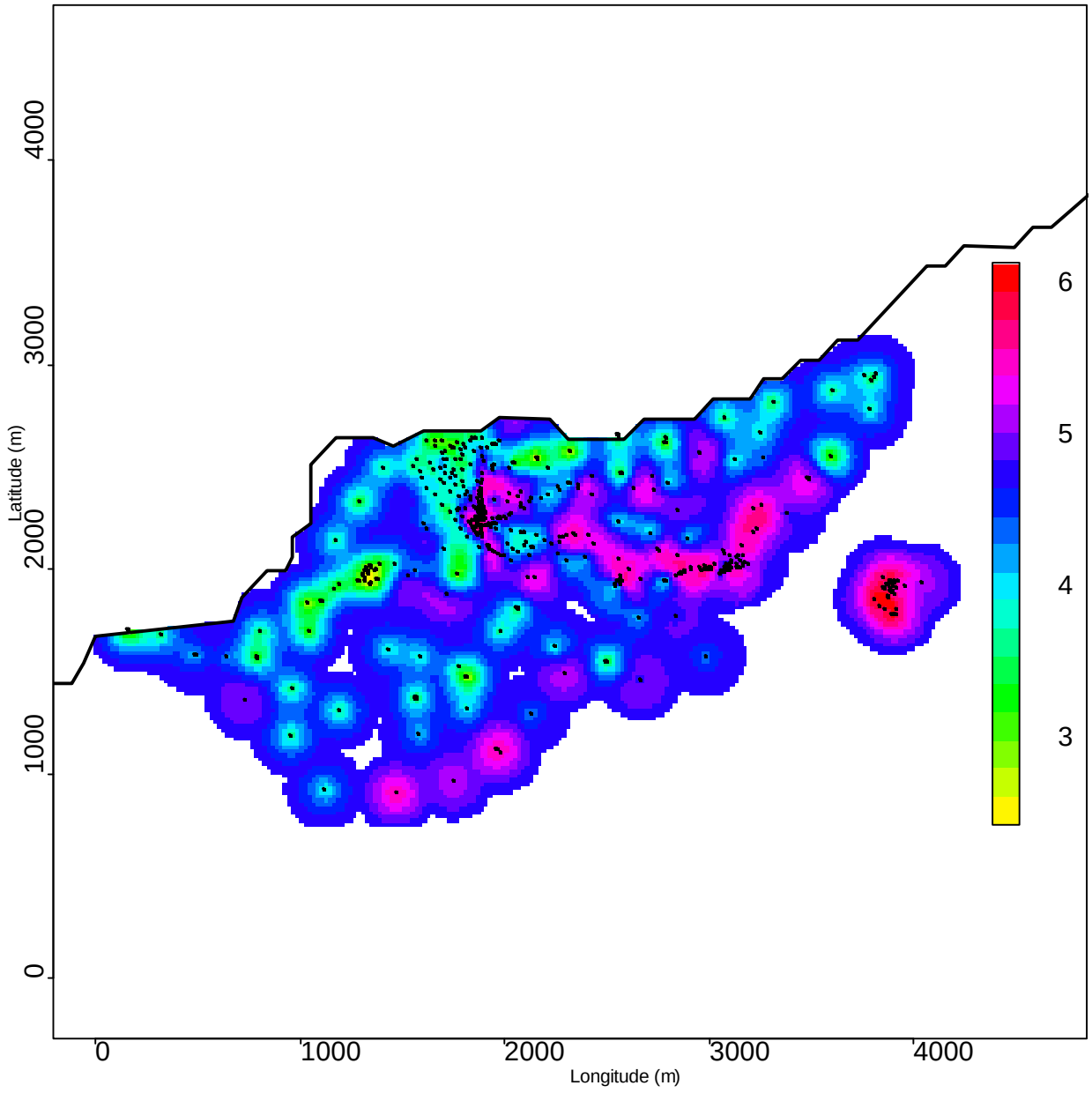
The relative standard error (standard error/predicted) of the predicted concentrations are shown in Figure 23. The relative error is in general small in the more densely sampled areas. It does of course also depend on the local variation.



Based on normal probability assumptions the probability that the concentration exceeds a certain threshold can be calculated. In Figure 24 and Figure 25 the probability of the $\log_{10}(^{241}\mathrm{Am})$ concentration exceeding 3 and 2 respectively is shown.




Latitude (m)
0
1000
2000
3000
4000
Longitude (m)
0
1000
2000
3000
4000
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1


4000
3000
2000
1000
0
Latitude (m)
0
1000
2000
3000
4000
Longitude (m)
6
5
4
3

4.7 Prediction of ^{239,240}Pu

Based on the predicted surface of $\log_{10}(^{241}\text{Am})$ the level of ^{241}Am and ^{239,240}Pu is calculated as described in section 2.5. The results for the three models are listed in the Table 10, with the optimal model including trend in bold.

Model	Total Area (km^2)	Area within *range* from nearest neighbour (km^2)	Proportion of total area within *range*	Total ^{241}Am GBq ($Bq\times10^9$) [95% CI]	Total ^{239,240}Pu GBq ($Bq\times10^9$) [95% CI]
Basic model	8.00	5.64	71%	87[63;143]	518[377;856]
+Aniso	8.00	7.24	91%	569[287;1476]	3410[1717;8854]
+ Trend	**8.00**	**4.93**	**61%**	**45[36;65]**	**270[212;387]**

Table 10 Predicted level of ^{241}Am and ^{239,240}Pu in region 2. Confidence intervals are found by simulation.

As can be seen in Table 10, the area where the level of ^{241}Am and ^{239,240}Pu is predicted depends on the model or more specifically on the estimated range of the model.

The higher the estimated range the larger area is included. Furthermore, measurements with high levels have influence on a larger area in models with larger range.

For the optimal model 61% corresponding to 4.93 km^2 of the total 8.00km^2 were included and the predicted total amount of ^{241}Am is 45 GBq ($Bq\times10^9$). The predicted total amount of ^{239,240}Pu is 270 GBq ($Bq\times10^9$).

Based on simulations, the confidence intervals for the total level of ^{241}Am and ^{239,240}Pu is Table 11

	^{241}Am [95%CI] (GBq)	^{239,240}Pu [95%CI] (GBq)
Simulations based on the estimated trend model.	[36;65]	[212;387]
Simulations as above and in addition including uncertainty of spatial parameters	[24;82]	[140;490]

Table 11 Confidence intervals of the total level of ^{241}Am and ^{239,240}Pu, based on simulation in region 2.

If locations with a distance to nearest measurement location larger than the estimated range is included in the prediction, the predicted total level of ^{241}Am and ^{239,240}Pu increase as described in the table below

Model	Total ^{241}Am TBq (Bq×10^{12})	Total ^{239,240}Pu TBq (Bq×10^{12})
Basic model	132	793
+Aniso	766	4594
+ Trend	73	442

Table 12 Predicted level of ^{241}Am and ^{239,240}Pu in region 2 including locations further than the estimated range away from the nearest measurement location.

4.8 Conclusion region 2

A clear spatial correlation structure is present in region 2. Data are correlated up to a distance of approximately 200m (estimated range=198m). A south-east trend was observed. Due to the size of the area and the sampling performed, some areas are further than 200m away from the nearest observation. In these areas the average concentration is best guess. This most likely doesn't reflect reality so interpretation in these areas should be done with care. The partial sill was estimated to be 1 and the nugget approximately 0.5, so the partial sill is 2 times the nugget.

Further sampling is needed to make more precisely conclusion about the concentrations in these areas.

Under the optimal spatial model, within the area of prediction, the predicted total amount of ^{241}Am is 45[95%CI: 36;65] GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 270[95%CI: 212;387] GBq (Bq×10^9).

Including the uncertainty of the estimated spatial parameters the confidence intervals gets wider, thus within the area of prediction, the predicted total amount of ^{241}Am is 45[95%CI: 24;82] GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 270[95%CI: 140;490] GBq (Bq×10^9).

Based on the topography of the area and the location of the hot-spots a hypothesis could be wind, snow and water transport particles to pools where the level therefore increase.

5 Analysis of a sub-region of Region 2

This section describes the analysis of data in a 1x1 km coastal sub-region of region 2

5.1 Geographical location

The locations in the sub-region is bound by the following coordinates

East (UTM zone 19)	Maximum	492000
	Minimum	491000
North (UTM zone 19)	Maximum	8486000
	Minimum	8485000

The minimum point (491000, 8485000) is subtracted from the coordinates in the spatial statistical presentation
A total of 312 measurements were taken on 215 locations, 225 soil samples and 87 CAP measurements. The location is shown in Figure 27

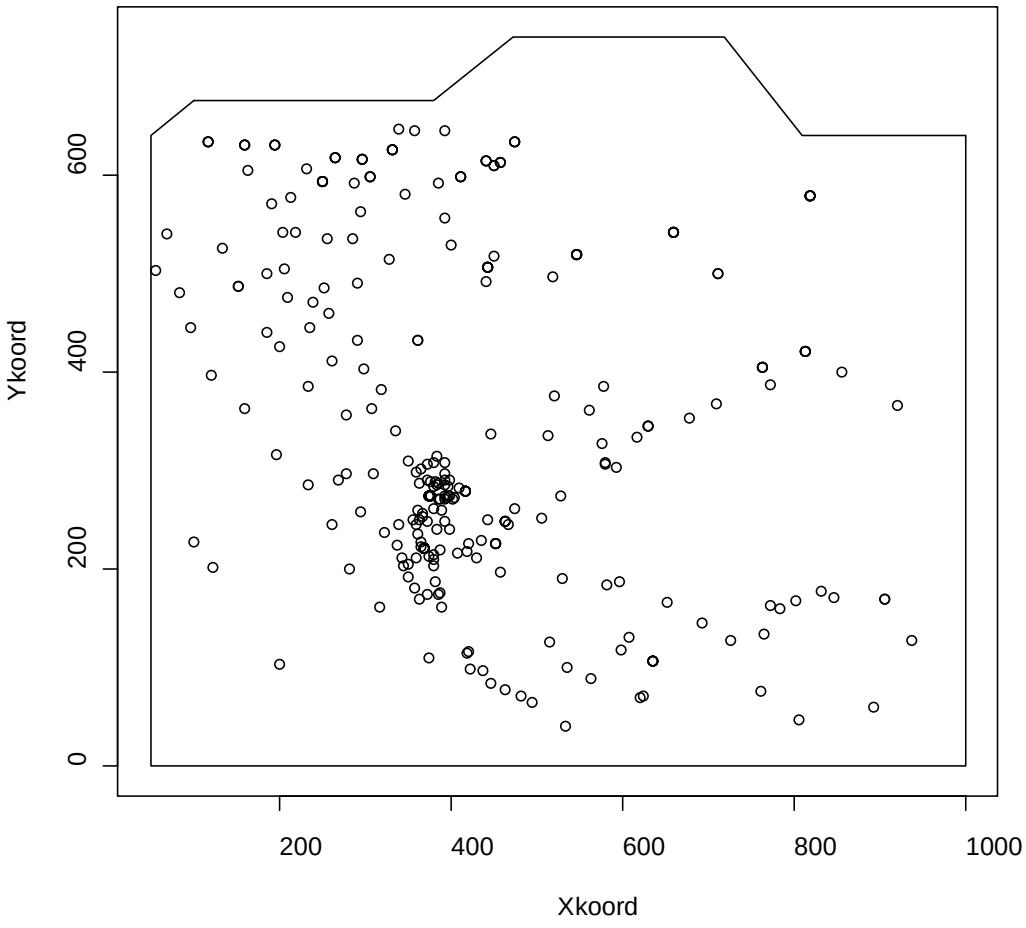


Figure 27 Location of a sub-region of Region 2

5.2 Descriptive statistics

A summary of the measurement is shown in Table 13.

	Minimum	25% quantile	Mean	Median	75% quantile	Maximum
Am (Bqm^{-2})	1	67	7.4×10^3	1.7×10^3	9.2×10^3	1.4×10^5
$Log_{10}Am$	0.04*	1.8	2.9	3.2	4.0	5.1
Distances (m)	0.16	160	301	302	418	965

Table 13 Summary statistics sub- region 2.

*0.1 was added to the concentrations to obtain numerical stability

All spatial statistics is performed on log_{10} transformed data. A summary plot is shown in Figure 28

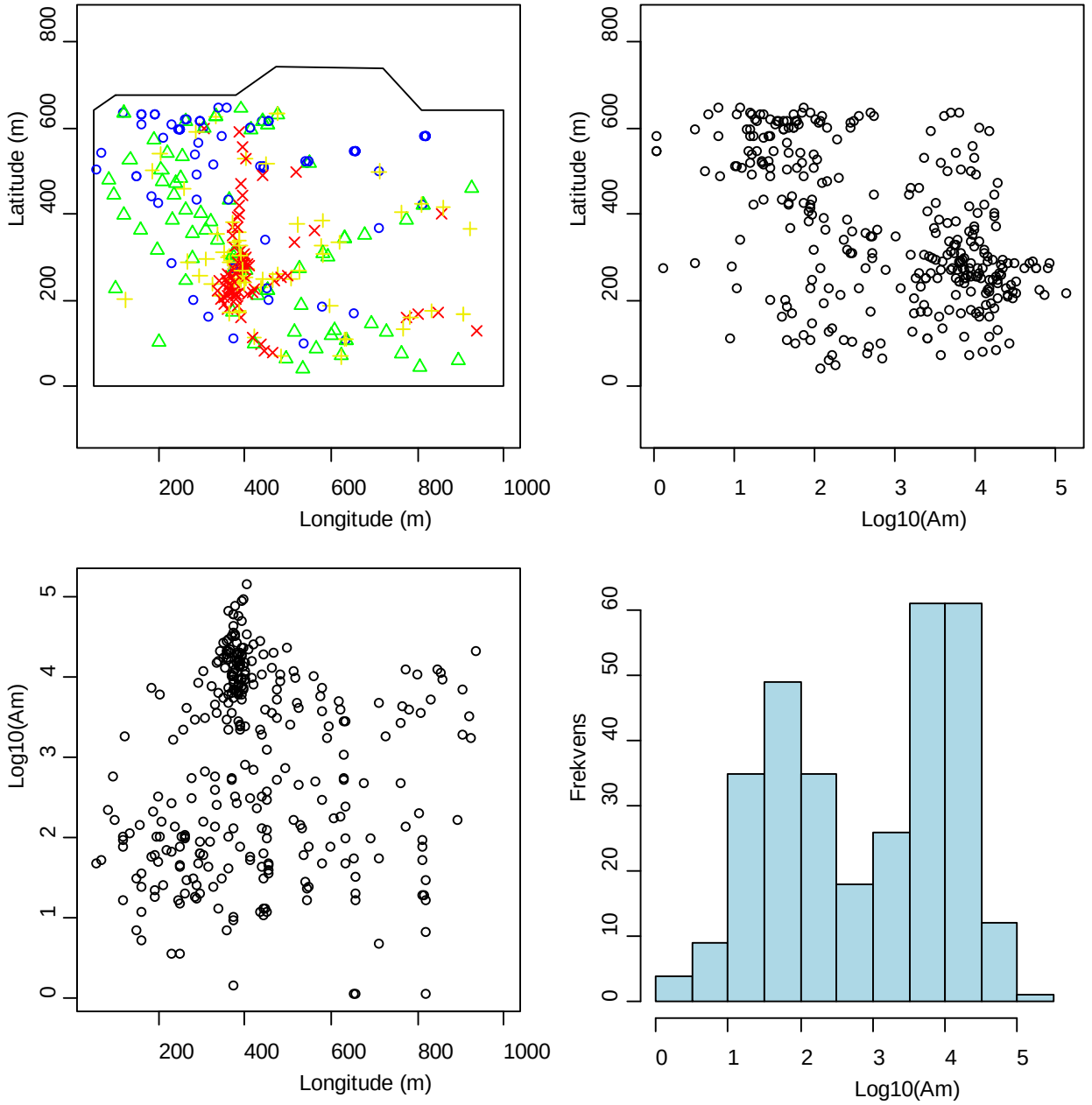


Figure 28 Summary plot of the log_{10}(Am) data in sub-region 2. In the upper left: Blue circle is the 0-25% quantiles, green triangle is the 25-50% quantiles, yellow plus is the 50-75% quantiles and red cross is the 75-100% quantiles.

