

# Quantifying uncertainty of geological 3D layer models, constructed with *a-priori* geological expertise

Jan Gunnink<sup>1</sup>, Denise Maljers<sup>2</sup> and Jan Hummelman<sup>2</sup>,

<sup>1</sup>TNO Built Environment and Geosciences – Geological Survey of the Netherlands, Utrecht, the Netherlands. E-mail: jan.gunnink@tno.nl

<sup>2</sup>TNO Built Environment and Geosciences – Geological Survey of the Netherlands, Utrecht, the Netherlands

---

## **Abstract**

Uncertainty quantification of geological models that are constructed with additional geological expert-knowledge is not straightforward. To construct sound geological 3D layer models we use a lot of additional knowledge, with an uncertainty that is hard to quantify. Examples of geological expert knowledge are trend surfaces that display a geological plausible basin, additional points that guide the pinching out of geological formations along its depositional extent, etc. All the added geological knowledge, together with the stringent assumptions of normality and second-order stationarity, makes the kriging standard error in our modeling not usable as a measure of uncertainty.

We developed a procedure to quantify the uncertainty of our geological 3D layer model that uses cross-validation in a moving window environment to calculate mean deviations and standard errors on a sub-regional scale. Subsequently, we rescaled the x-validation standard error to account for local data configuration and clustering. Summary statistics (Root Mean Squared Prediction Error, Root Mean Error Variance and prediction interval) indicate that there is no bias in the geological model estimation and that the absolute values are trustworthy. An additional check on the above described results was provided by a spatial bootstrapping procedure. Based on 100 bootstrap samples that were “redrilled” in the model, the variance was not comparable to the cross-validated results.

To validate the results of the uncertainty quantification we used a sample of (6% randomly selected) drillings as an independent dataset. Results indicate that for datasets with lots of data, the uncertainty quantification provided satisfying results, in terms of RMSE and RMEV. In cases of sparse data, setting aside 6% of the drillings leads to unfavorable statistics, indicating that a minimum of datapoints is needed to obtain a reliable quantification of uncertainty.

---

*Keywords: uncertainty quantification, geological modeling, cross-validation*

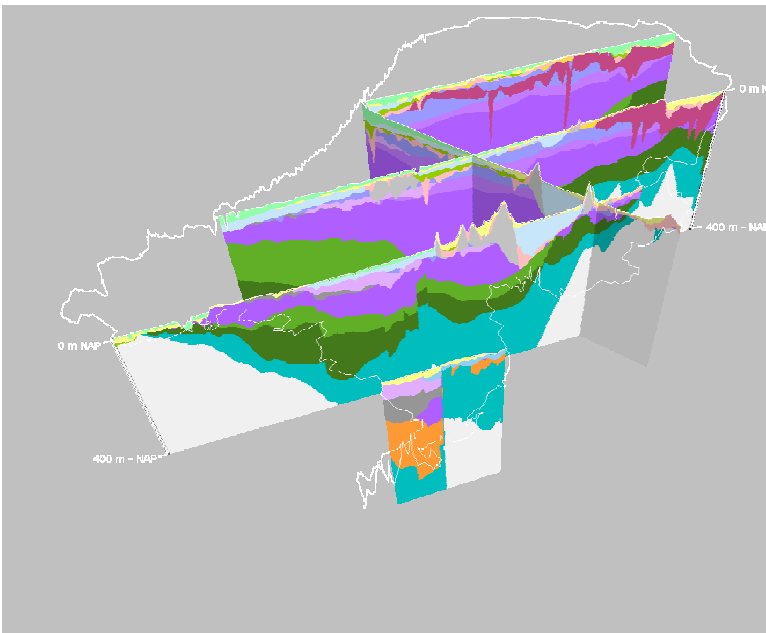
# 1. INTRODUCTION

The use of 3D geological models for the shallow subsurface (<500m depth) is increasing. Besides scientific interest, groundwater and management of resources like sand and gravel, geotechnical issues, are increasingly using 3D geological models. To the users of 3D geological models it is important to stress the fact that the model is a representation of the subsurface, and that there is uncertainty about the modeled surface. The quantification of the uncertainties is not a straightforward procedure. The calculation of plausible 3D geological models requires a lot of expert knowledge. The application of standard interpolation and modeling routines often do not result in a coherent representation of the geological surfaces and often do not conform to the general geological setting of the area under investigation.