In Figure 28 it can be seen that high concentrations is evenly spread, although there seems be one area with higher concentrations. The histogram shows a bi-modular distribution.

5.3 Spatial correlation

Before doing the maximum likelihood estimation a variogram is calculated. The calculated values and the fitted spherical variogram function is shown in Figure 29

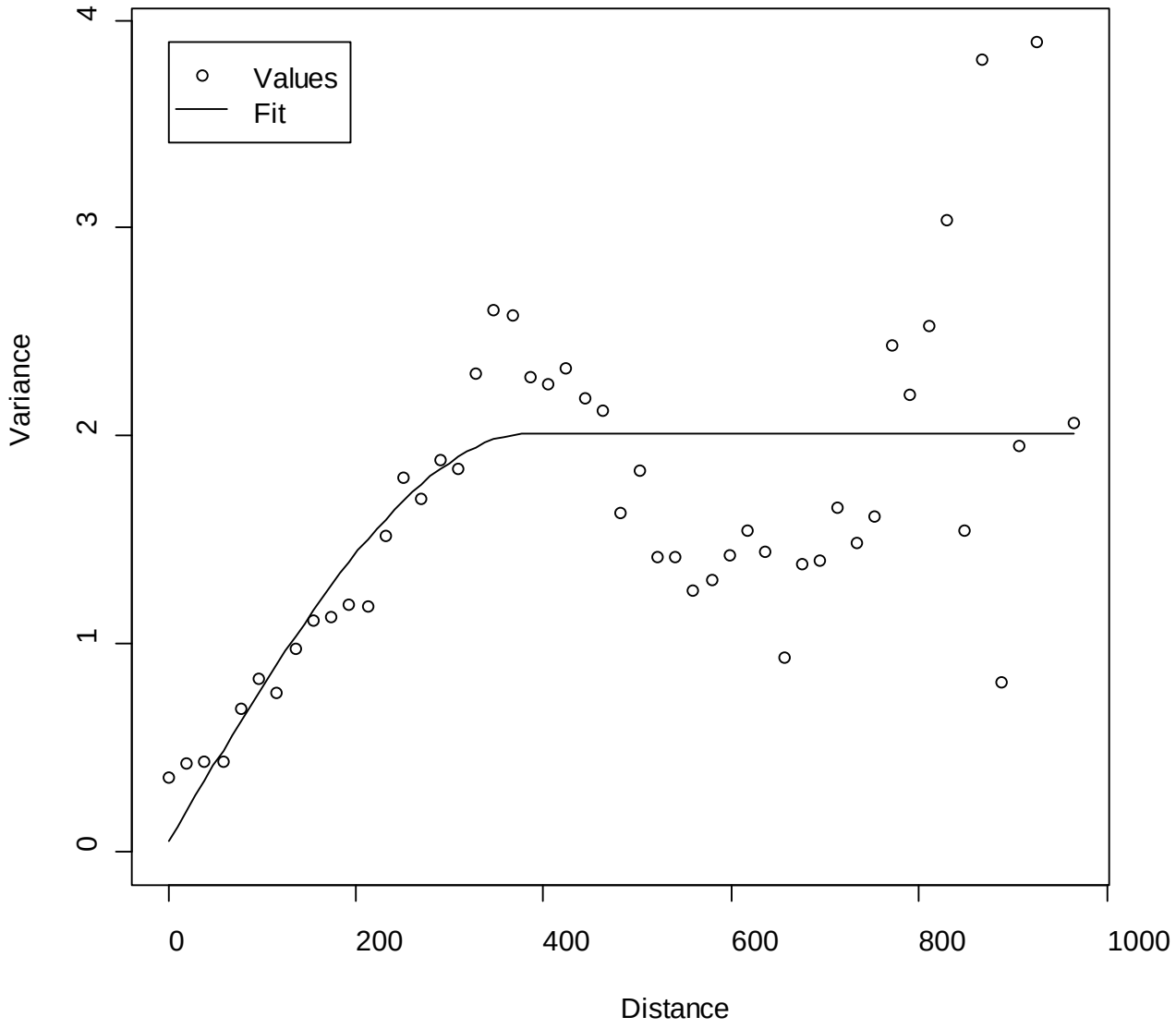


Figure 29 Variogram for log_{10}(Am) in the sub-region of Region 2

A very clear spatial correlation is seen. The parameters in the fitted spherical variogram function is shown in Table 14

Spatial parameter	Nugget	Sill	Partial Sill	Range

Estimated value	0.05	2.01	1.96	387m

Table 14 Parameter values from nonlinear weighted regression of the variogram, region 2

So, preliminary it is assumed that data is correlated within a range of approximately 400 m

5.4 Prediction of contamination level

As for region 1 and 2 the basic spatial model and more complex models were investigated: 1) Including a 1st order trend and/or 2) including anisotropy, meaning that the spatial correlation is independent on direction. The spatial maximum likelihood estimation resulted in the correlation structure shown in Table 15

Model	Nugget	Sill	Partial Sill	Range (m)	Lambda	Beta0, BetaX, BetaY	Angle	Ratio	AIC	BIC
Basic model	**0.5357**		**1.046**	**108**	**1.159**	**1.7601**			**803**	**822**
+Anisotropy*	0.5131	1.475	0.9615	90.5	1.146		2.763	2.033	801	827
+trend	0.5331	1.401	0.8678	98.9	1.148	2.0274 8.2×10^{-4} - 1.8×10^{-3}			798	825

Table 15 Maximum likelihood parameter estimates for spatial models, in sub-region 2

Based on the AIC a model with a first order trend is the most optimal. Based on the BIC the basic model without trend or anisotropy is optimal. The two model selection criteria disagree.

In this sub-area of region 2 the variability is assumed to be local, no overall trend is assumed to exist. Therefore it is decided to use the model without an overall spatial trend.

Based on the ML-estimation data is correlated within a distance of approximately 108 m. The correlation is not depending on direction. This implies that locations further away than 108 m from a measurement location can not be predicted based on the measurements. The partial sill is approximately 2 times higher than the nugget. The 90% and 95% profile likelihood confidence intervals for the spatial parameters (nugget, partial sill and range) of the selected model are shown in Table 16

	ML-estimate and 95% profile likelihood confidence intervals		
	Nugget	Partial sill	Range
Est	0.54	1.04	108
90% CI	[0.38;0.82]	[1.04;1.05]	[88;157]
95%CI	[0.36;0.85]	[1.04;1.05]	[86;207]

Table 16 Maximum likelihood estimates with 95% profile likelihood intervals in sub-region 2.

The spatial prediction on a on a 5m×5m grid of the $\log_{10}$(Am) concentrations in sub region 2 is shown in Figure 30
Locations further than 108 m from a measurement location is not predicted.

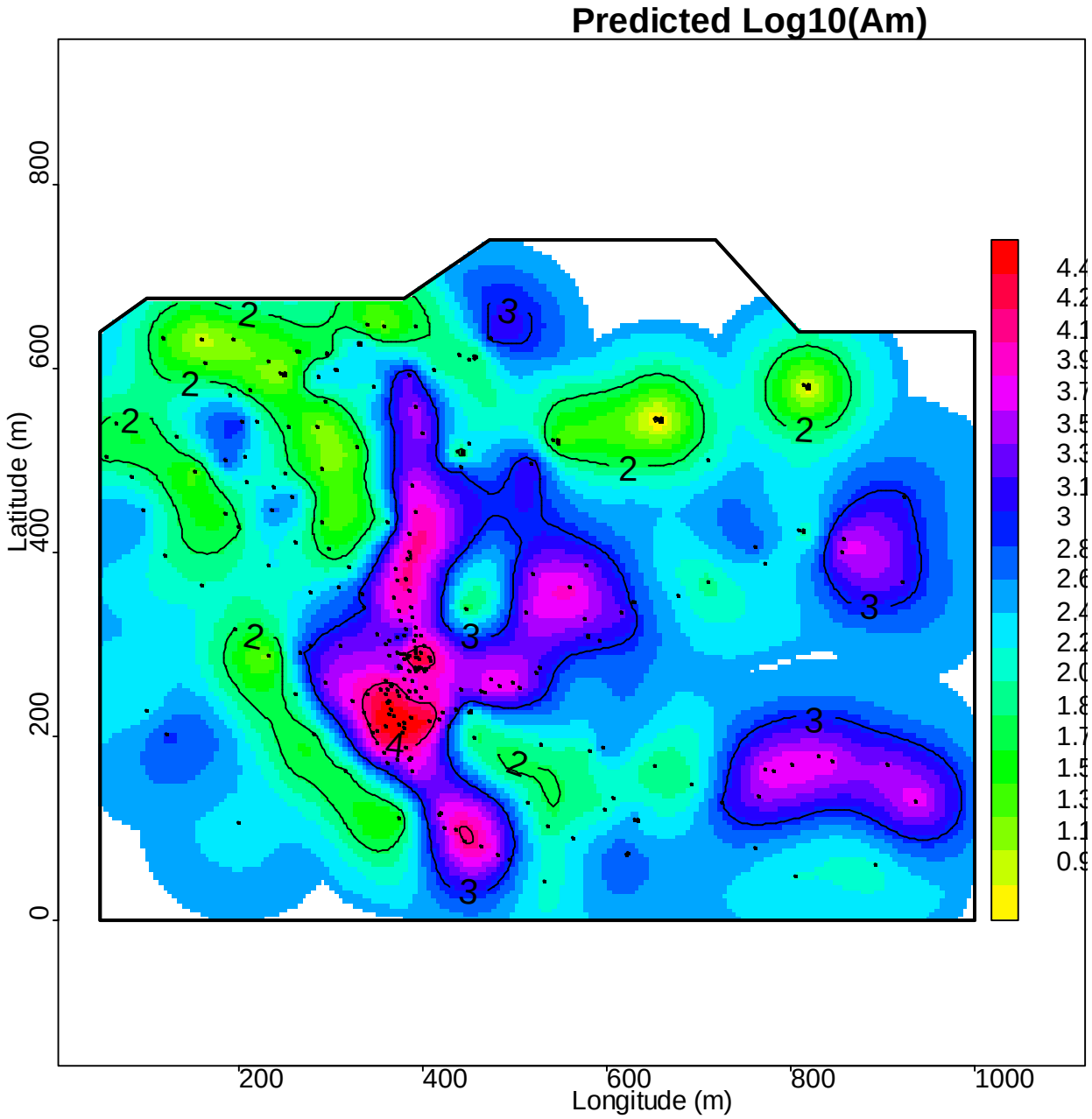
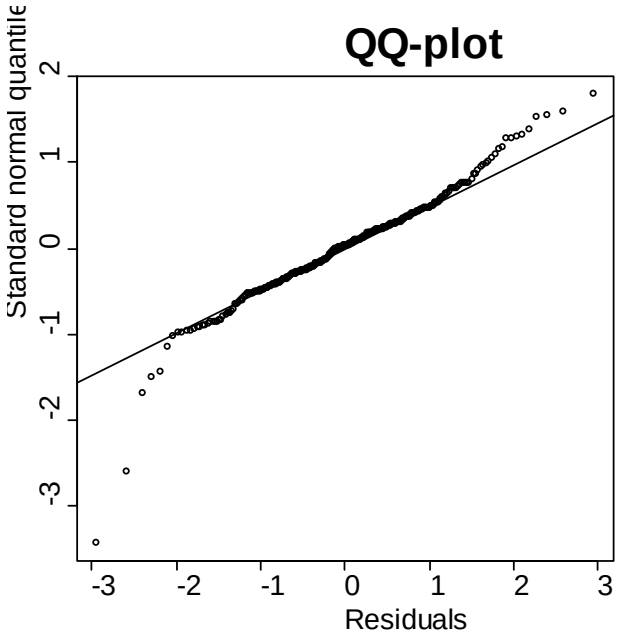
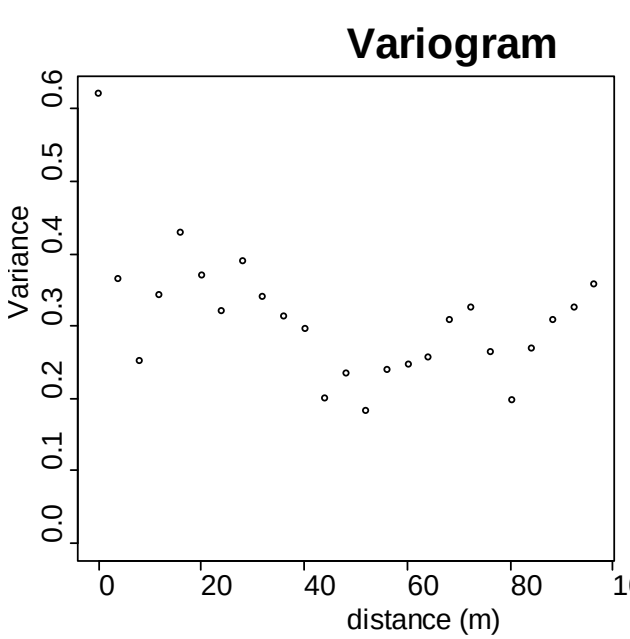


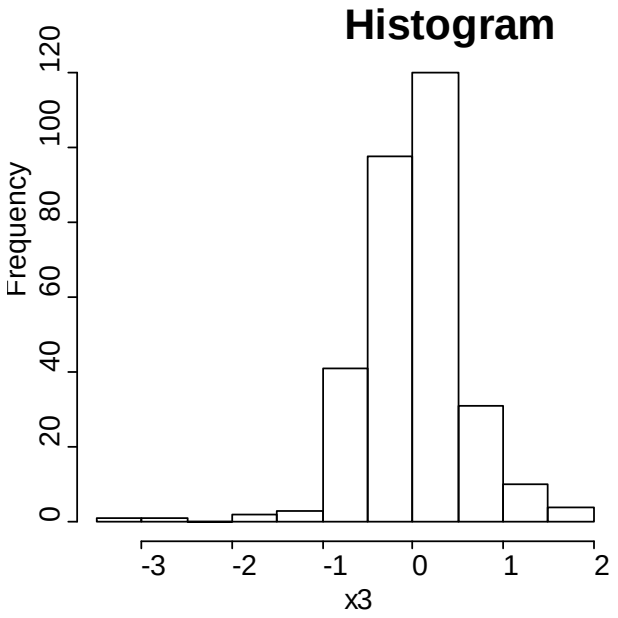
Figure 30 Spatial prediction of the $\log_{10}$(Am) concentration in sub-region2. Black dots are measurement locations.

5.5 Validation of spatial maximum likelihood estimation

In Figure 31 the residuals from the maximum likelihood estimation is evaluated. The QQ-plot, Histogram and the Box-Whisker plot show approximately normal distributed data. The variogram show no spatial dependency.

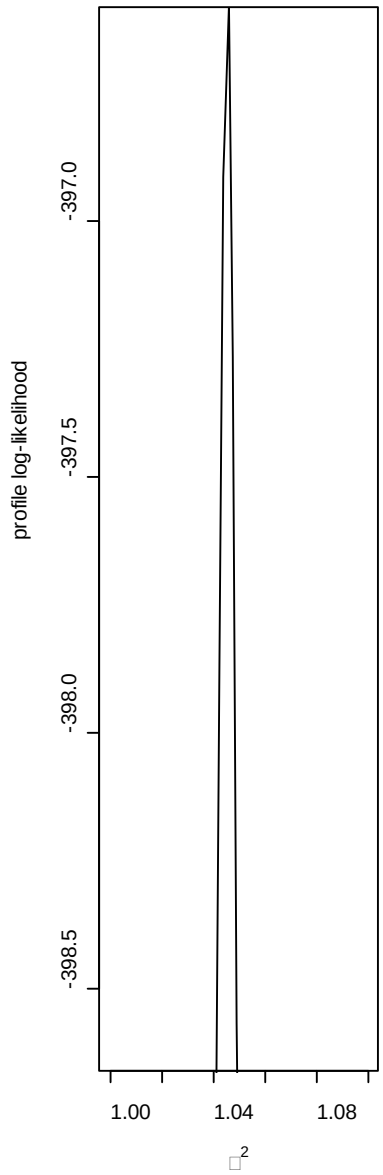






Profile-likelihood curves

In Figure 22 the profile likelihood is shown for the final model of sub-region 2. The dotted horizontal lines indicate approximate 90% and 95% confidence intervals for each parameter.



5.6 Uncertainty of predicted contamination level

The relative standard error (standard error/predicted) of the predicted concentrations are shown in Figure 33. The relative error is in general small in the more densely sampled areas. It does of course also depend on the local variation

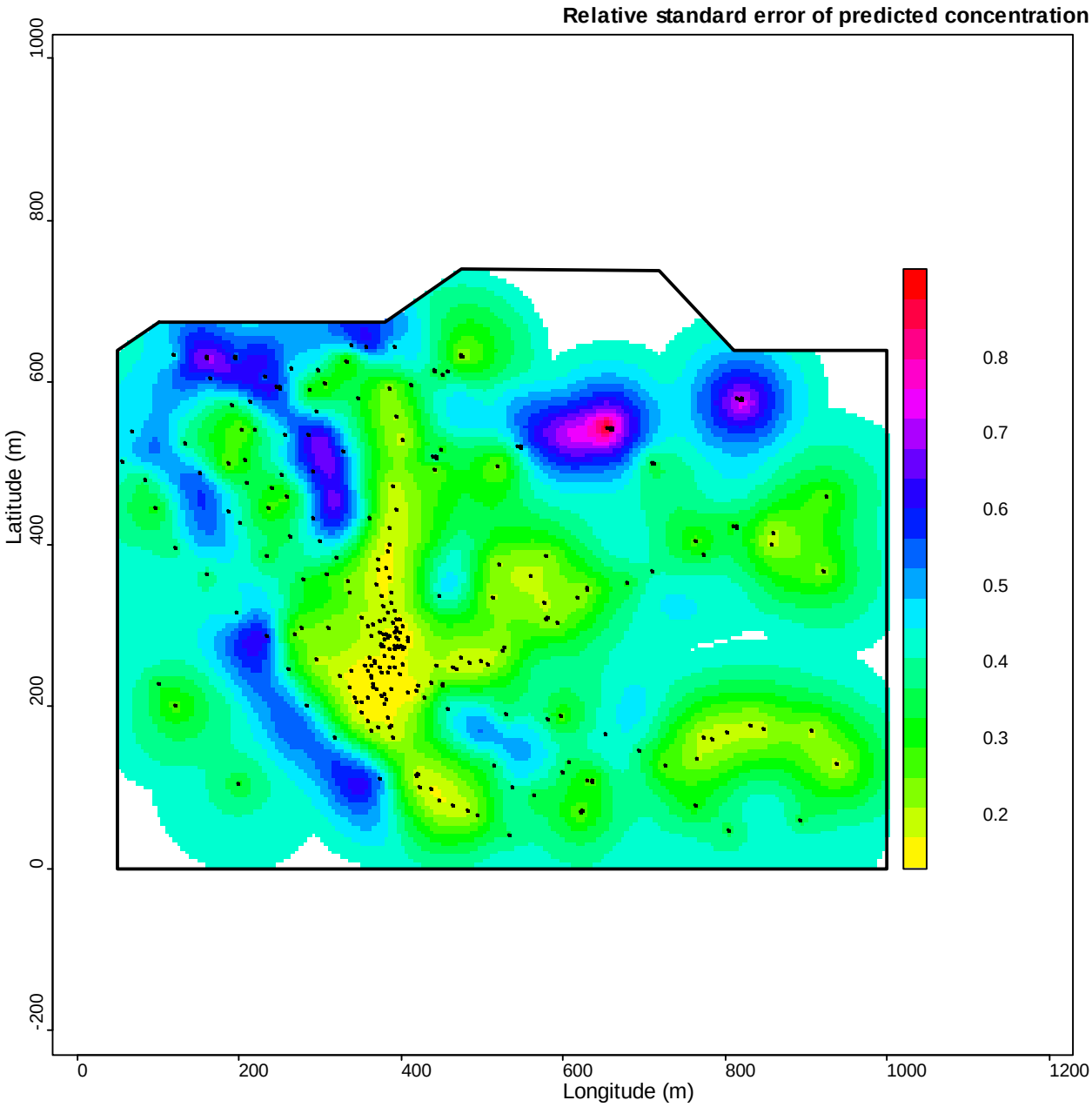


Figure 33 Relative standard error (standard error/predicted) of the log_{10}(Am) predicted concentrations in sub-region 2

Based on normal probability assumptions the probability that the concentration exceeds a certain threshold can be calculated. In Figure 34 and Figure 35 the probability of the $\log_{10}$ concentration exceeding 2 and 3 respectively is shown. As could be seen from the relative standard error little information is available far from the sampling points.

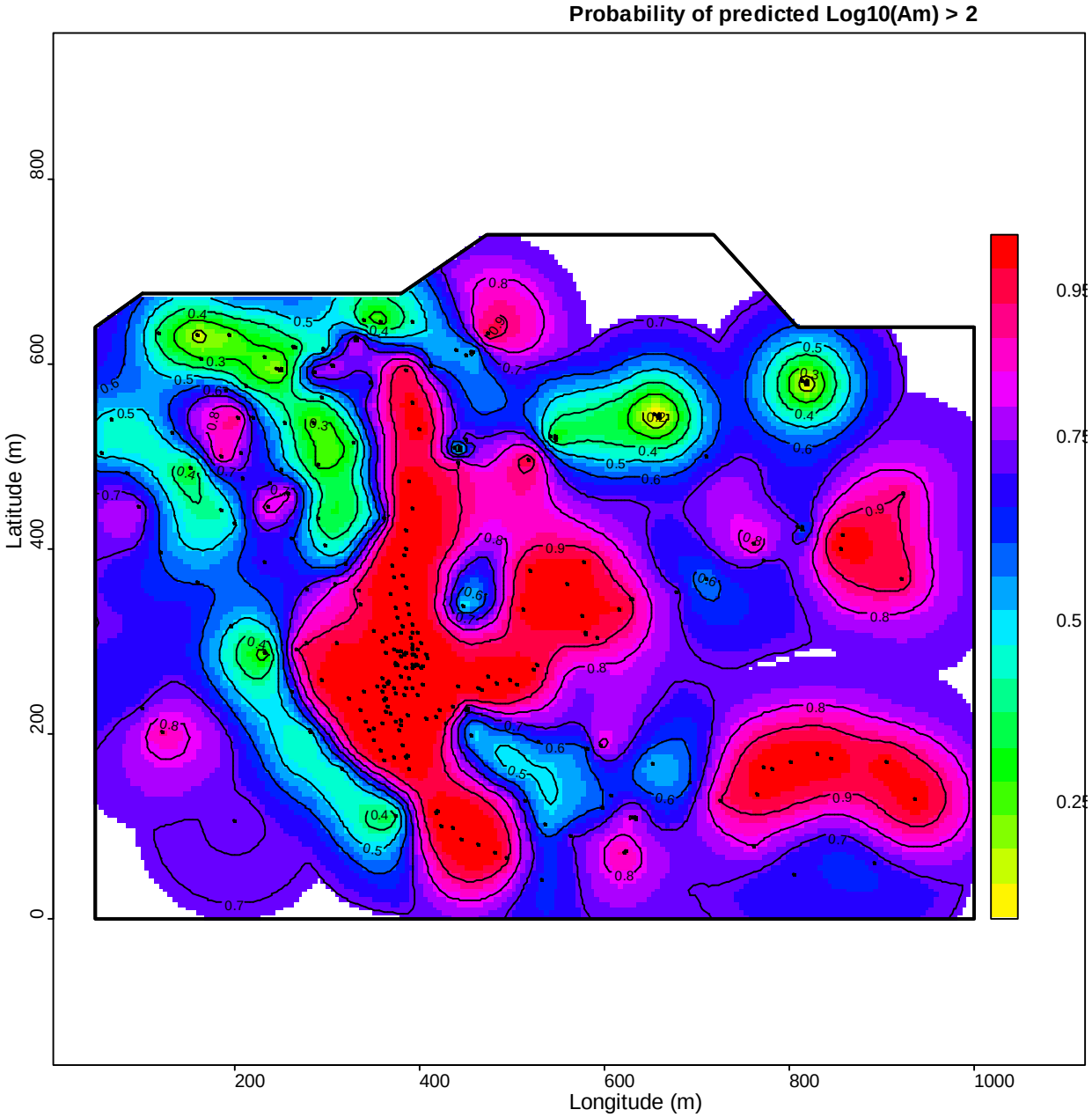


Figure 34 Probability of the $\log_{10}$ concentration exceeding 2

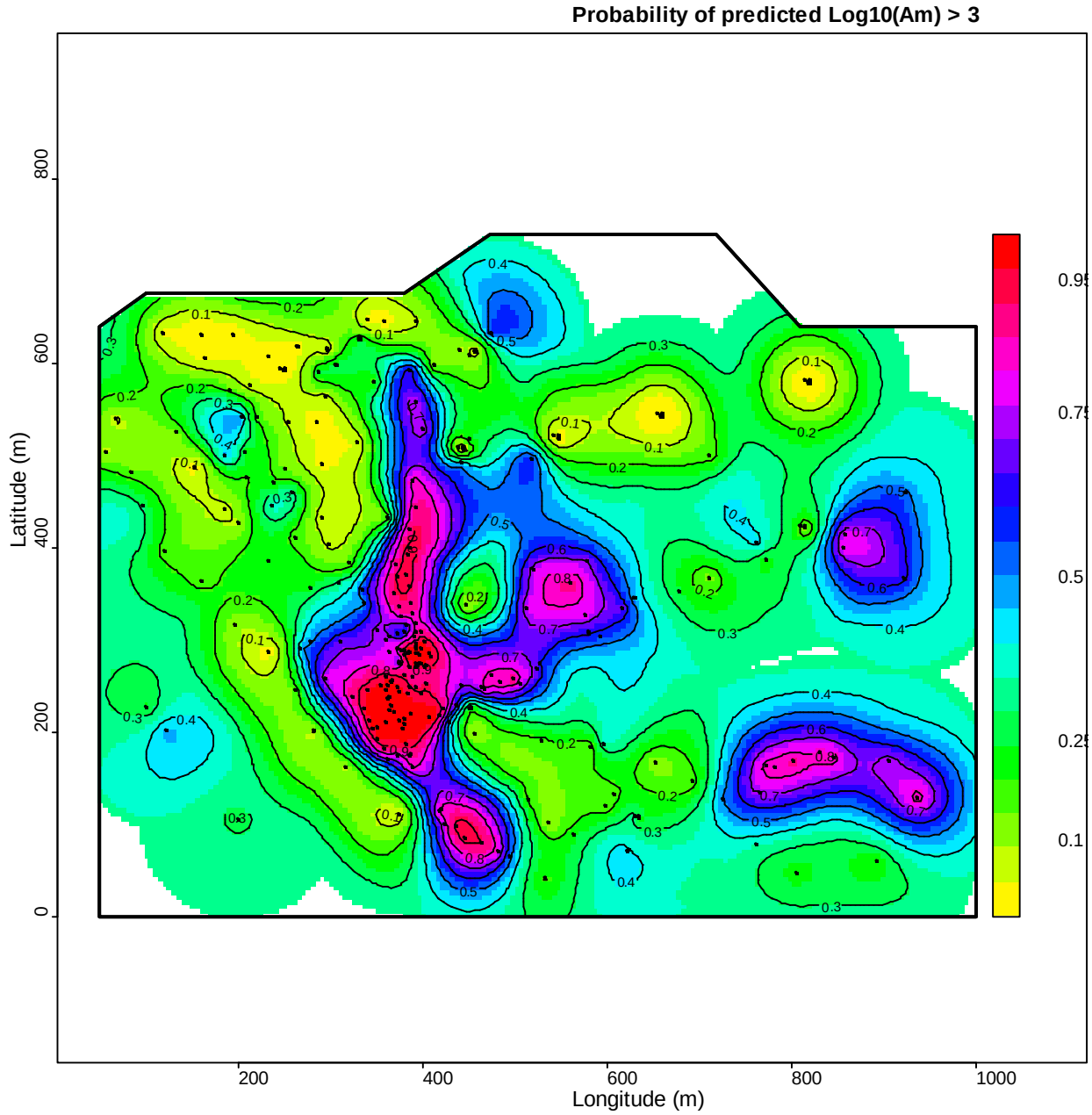


Figure 35 Probability of the log_{10} concentration exceeding 3

Based on the standard error of each predicted value the upper 95% confidence limit for the predicted values is shown in Figure 36. Notice that it is a point-wise 95% confidence interval, so it is not a 95% confidence interval for the surface as such. It can however be used to evaluate the upper limit for each grid-point

.