The modeling of the shallow subsurface of geological Formations in The Netherlands is based on a selection of drillings from the DINO database ([www.dinoloket.nl](http://www.dinoloket.nl)). The selection criteria are based on the quality of the borehole description, the depth of the drilling and the (more or less) even distribution of the drillings over the area.

After geological interpretation of the borehole intervals in terms of geological Formations (33 different Geological Formations are distinguished), subsequent interpolation onto a regular grid (cellsize: 100mx100m) and modeling leads to the basal surface of each formation. After stacking the basal surfaces in the correct geological sequence, the result is a complete 3D geological model: Digital geological Model of The Netherlands (DGM), see Fig. 1.

The estimation of the uncertainty of the basal surface of each Formation is the issue that will be explored here.



**Fig.1:** Digital Geological Model (DGM) of The Netherlands.

## 2. Geological modeling and *a-priori* expertise

The interpolation and modeling of the Digital Geological Model is carried out by using the interpreted borehole data as primary source of information, together with the depositional extent and the faults that influence the base of each Formation.

To be able to produce plausible geological models, we use geological expertise to guide the interpolation and modeling. Some examples are provided in the following paragraphs.

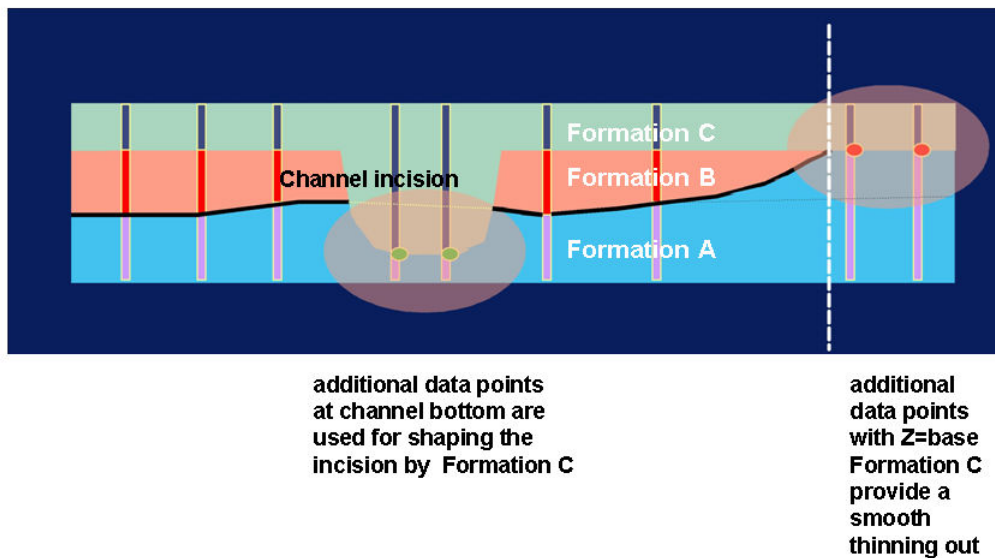
To assist the modeling of bowl-shaped basins that are formed by glacial scouring, we create a trend surface that conforms with the general ideas of the geological experts about the form of the basin. Based on drillings, expert knowledge of the local situation and general knowledge about depth / width ratio's, a shape is created that is used as a trend surface in the modeling.

Deeper marine formations show a general trend of deepening to the North-West. Especially in the deeper parts of the marine sequence, only a few borehole data are available and the trend is used to guide the interpolation to produce plausible geological surfaces.

At the border of the depositional extent, the base of some Formations pinch out. The amount of borehole data is often not sufficient to model this correctly. Therefore, in areas where the pinching out (according to the geological expert) occurs, additional data is added to guide the interpolation to produce correct geological surfaces. Also, additional points are sometimes added to steer the interpolation into the desired result, see Fig. 2.

In the case of faults, the interpolation is prevented “to look around the corner” by adding additional datapoints at both ends of the fault.