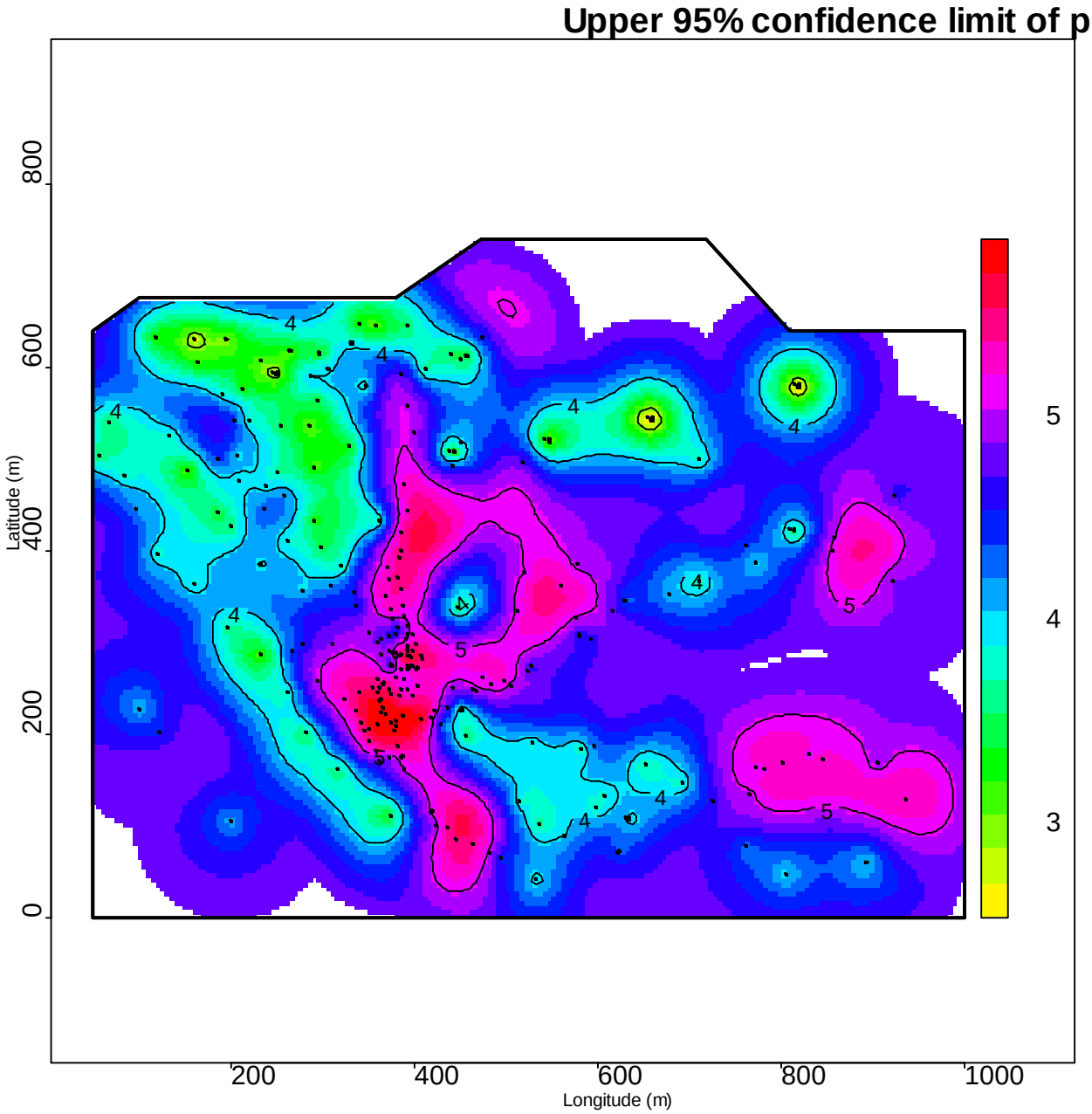


Figure 36 upper 95% confidence limit for the predicted values

5.7 Prediction of 239,240Pu

Based on the predicted surface of ^{241}Am the level of 239,240Pu is calculated as described in section 2.5. The results for the three models are listed in the table below, with the optimal model including trend in bold.

Model	Total Area (km^2)	Area within *range* from nearest neighbour (km^2)	Proportion of total area within *range*	Total ^{241}Am GBq ($Bq\times10^9$) [95% CI]	Total ^{239,240}Pu GBq ($Bq\times10^9$) [95% CI]
Basic model	**0.65**	**0.62**	**95%**	**4.8 [4.2;5.7]**	**28 [24:34]**
+Aniso	0.65	0.59	90%	3.9 [3.5;4.6]	23 [20;27]
+ Trend	0.65	0.61	93%	4.0 [3.6;4.8]	24 [21;28]

Table 17 Predicted level of ^{241}Am and ^{239,240}Pu in sub-region 2. Confidence intervals are found by simulation

For the optimal model 95% corresponding to 0.62 km^2 of the total 0.65km^2 were included and the predicted total amount of ^{241}Am is 4.8 GBq ($Bq\times10^9$) and the predicted total amount of ^{239,240}Pu is 28 GBq ($Bq\times10^9$).

Based on simulations, the confidence intervals for the total level of ^{241}Am and ^{239,240}Pu is in Table 18

	^{241}Am [95%CI] (GBq)	^{239,240}Pu [95%CI] (GBq)
Simulations based on the estimated trend model.	[4.2;5.7]	[24;34]
Simulations as above and in addition including uncertainty of spatial parameters	[2.9;7.0]	[17;42]

Table 18 Confidence intervals of the total level of ^{241}Am and ^{239,240}Pu, based on simulation in region 2.

If locations with a distance to nearest measurement location larger than the estimated range is included in the prediction, the predicted total level of ^{241}Am and ^{239,240}Pu increase as described in the table below

Model	Total ^{241}Am GBq ($Bq\times10^9$)	Total ^{239,240}Pu GBq ($Bq\times10^9$)
Ord	5.1	30
+Aniso	4.4	26
+ Trend	4.3	26

Since most of the area in this sub-region 2 is within a distance to nearest measurement location less than the estimated range the results are almost identical to

leaving out location that are further away from a measurement location than the estimated range.

5.8 Conclusion sub-region of region 2

A clear spatial correlation structure is present in the 0.65 km^2 sub-region of region 2. Data are correlated up to a distance of approximately 108m (estimated range).

Under the optimal spatial model, within the area of prediction, the predicted total amount of ^{241}Am is 4.8 [95%CI: 4.2;5.7]GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 28 [95%CI: 24:34] GBq (Bq×10^9).

Including the uncertainty of the estimated spatial parameters the confidence intervals gets wider, thus within the area of prediction, the predicted total amount of ^{241}Am is 4.8[95%CI: 2.9;7.0] GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 28[95%CI: [17;42]] GBq (Bq×10^9).

6 Summary of region 3

The locations in region 3 is bound by the following coordinates

East (UTM zone 19)	Maximum	
	Minimum	
North (UTM zone 19)	Maximum	
	Minimum	8516100

A summary table of the ^{241}Am concentrations is shown in the table below

N	Min	25% quantile	Mean	Median	75% quantile	Max
8	1	6.2	12.7	12.4	16.5	30.9

7 Summary of region 4

The locations in region 4 is bound by the following coordinates

East (UTM zone 19)	Maximum	491000
	Minimum	
North (UTM zone 19)	Maximum	8505900
	Minimum	8490650

A summary table of the ^{241}Am concentrations is shown in the table below

N	Min	25% quantile	Mean	Median	75% quantile	Max
17	1	4.5	15.0	15.1	20.6	55.6

8 Summary of region 5

The locations in region 5 is bound by the following coordinates

East (UTM zone 19)	Maximum	478000
	Minimum	
North (UTM zone 19)	Maximum	8490650
	Minimum	8475300

A summary table of the ^{241}Am concentrations is shown in the table below

N	Min	25% quantile	Mean	Median	75% quantile	Max
16	1	1	8.8	1	9.4	52.7

9 Summary of region 6

The locations in region 6 is bound by the following coordinates

East (UTM zone 19)	Maximum	
	Minimum	500000
North (UTM zone 19)	Maximum	
	Minimum	8485000

A summary table of the ^{241}Am concentrations is shown in the table below

N	Min	25% quantile	Mean	Median	75% quantile	Max
16	1	1	9.3	1	16.6	37.7

10 Soil samples in region 2

10.1 Descriptive statistics

A summary of the measurement is shown below.

	Minimum	25% quantile	Mean	Median	75% quantile	Maximum
Am	1	26	6.3×10^3	99	2.5×10^3	2.8×10^3
$Log_{10}Am$	0.04*	1.4	2.4	2.0	3.4	5.4
Distances (m)	0.2	440	971	858	1352	3867

Table 19 Summary statistics sub- region 2.

*0.1 was added to the concentrations to obtain numerical stability

All spatial statistics is performed on $\log_{10}$ transformed data. A summary plot is shown in Figure 37

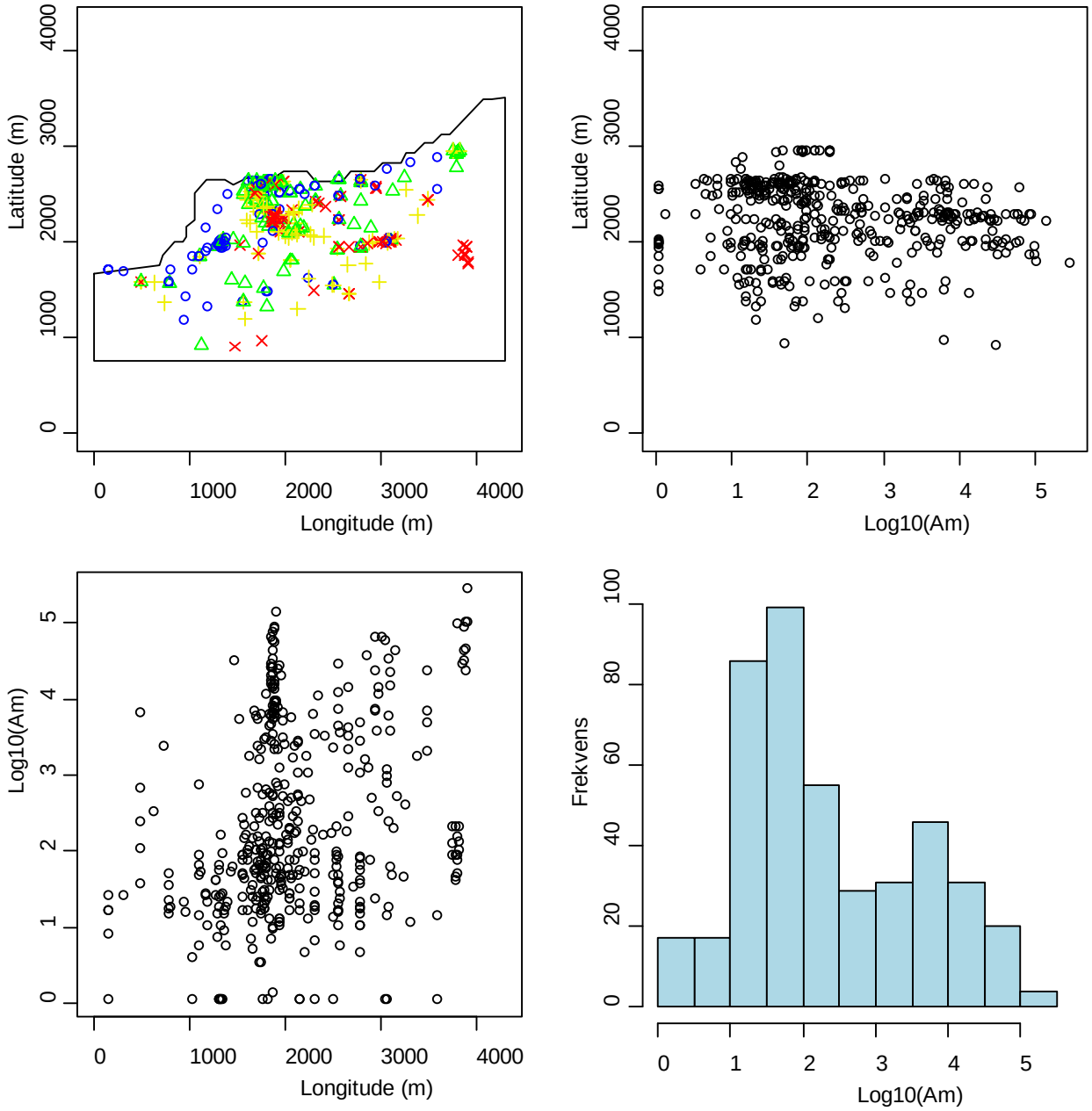


Figure 37 Summary plot of the $\log_{10}$(Am) for soil samples in region 2. In the upper left: Blue circle is the 0-25% quantiles, green triangle is the 25-50% quantiles, yellow plus is the 50-75% quantiles and red cross is the 75-100% quantiles.