By using all these additional expert knowledge, the meaning of the kriging variance is questionable. It is often not possible to quantify the uncertainty that belongs to the geological expert knowledge. Also, the *a-priori* decisions about stationarity and multi Gaussian distribution is not always easy to make, and they are crucial in accepting the kriging variance as a measure of uncertainty.



**Fig.2:** Adding a-priori data to steer the geological modeling.

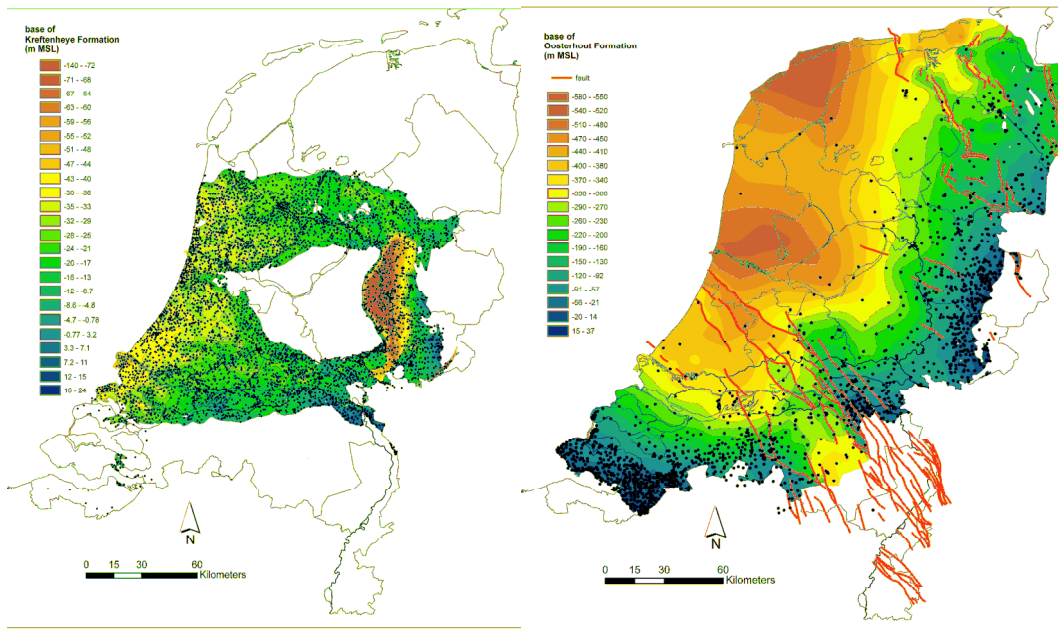
### 3. Uncertainty modeling

To try to incorporate the *a-priori* knowledge in uncertainty estimation, we decided to try a different approach that is based on cross-validation. The general idea of cross-validation is to leave out each data point – one at a time – and estimate the variable at that location with the remaining datapoints.

By comparing the true and the estimated value at the cross-validated location the uncertainty for that location can be assessed.

Besides the cross-validation, we also used a jack-knifing procedure, in which 6% of the dataset was set aside to compare the results of the modeling with. This will provide an idea about the robustness of the cross-validation method: if both methods give comparable results, we have more confidence in the results of the cross-validation method.

We tested both the cross-validation and the jack-knifing on two geological Formations in The Netherlands. The Krefetenheye Formation is a late Pleistocene fluvial deposit, located in the central part of The Netherlands, at a depth of approx 20 m below mean sea level. There are more than 2500 drillings that have reached the base of the Krefetenheye Formation. The other geological surface is the Oosterhout Formation - a near-shore marine deposit that deepens to the north-west – and has only 550 drillings that have reached the base and they are mostly located in the eastern and southern part of the country. The average depth of the base of this Formation ranges from almost at the surface in the east to more than 500m depth in the north-west. In Fig. 3 the depth of the base of these Formations is depicted.