10.2 Spatial correlation

Before doing the maximum likelihood estimation a variogram is calculated. The calculated values and the fitted spherical variogram function is shown in Figure 38

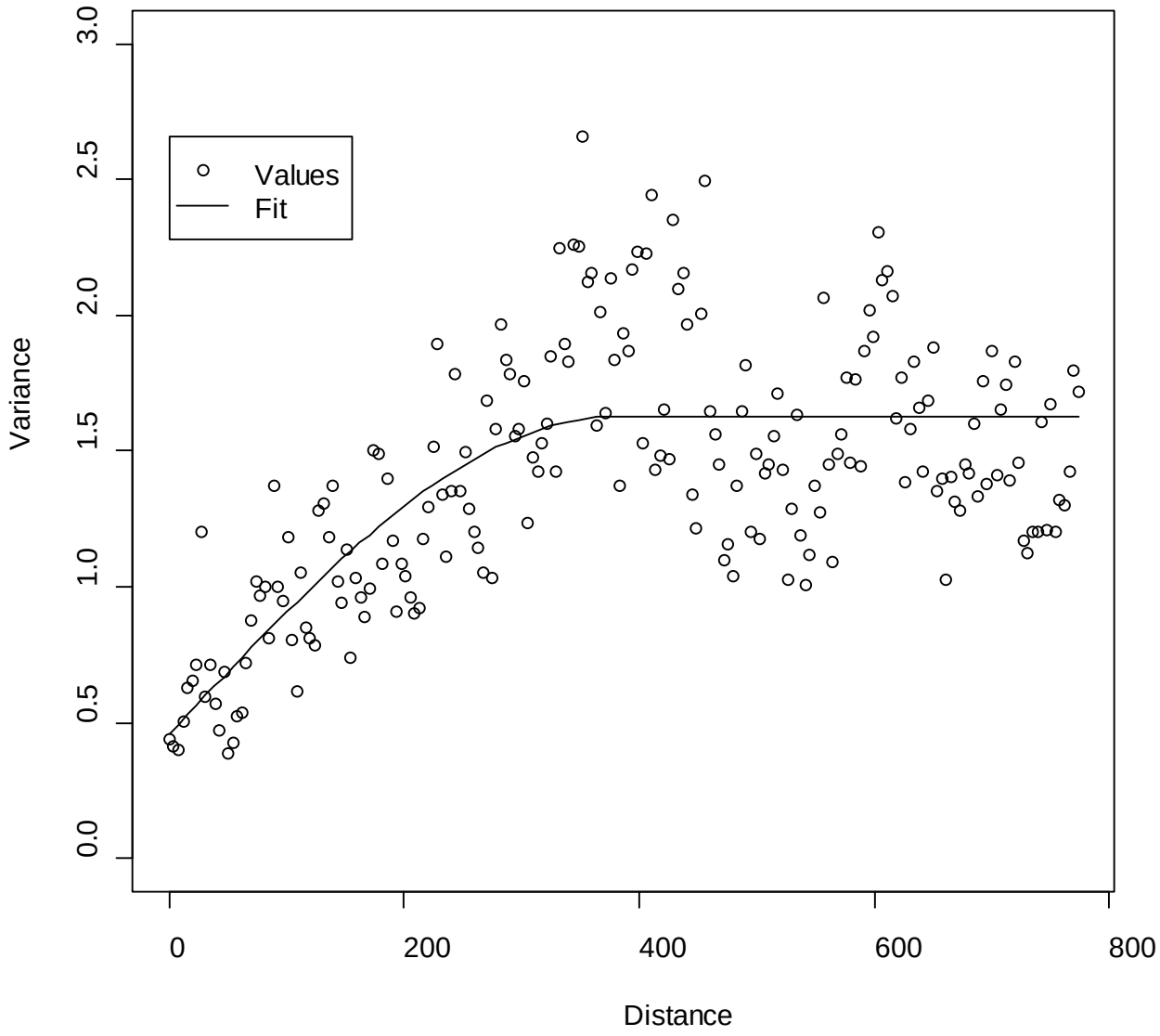


Figure 38 Variogram for log_{10}(Am) for soil samples in Region 2

A very clear spatial correlation is seen. The parameters in the fitted spherical variogram function is

Spatial parameter	Nugget	Sill	Partial Sill	Range
Estimated value	0.46	1.63	1.17	379m

Table 20 Parameter values from nonlinear weighted regression of the variogram, soil samples region 2

So, preliminary it is assumed that data is correlated within a range of approximately 400 m

10.3 Prediction of contamination level

As for region 1 and 2 the basic spatial model and more complex models were investigated: 1) Including a 1st order trend and/or 2) including anisotropy, meaning that the spatial correlation is independent on direction. The spatial maximum likelihood estimation resulted in the correlation structure shown in the table below

Model	Nugget	Sill	Partial Sill	Range (m)	Lambda	Beta0, BetaX, BetaY	Angle	Ratio	AIC	BIC
Basic model	0.471	1.058	0.5872	272	0.8117	1.0191			1199	1220
+Anisotropy*	0.453	1.054	0.6009	220	0.8118	1.0167	2.492	1.562	1200	1228
+trend	**0.4824**	**0.8933**	**0.4109**	**222**	**0.8207**	**1.4321** **5.6×10^{-4}** **-** **7.8×10^{-4}**			**1181**	**1210**

Table 21 Maximum likelihood parameter estimates for spatial models, in sub-region 2

.

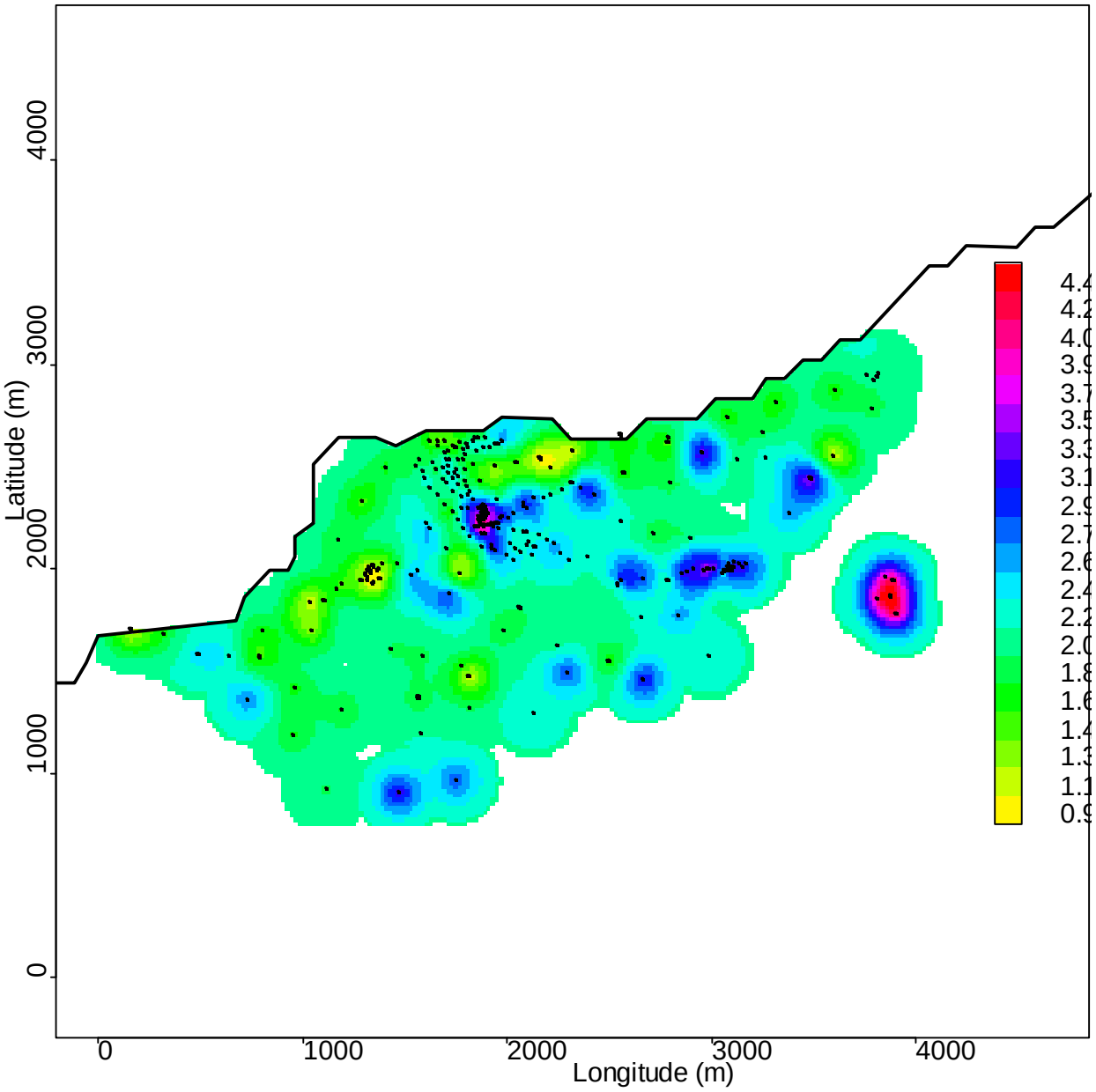
Based on the AIC and BIC a model with a first order trend is the most optimal. The two model selection criteria agree.

Therefore it is decided to use the model without an overall spatial trend.

Based on the ML-estimation data is correlated within a distance of 222m. The correlation is not depending on direction. This implies that locations further away than 222 m from a measurement location can not be predicted based on the measurements. The partial sill is approximately on the same level as the nugget.

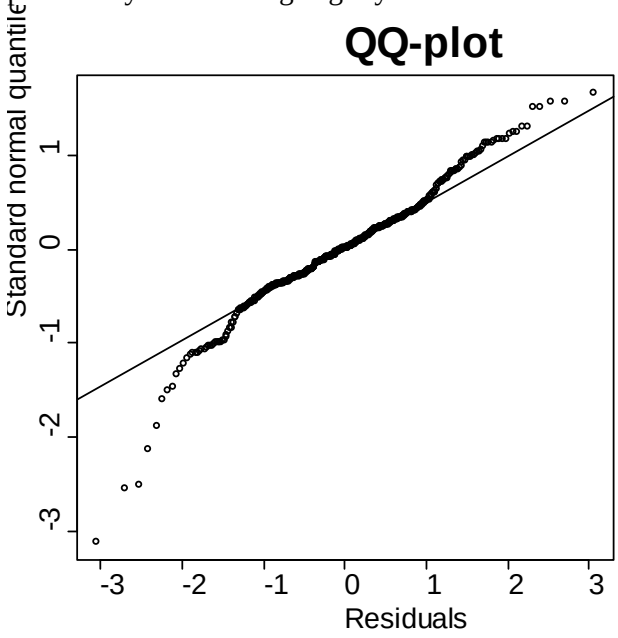
The spatial prediction on a 20m×20m grid of the $\log_{10}$(Am) soil concentrations in region 2 is shown in Figure 39 with height curves overlaid. Locations further than 222 m from a measurement location is not predicted.

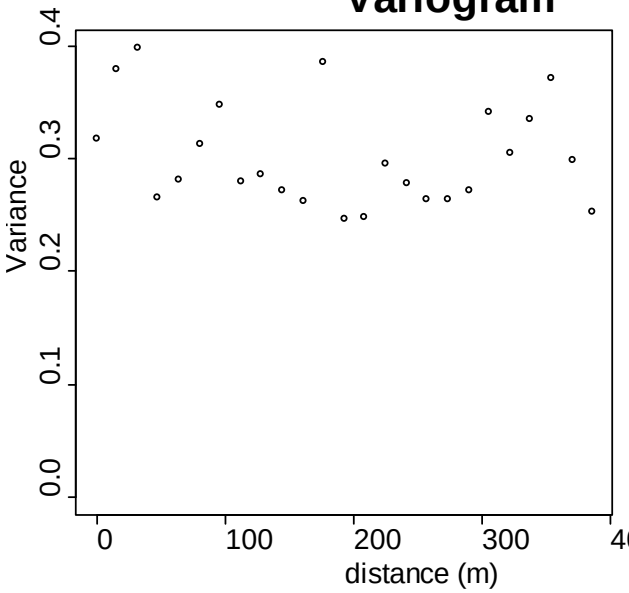
The trend parameters (Xcoord,Ycoord) show that concentration increase in south-east direction

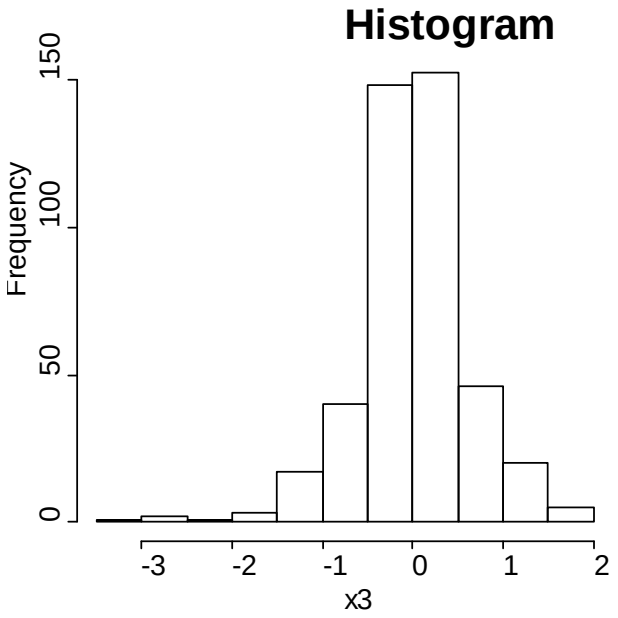

Latitude (m)
0
1000
2000
3000
4000
Longitude (m)
0
1000
2000
3000
4000

10.4 Validation of spatial maximum likelihood estimation

In Figure 40 the residuals from the maximum likelihood estimation is evaluated. The QQ-plot, Histogram and the Box-Whisker plot show approximately normal distributed data. The variogram show perhaps a small spatial dependency with data points very close having slightly less variance