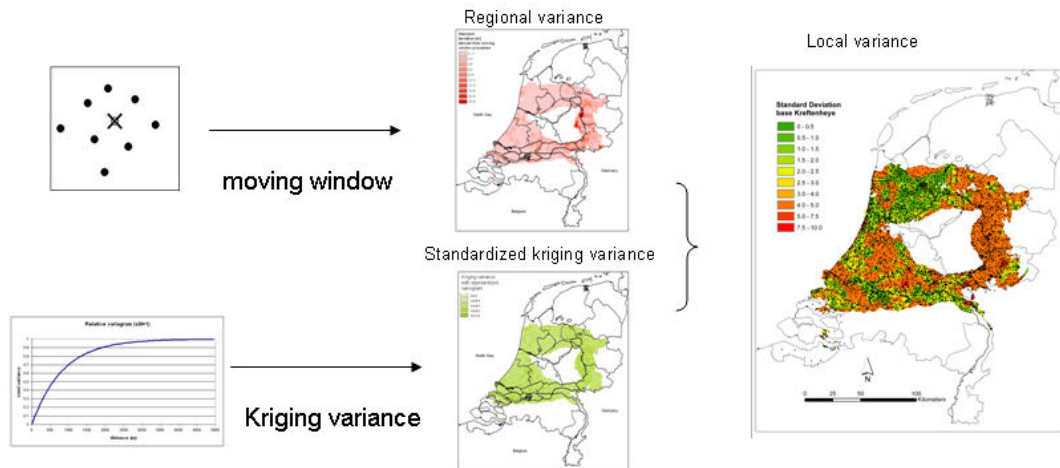
**Fig.3:** Depth of the bases of the Kreftenheye and the Oosterhout Formation.

The cross-validation was carried out for both the Kreftenheye and the Oosterhout Formation. This results in the estimation error for each location. Our objective is to provide maps with uncertainty for the entire area of the respective Formations. To get from uncertainty at a location to uncertainty for each grid-cell, we used a moving window technique. In a moving window, the mean and the variance of the estimation errors were calculated, we called this the regional variance. The size of the moving window and the amount of datapoints in the moving window are crucial to the calculation of the mean and variance in that window. We determined the optimum size of the moving window by trying out several sizes and comparing the mean error and standard deviation for the different moving window sizes. The size of the moving window with the mean error closest to 0 was chosen. This amounted to 4 km for the Kreftenheye and 40 km for the Oosterhout Formation.

Besides this, we demanded that at least 15 datapoints should be present in the moving window to get an estimate of the mean error and the variance. In this way, for each gridcell location, the variance for the moving window, centered on the gridcell was calculated and stored as “regional variance”.

The regional variance can be regarded as the average deviation from the true value in a window centered on the gridcell. This does not account for any clustering or closeness to the datapoints. We used a variogram based approach to account for this. Kriging is used to obtain the kriging variance, and the kriging variance is used as a scaling factor to the regional variance. The variogram we used had a sill of 1, while the range was the same as was used for the original geological interpolation: 2500m for the Kreftenheye and 5000m for the Oosterhout Formation.

By multiplying the regional variance by the kriging variance we obtained the so called “local variance”, which is thus based on a regional measure of uncertainty, scaled to take into account the data configuration (clustering and distance). The procedure is depicted in Fig. 4



**Fig.4:** Flow chart for calculating local variance.

There are areas in the Netherlands with few data, in which it was not possible to calculate the regional variance. We decided to calculate the variance of the error for all datapoints that are in these areas and assign this value to the gridcells with too few data in the moving window.

The jack-knifing procedure – leaving out 6% of the datapoints - resulted in the estimation of the base of the Kreftenheye and the Oosterhout Formation for those locations that were discarded, and these bases can then be compared to the true value of the base of each Formation

### 3.2. Results

Results are checked and interpreted using cross-validation statistics that are described by Cressie (1993), Cressie and Ver Hoef (2001). Also, graphs showing the accuracy are used, according Goovaerts (2001)

The following statistics are defined:

The mean error (ME) is defined as

$$(1/n_1) \sum_{i=1}^{n_1} \left( \hat{Z}_1(s_{1i}) - Z_1(s_{1i}) \right)$$

The standardized prediction residual (SPR) is:

$$\sum_{i=1}^{n_1} \left( \hat{Z}_1(s_{1i}) - Z_1(s_{1i}) \right) / local\_stdev$$

The root mean square prediction error (RMSPE) is defined as

$$\left\{ \sum_{i=1}^{n_1} (\hat{Z}_1(s_{1i}) - Z_1(s_{1i}))^2 / n_1 \right\}^{1/2}$$

The RMEV is defined as:

$$\left\{ \sum_{i=1}^{n_1} (\text{var}(\hat{Z}_1(s_{1i})) / n_1) \right\}^{1/2}$$

The mean error should be close to 0 and the mean and the standard deviation of the standardized prediction residual should have sample mean 0 and sample standard deviation 1, respectively.