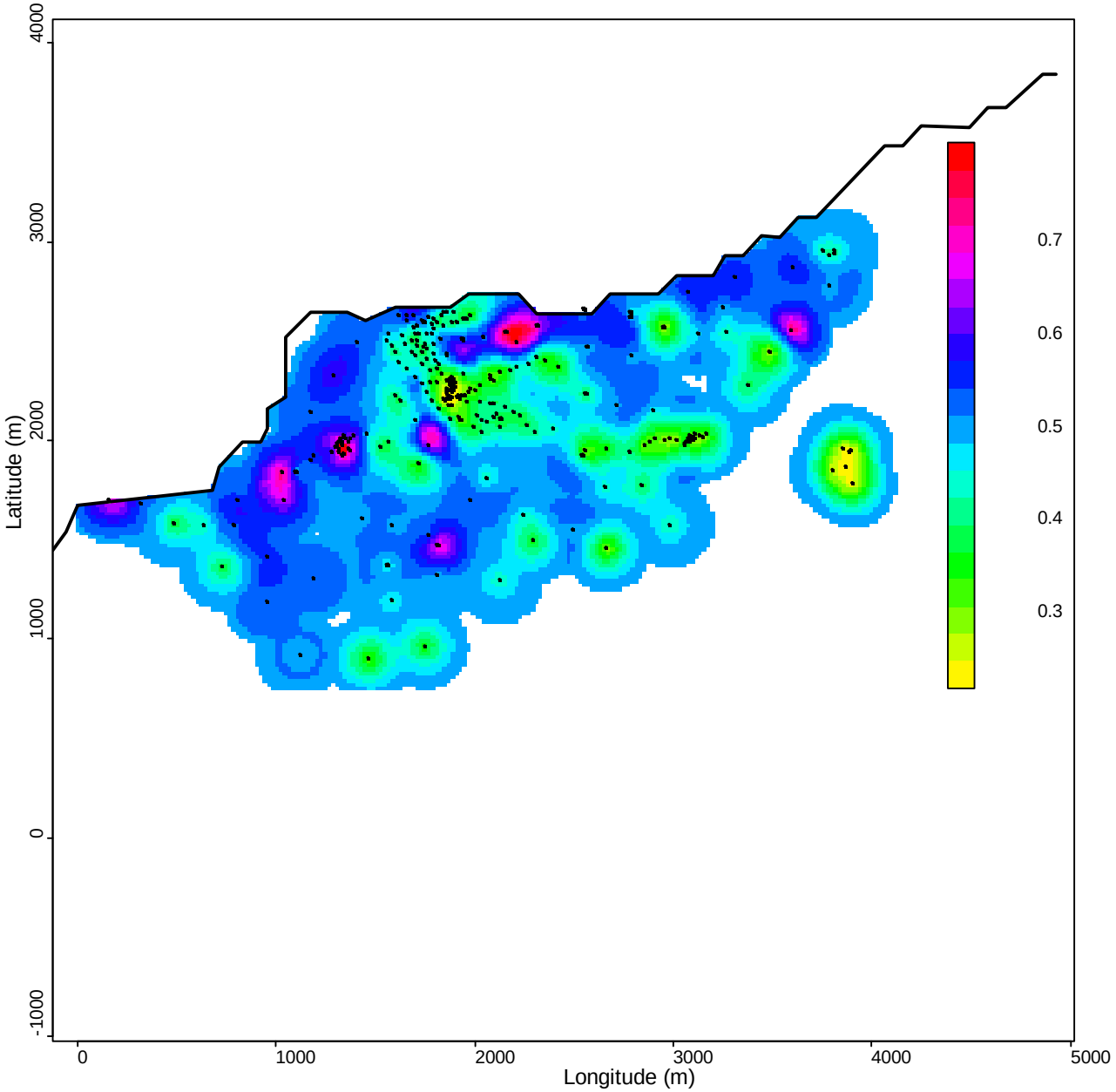




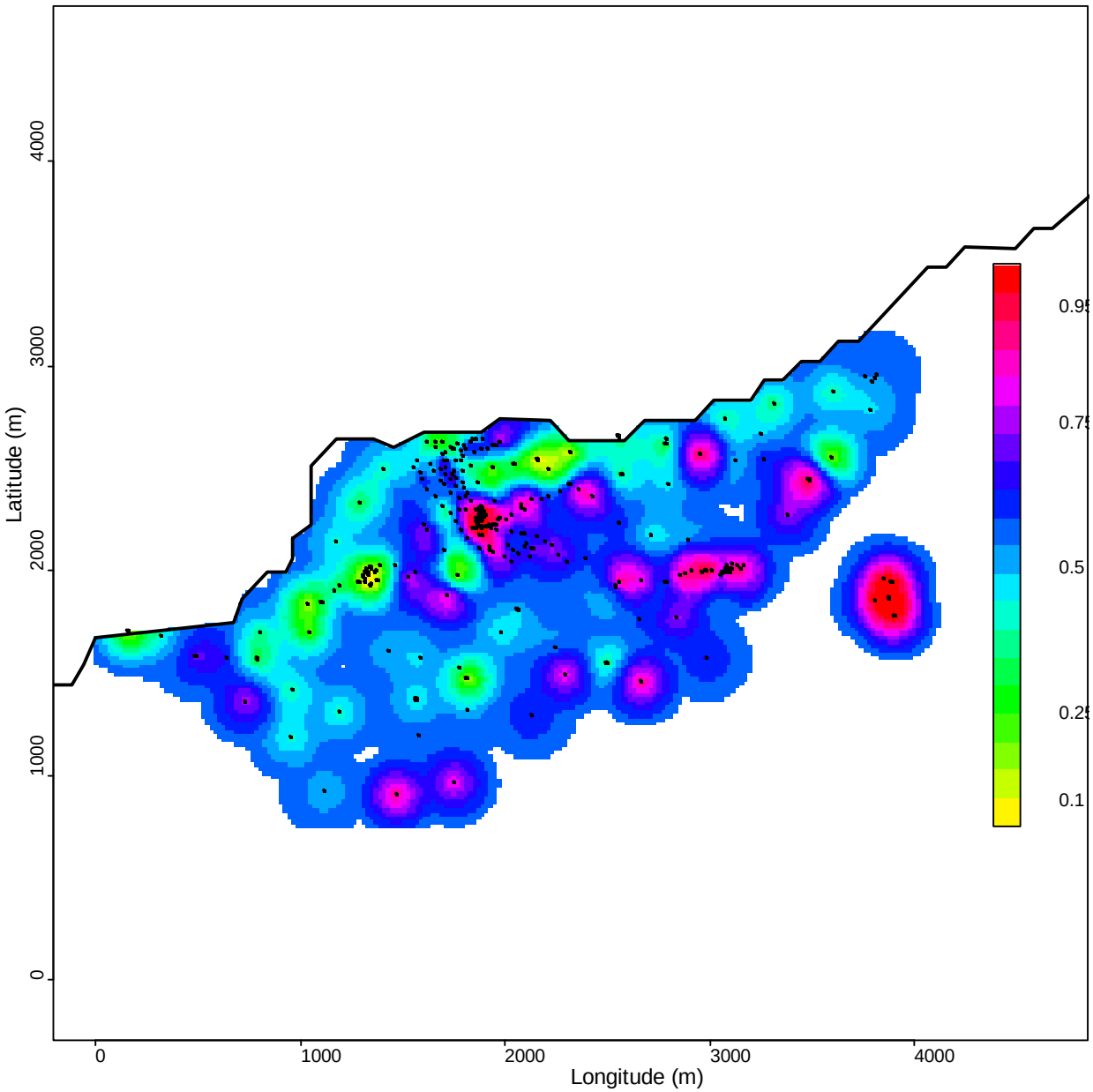


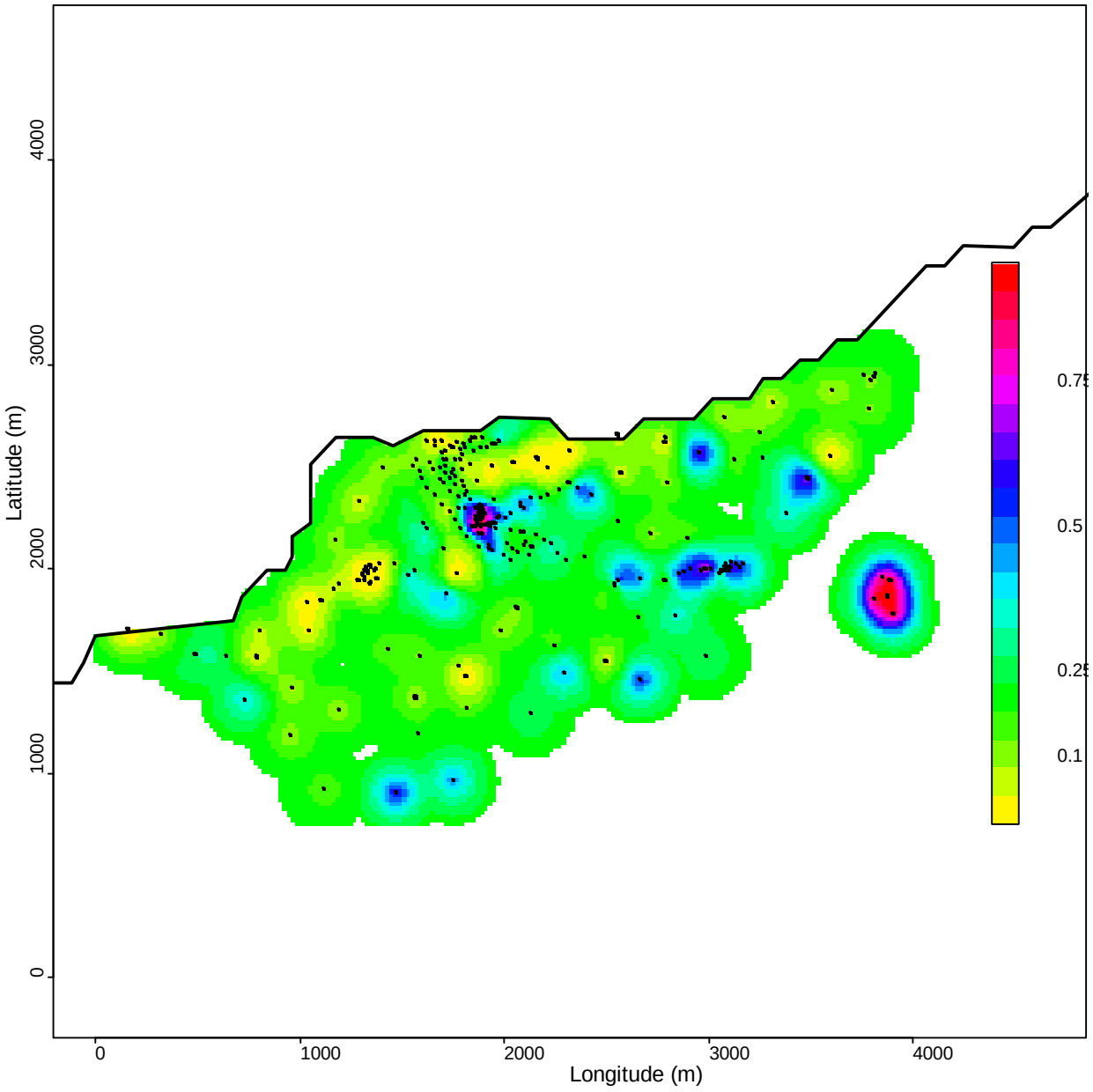
10.5 Uncertainty of predicted contamination level

The relative standard error (standard error/predicted) of the predicted concentrations are shown in Figure 41. The relative error is in general small in the more densely sampled areas. It does of course also depend on the local variation

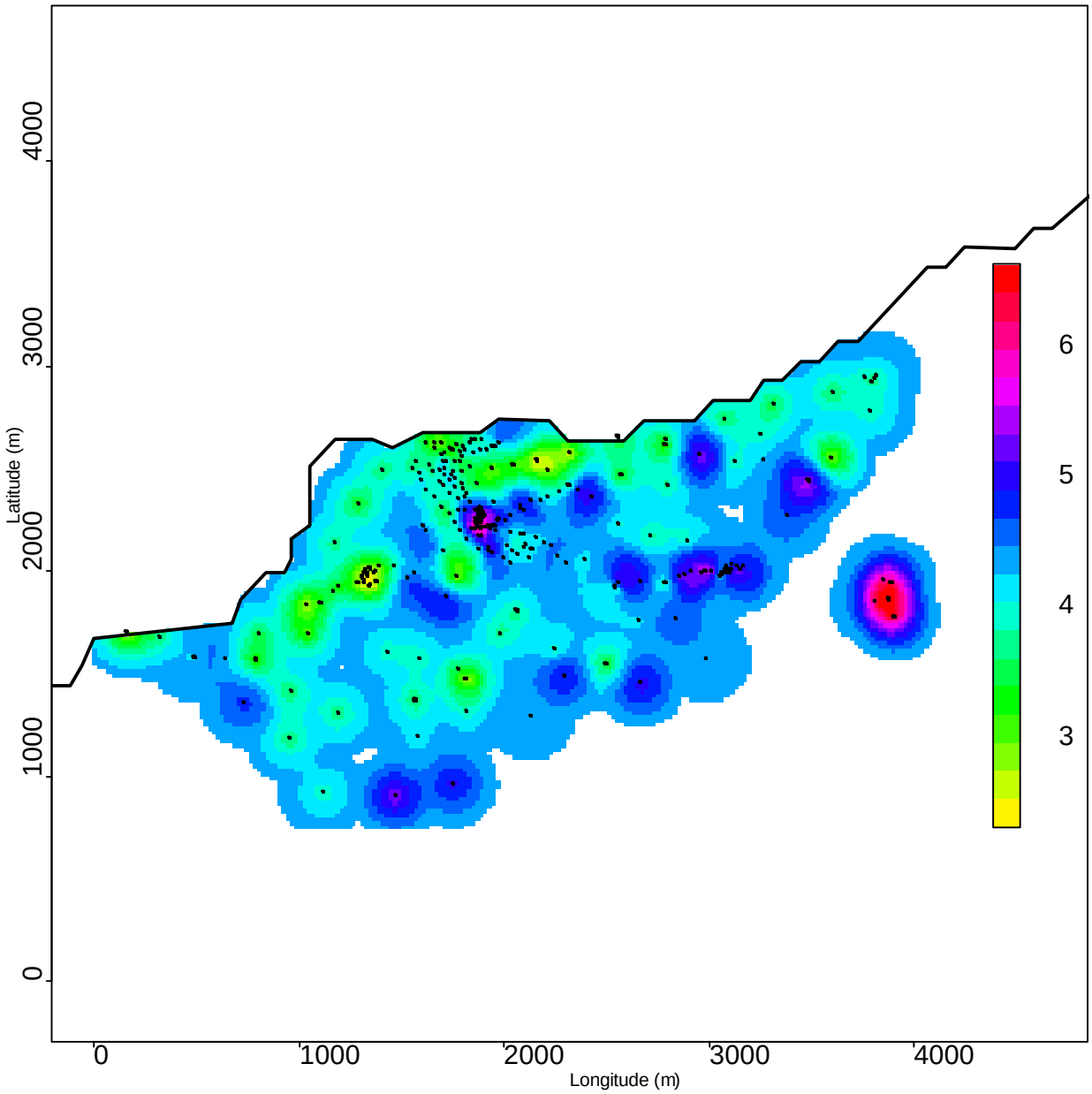


Based on normal probability assumptions the probability that the concentration exceeds a certain threshold can be calculated. In Figure 42 and Figure 43the probability of the $\log_{10}$ concentration exceeding 2 and 3 respectively is shown. As could be seen from the relative standard error little information is available far from the sampling points.




Latitude (m)
0
1000
2000
3000
4000
Longitude (m)
0
1000
2000
3000
4000
0.75
0.5
0.25
0.1

Based on the standard error of each predicted value the upper 95% confidence limit for the predicted values is shown in Figure 44.



10.6 Prediction of 239,240Pu

Based on the predicted surface of ^{241}Am the level of 239,240Pu is calculated as described in section 2.5. The results for the three models are listed in the table below, with the optimal model including trend in bold.

Model	Area (km^2)	Area within *range* from nearest neighbour (km^2)	Proportion area within *range*	Total ^{241}Am GBq ($Bq\times10^9$) [95%CI]	Total ^{239,240}Pu GBq ($Bq\times10^9$) [95%CI]
Basic model	8.00	5.48	69%	58	350
+Aniso	8.00	5.04	63%	53	319
+ Trend	**8.00**	**5.06**	**63%**	**27[18;58]**	**162[112;350]**

Table 22 Predicted level of ^{241}Am and ^{239,240}Pu for Soil samples in region 2. Confidence intervals are found by simulation.

As can be seen the area, where the level is predicted depends of the model or more specifically on the estimated range. The higher range the larger area is included. For the optimal model 63% corresponding to 5.06 km^2 of the total 8.00km^2 were included.

For the optimal model 63% corresponding to 4.93 km^2 of the total 8.00km^2 were included and the predicted total amount of ^{241}Am is 45 GBq ($Bq\times10^9$). The predicted total amount of ^{239,240}Pu is 270 GBq ($Bq\times10^9$).

Based on simulations, the confidence intervals for the total level of ^{241}Am and ^{239,240}Pu is Table 23

	^{241}Am [95%CI] (GBq)	^{239,240}Pu [95%CI] (GBq)
Simulations based on the estimated trend model.	[18;58]	[112;350]
Simulations as above and in addition including uncertainty of spatial parameters	[7;41]	[40;342]

Table 23 Confidence intervals of the total level of ^{241}Am and ^{239,240}Pu for soil samples, based on simulation in region 2.