If the estimated prediction variances are correct, then RMEV should be close to RMSPE. In the accuracy plots, the fraction of the datapoints falling in a particular, symmetric, probability interval is compared to the theoretical value. This gives an indication of the prediction errors are valid.

In the accuracy plot, the theoretical probability and the actual proportion in the data are compared to inspect for bias.

For the Kreftenheye Formation, the results are summarized in Table 1 and in Fig. 5. The statistics are looking promising, with most of the results being close to the expected values. We therefore conclude that there is little bias in the prediction errors and that the prediction variances appear to be statistically valid, according to the accuracy plot, Fig. 6. We calculated the same statistics for the jack-knife dataset, and came to the same conclusion as with the cross-validation.

Table 1. Validation statistics for the Kreftenheye Formation

	<i>cross-validation</i>	<i>jack-knife</i>
n	2534	160
ME	0.12	0.27
SPR:mean	0.04	0.06
SPR:stdev	1.1	1.11
RMSPE	2.62	3.99
RMEV	2.41	3.69

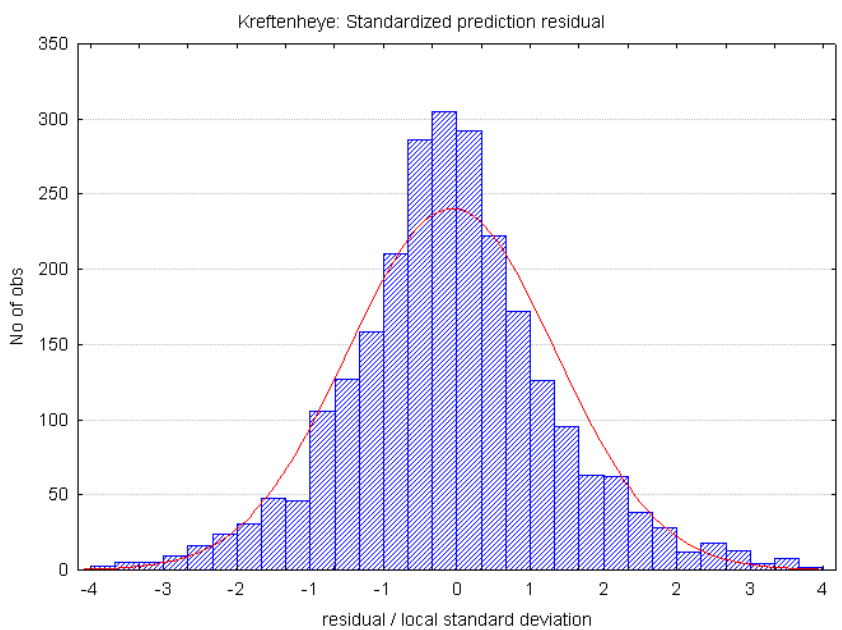


Fig.5: Standardized prediction residuals for the base of the Kreftenheye Formation