If locations with a distance to nearest measurement location larger than the estimated range is included in the prediction the predicted total level of ^{241}Am the level of 239,240Pu increase as described in the table below

Model	Total ^{241}Am GBq ($Bq\times10^9$)	Total ^{239,240}Pu GBq ($Bq\times10^9$)
Basic model	74	446
+Aniso	71	427
+ Trend	37	223

Table 24 Predicted level of ^{241}Am and ^{239,240}Pu for soil samples in region 2 including locations further than the estimated range away from the nearest measurement location

10.7 Conclusion soil samples of region 2

A clear spatial correlation structure is present in the soil samples of region 2. Data are correlated up to a distance of approximately 222m (estimated range). The trend parameters show that concentration increase in south-east direction. The predicted values a very similar to the combined Soil and CAP data. Due to the size of the area and the sampling performed, some areas are further than 222m away from the nearest observation. The partial sill was estimated to be .4 and the nugget approximately 0.5, so the partial sill and nugget are on the same level

Under the optimal spatial model, within the area of prediction, the predicted total amount of ^{241}Am is 27[18;58]GBq (Bq$\times 10^9$) and the predicted total amount of ^{239,240}Pu is 162[112;350]GBq (Bq$\times 10^9$).

Including the uncertainty of the estimated spatial parameters the confidence intervals gets wider, thus within the area of prediction, the predicted total amount of ^{241}Am is is 27[7;41]GBq (Bq$\times 10^9$) and the predicted total amount of ^{239,240}Pu is 162[40;342]GBq (Bq$\times 10^9$).

11 Overall discussion and conclusion

A clear spatial correlation structure is present. Data are correlated up to a distance of approximately 100-400 meters (estimated ranges) depending on the area investigated. The tendency is that smaller areas results in smaller estimated ranges.

Based on the analysis in region 2 a south-east trend is identified with increasing concentration from the coast in the south-east direction. There is no natural explanation of a trend in this direction. However, indication of the trend was consistently there in region 1, 2 and also in the sub-region 2.

The strength of spatial statistics is that predicted values as well as an uncertainty measure of the predicted values are estimated taking the spatial correlation into account. Due to the size of the area and the sampling performed, some areas are further than the estimated range from the nearest observation. This means that the information in the sampled points is of no use in these areas, resulting in the average concentration being the best guess. Despite a Log_{10} transformation of the data the few high concentration still result in a relatively high mean. This most likely doesn't reflect reality and also give rise to strange discontinuities where remote areas are surrounded of low level measurements. Then at a distance of range the level increases to the average level. So interpretation in these areas should be done with great caution. Further sampling is needed for more precise conclusions in these areas, and therefore prediction is not done in these areas.

Several measures of uncertainty is presented, all contributing to the interpretation of the accuracy of the predicted level of ^{241}Am.
The relative standard error, describe the uncertainty compared to the mean.
The upper 95% confidence interval limit gives an upper limit of the possible level at each location i.e. a limit where the true value with 95% certainty is smaller. Due to the relative large variation the 95% confidence limit is wide. Notice that it is a point-wise 95% confidence interval, so it is not a 95% confidence interval for the surface as such. It can however be used to evaluate the upper limit for each grid-point
Furthermore two threshold maps, showing the probability of $\log_{10}$(Am) exceeding a value of 2 and 3 respectively gives information on the position of the threshold in the estimated normal distribution in the prediction locations.

With the current data it has been a rather complex task to find the optimal solution in the parameter space consisting of nugget, sill, range, mean component, trend and anisotropy. Overall the predicted levels of ^{241}Am looks pretty much the same in densely sampled areas for all models. The parameters become increasingly important in sparsely sampled areas. The range parameter and the nugget compared to the partial sill determine the weighting of neighbourhood measurements. A large range will allow prediction further away from measurements and in general a more smooth prediction surface.

The combined soil and CAP data gave approximately the correlation structure as the soil data alone. Furthermore the confidence intervals for the total predicted level of ^{241}Am and ^{239,240}Pu were overlapping.

In region 1 the maximum observed level of ^{241}Am was 2.8×10^{5} Bqm^{-2}. Most high levels were observed near Narsaarsuk. This area was also sampled most intensively. However, in Grønnedal the maximum observed level of ^{241}Am was 1.9×10^{4} Bqm^{-2}. Based on the topography of the area and the location of the hot-spots a hypothesis could be wind, snow and water transport particles to pools where the level therefore

increase. The area is so sparse sampled that it is of no use to try to predict the overall amount of ^{241}Am and ^{239,240}Pu. This is only done for region 2.

Prediction of the overall amount of ^{241}Am and ^{239,240}Pu is only based on grid points within the range from the nearest measurement location. The overall amount is therefore highly depending on the estimated model.

Under the optimal spatial model in region 2, within the area of prediction, the predicted total amount of ^{241}Am is 45[95%CI: 36;65] GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 270[95%CI: 212;387] GBq (Bq×10^9). Including the uncertainty of the estimated spatial parameters the confidence intervals gets wider, thus within the area of prediction, the predicted total amount of ^{241}Am is 45[95%CI: 24;82] GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 270[95%CI: 140;490] GBq (Bq×10^9).

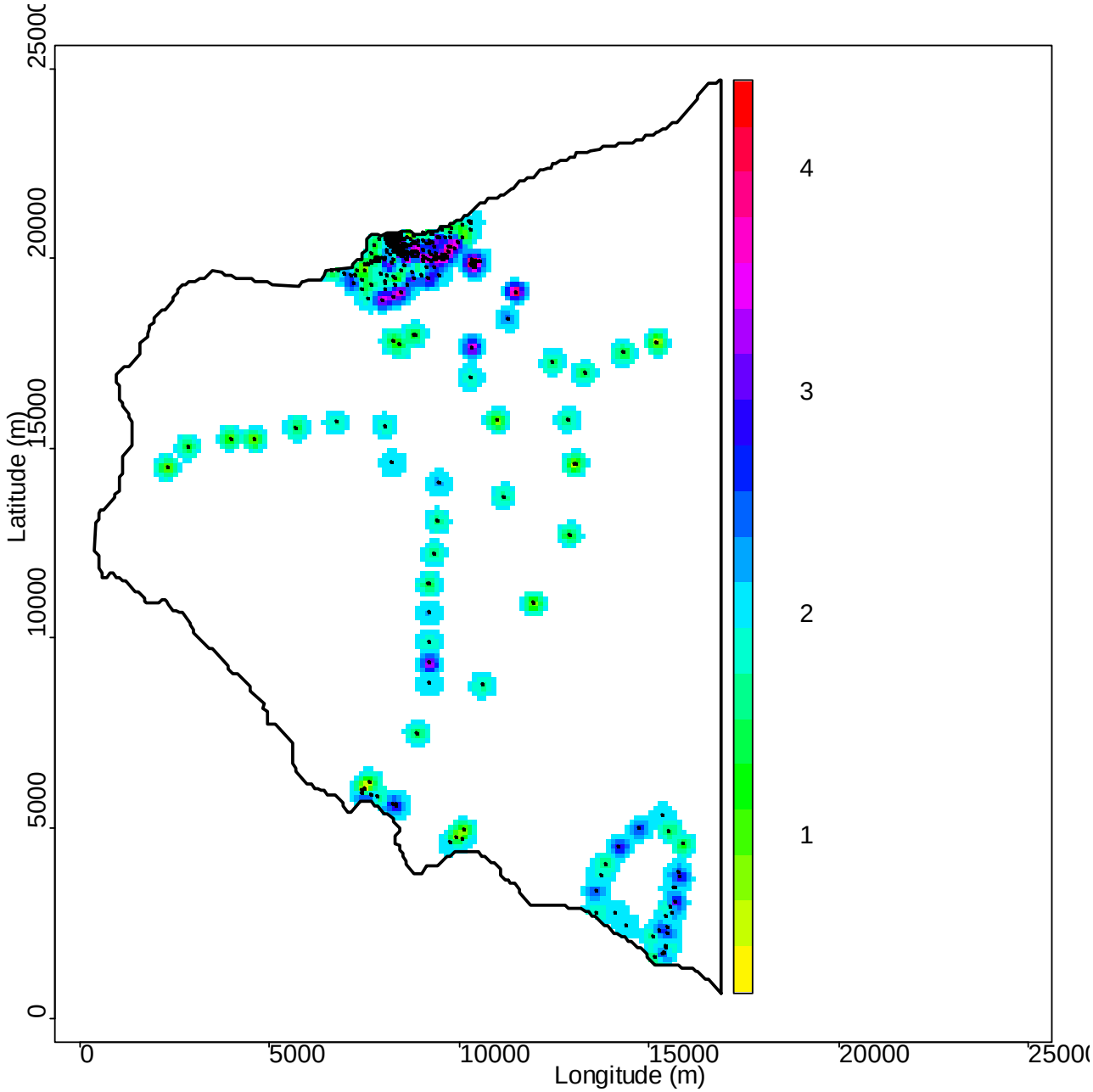
A clear spatial correlation structure is present in the 0.65 km^2 sub-region of region 2. Data are correlated up to a distance of approximately 108m (estimated range).

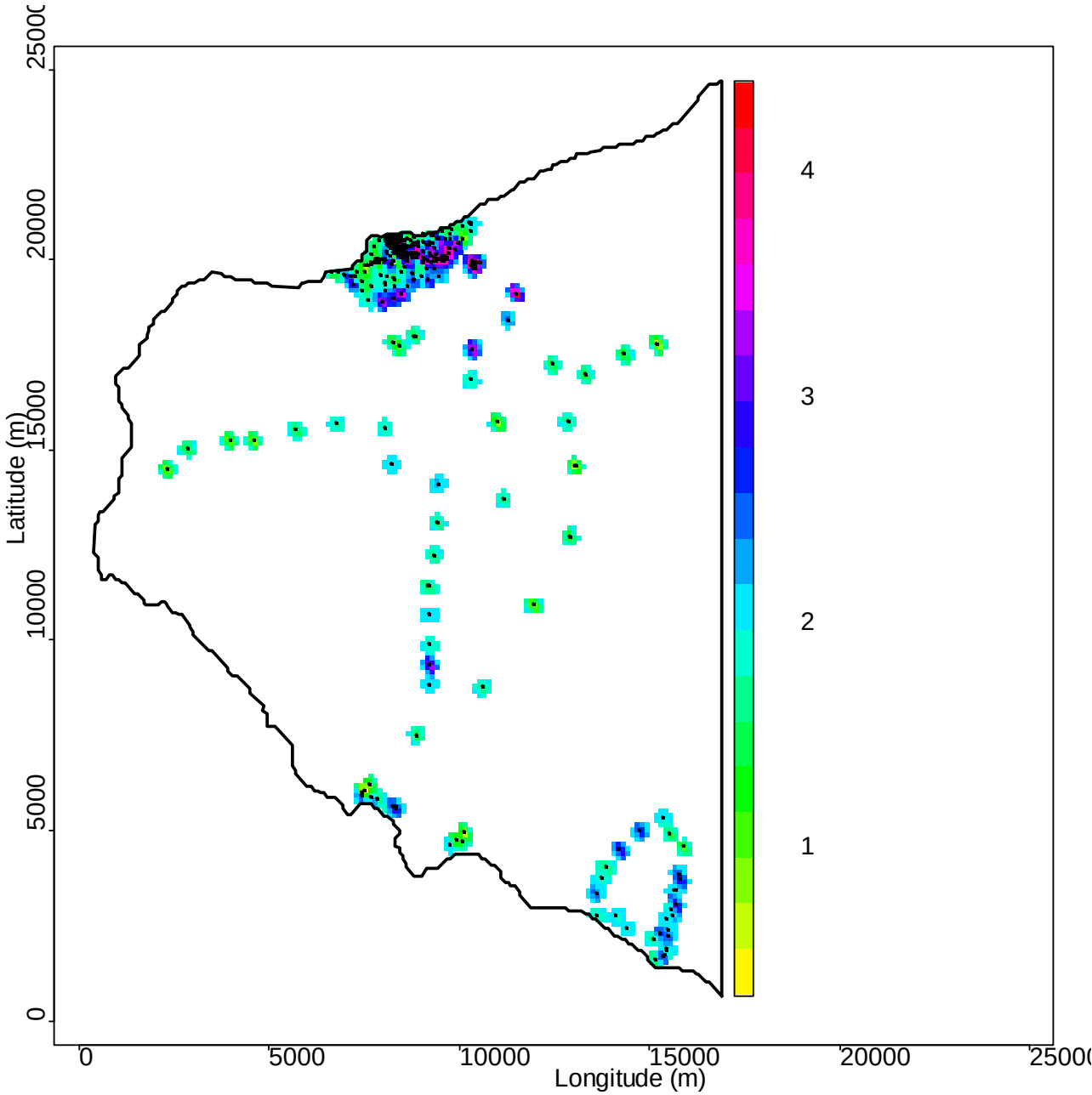
Under the optimal spatial model in sub-region 2, within the area of prediction, the predicted total amount of ^{241}Am is 4.8 [95%CI: 4.2;5.7]GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 28 [95%CI: 24:34] GBq (Bq×10^9). Including the uncertainty of the estimated spatial parameters the confidence intervals gets wider, thus within the area of prediction, the predicted total amount of ^{241}Am is 4.8[95%CI: 2.9;7.0] GBq (Bq×10^9) and the predicted total amount of ^{239,240}Pu is 28[95%CI: [17;42]] GBq (Bq×10^9).

The level of ^{241}Am was low in region 3-6, the area around Moriusaq, Saunders Ø, Wolstenholme Ø and the area near the Thule airbase. The maximum observed level of ^{241}Am was 56 Bqm^{-2}.

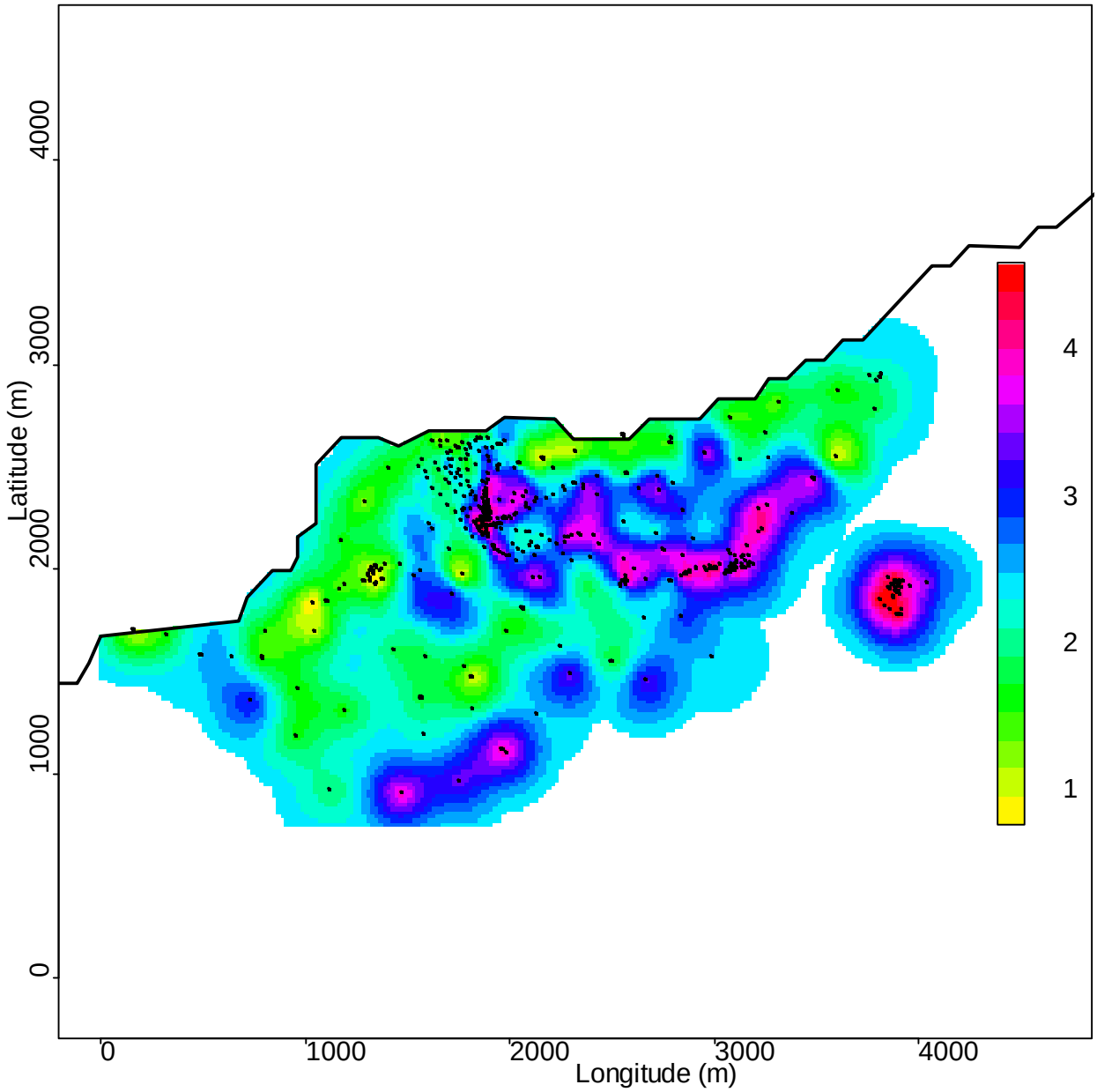
12 Appendix

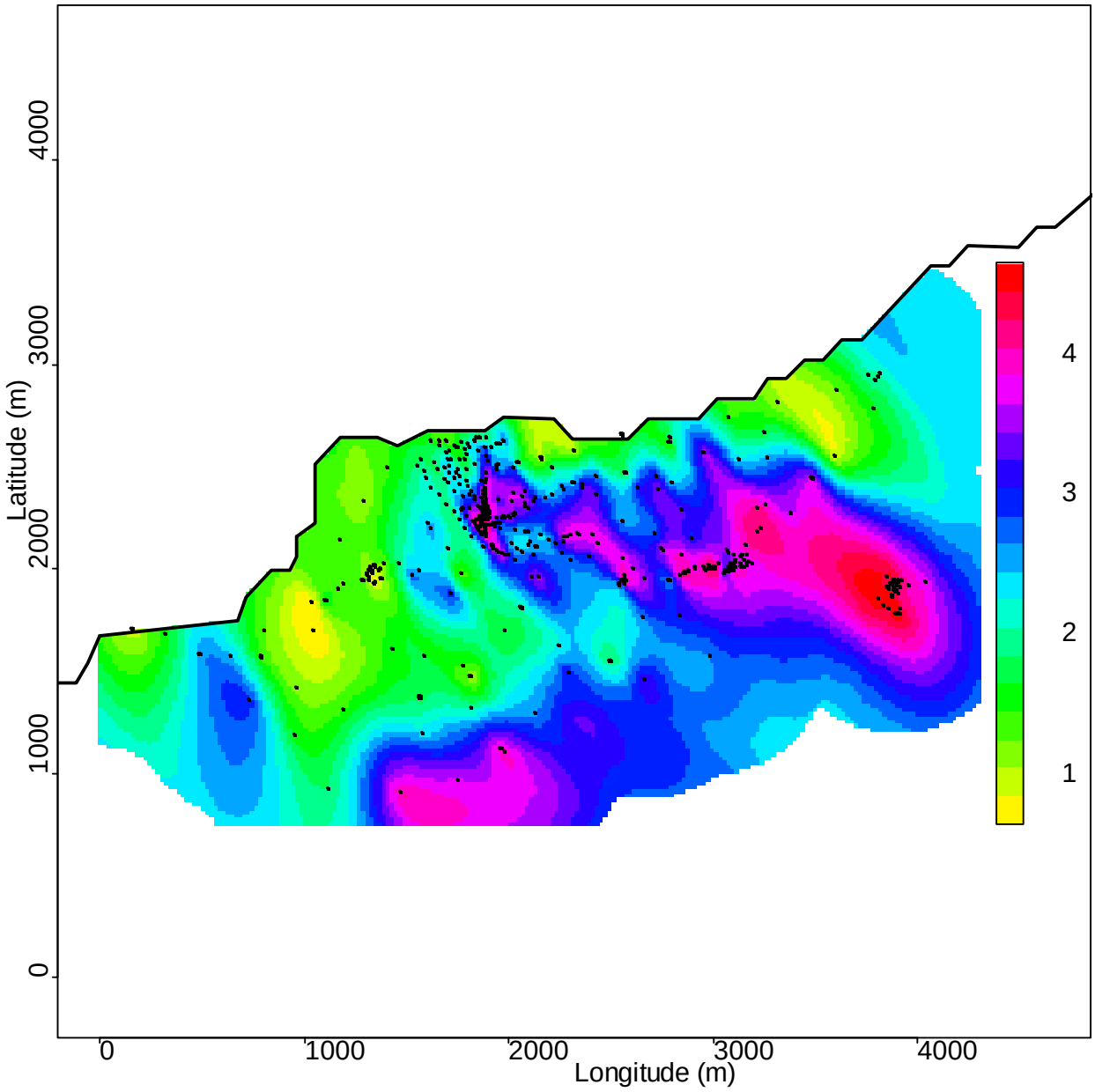
12.1 Region 1




25000
20000
15000
10000
5000
0
Latitude (m)
0
5000
10000
15000
20000
25000
Longitude (m)
4
3
2
1

12.2 Region 2


Latitude (m)
0
1000
2000
3000
4000
Longitude (m)
0
1000
2000
3000
4000
1
2
3
4


Latitude (m)
0
1000
2000
3000
4000
0
1000
2000
3000
4000
Longitude (m)
1
2
3
4

12.3 Sub-Region 2

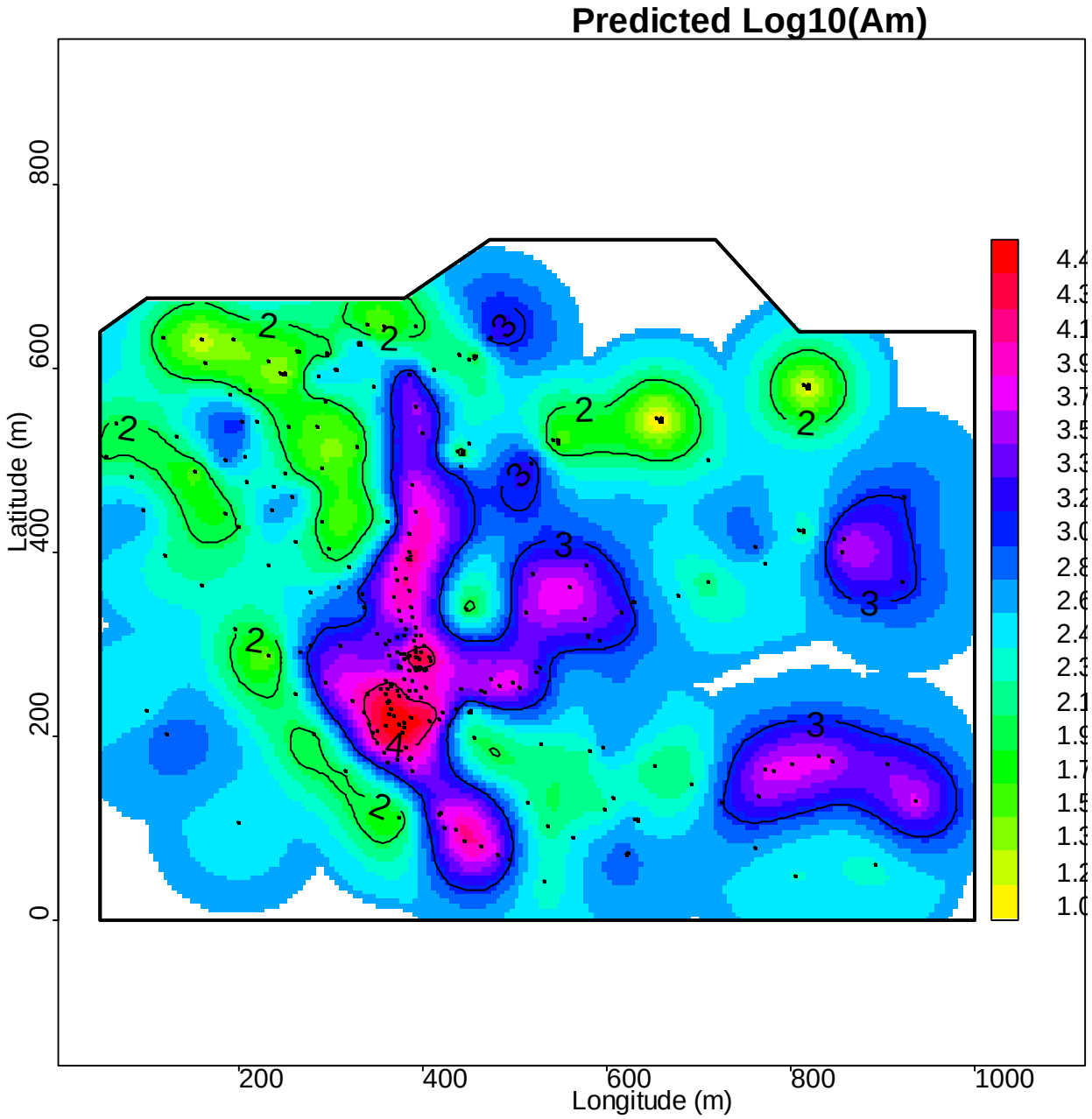


Figure 49 Spatial prediction of the log_{10}(Am) concentration in sub-region 2. Model with trend. Black dots are measurement locations.

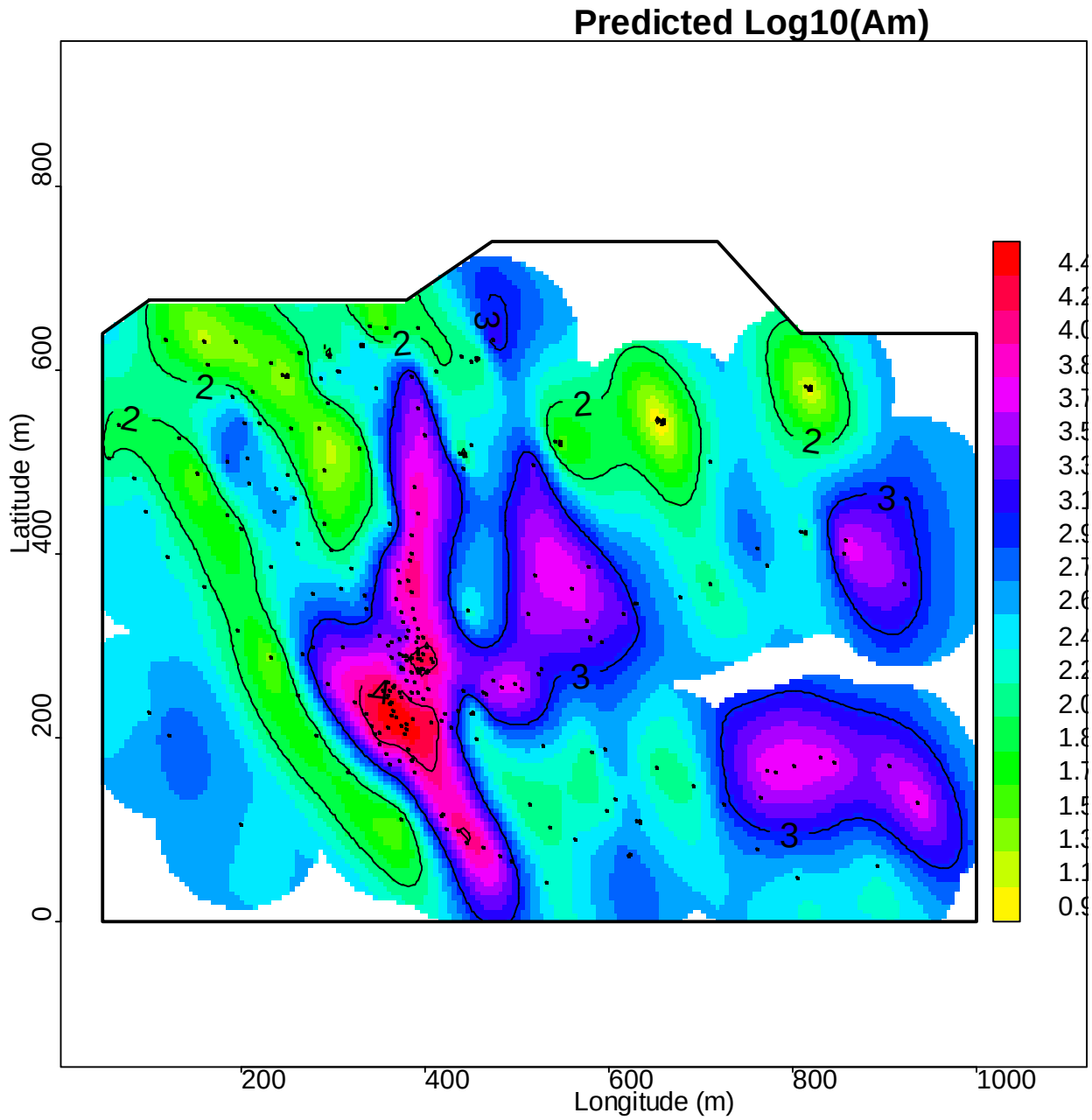


Figure 50 Spatial prediction of the log_{10}(Am) concentration in sub-region 2. Model with Anisotropy. Black dots are measurement locations.

Risø DTU is the National Laboratory for Sustainable Energy. Our research focuses on development of energy technologies and systems with minimal effect on climate, and contributes to innovation, education and policy. Risø has large experimental facilities and interdisciplinary research environments, and includes the national centre for nuclear technologies.

Risø DTU
National Laboratory for Sustainable Energy
Technical University of Denmark

Frederiksborgvej 399
PO Box 49
DK-4000 Roskilde
Denmark
Phone +45 4677 4677
Fax +45 4677 5688

www.risoe.dtu.dk