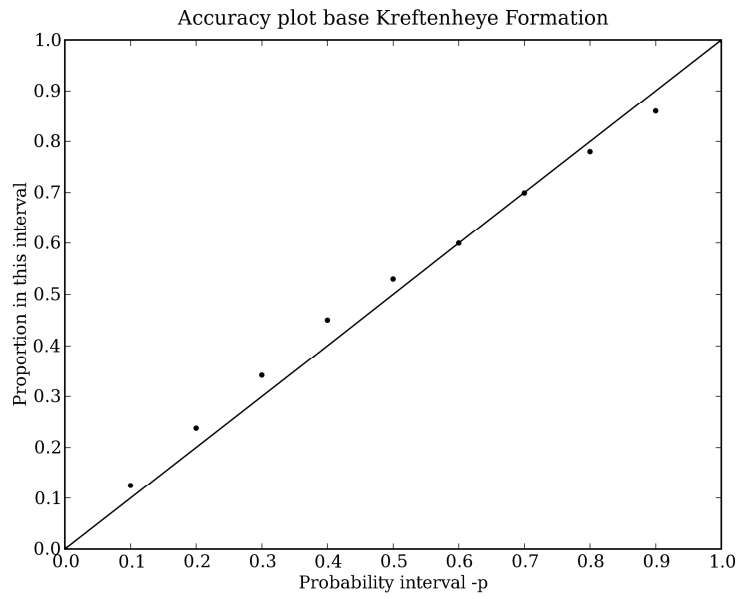


Fig.6: Accuracy plot for the base of the Kreftenheye Formation

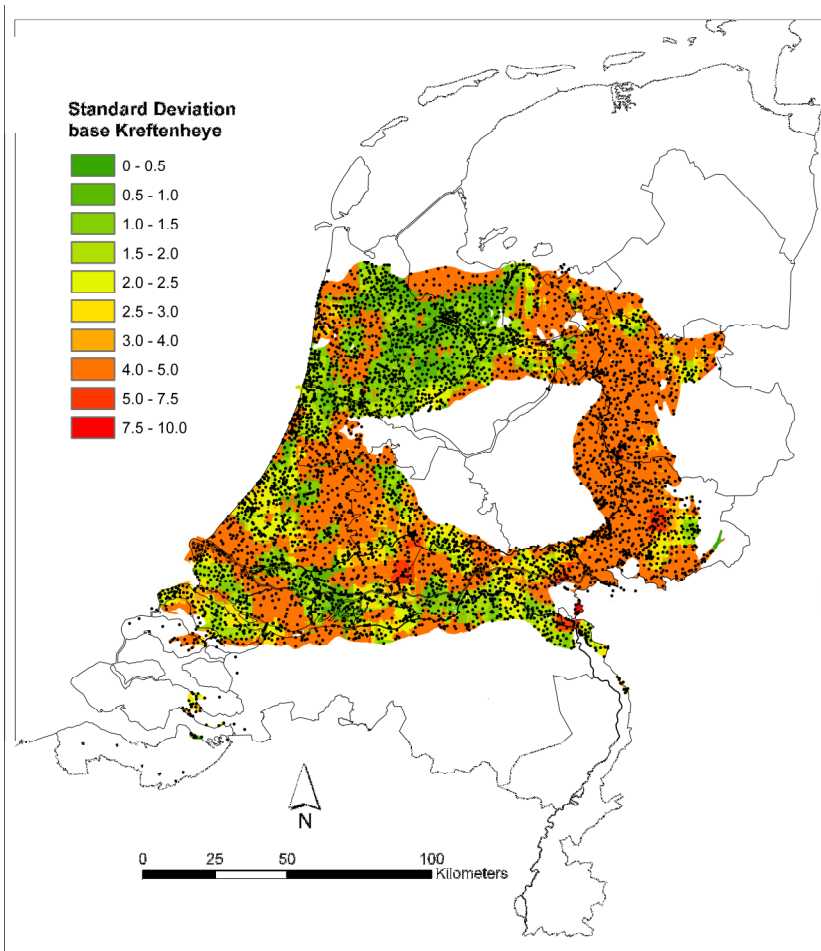


Fig.7: Local standard deviation for the base of the Kreftenheye Formation

We applied the same procedure – cross-validation and jack-knifing - to the Oosterhout Formation, which has a much lesser data density. The results are presented in table 2 and Fig. 8.

Table 2. Validation statistics for the Oosterhout Formation

	<i>cross-validation</i>	<i>jack-knife</i>
n	565	35
ME	-0.15	-0.89
SPR:mean	-0.06	-0.56
SPR:stdev	1.21	4.54
RMSPE	3.36	9.81
RMEV	3.1	2.72

The results of the Oosterhout cross-validation are in line with what we expect: the statistics are close to the expected values, only the standard deviation of the SPR is a bit high. On the other hand, the jack-knifing results are not good: the standard deviation of the SPR and the RMSPE are much too high. We think that the reason for this misbehavior is in

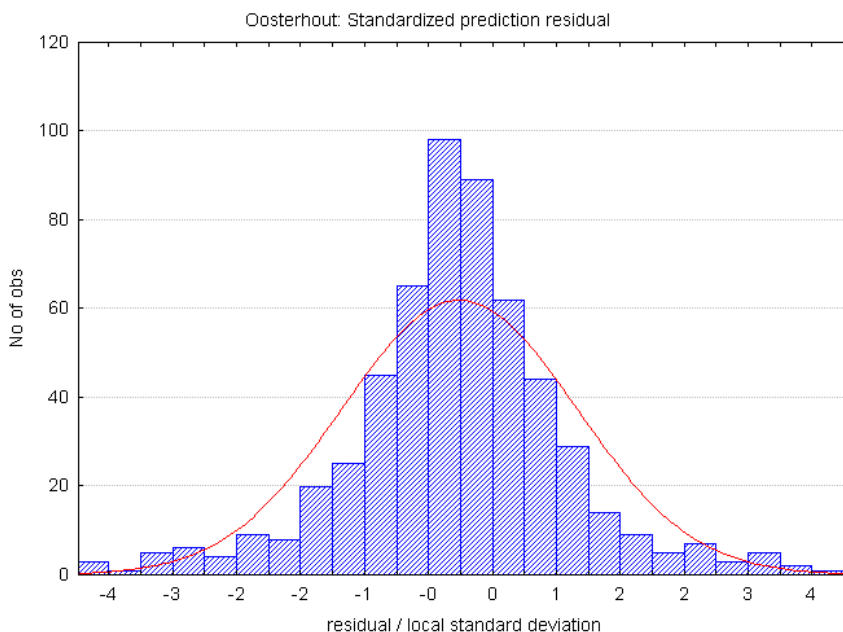
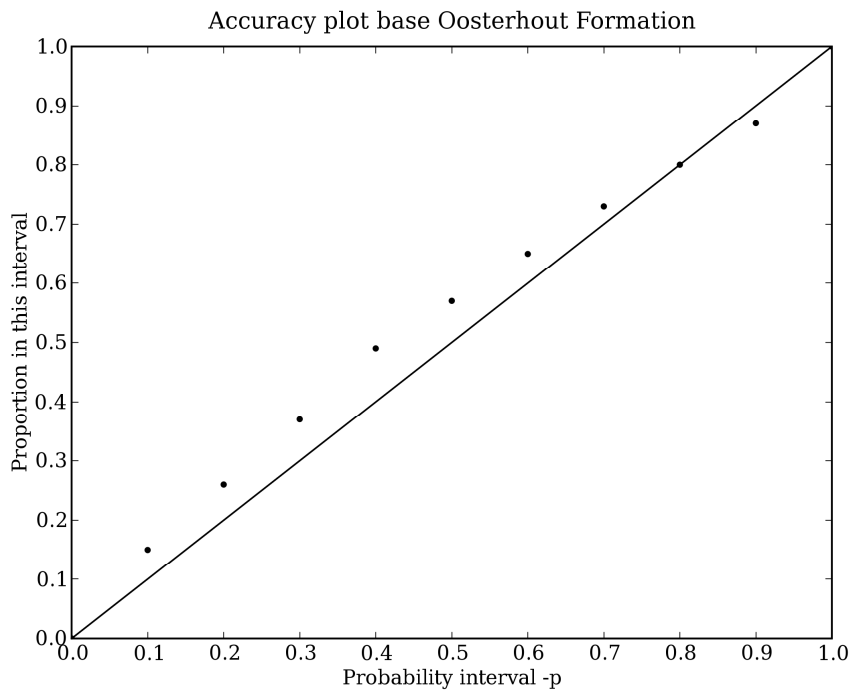


Fig.8: Standardized prediction residuals for the base of the Oosterhout Formation



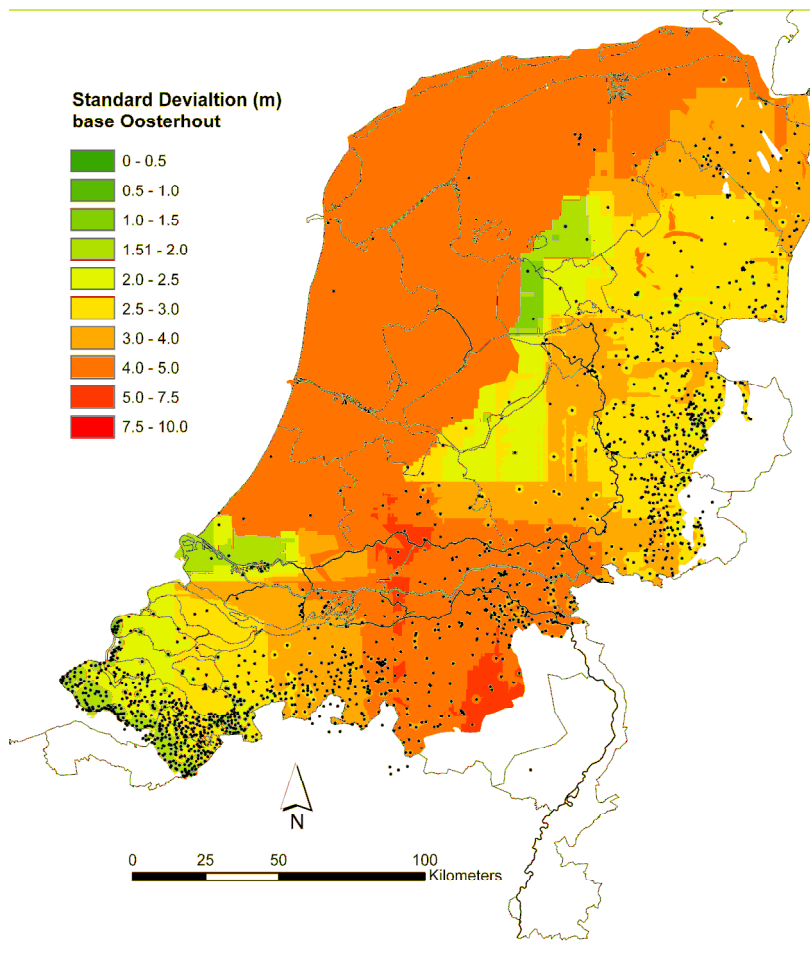
**Fig.9:** Accuracy plot for the base of Oosterhout Formation

the amount of data that is set aside for the jack-knifing. It appears that we need all the data we have to model the Oosterhout Formation and that by leaving out 6% of the datapoints the modeling results in a biased estimate of the uncertainties.

The accuracy plot (Fig. 9) shows a consistent deviation from the expected proportion in 0 – 0.7 probability intervals. The reason for this is not clear, but we think that this behavior is not cause for great concern, since the higher probability intervals are anyway the most interesting for the end-user.

In Fig. 10 the final results are presented for the Oosterhout Formation.

The large areas with similar standard deviations (especially in the north-west) are caused by too little data, in which case the standard deviation of all the points in these areas was assigned to the entire area.



**Fig.10:** Local standard deviation for the base of the Oosterhout Formation

We used “spatial bootstrapping” to compare the results derived from the cross-validation with the variances from the bootstrapping method.

The same amount of datapoints as that were used in the geological modeling, were “redrilled” in the modeled surface of the Kreftenheye and the Oosterhout Formation. This dataset was subsequently used to model the base of the Formations. By applying this procedure 100 times, with different spatial layout of the “redrilled” data using Monte-Carlo techniques, we calculated 100 different models of the base of the Formations. As expected, the average of the 100 surfaces is similar to the initial surface. On the other hand, the standard deviation appeared to be quite different from the standard deviation calculated with the cross-validation method. Especially in areas where a smooth surface was modeled, the variance of the bootstrapping method appeared to be much smaller than the variance of the cross-validation method. This seems to be an undesirable consequence due to the fact that the geologists have forced the surface to be smooth by applying a smooth trend surface. We therefore discarded the results of the bootstrapping method.

## 5. Conclusion

The quantification of the uncertainty of geological surfaces is important, because it gives the potential user knowledge about the usability and reliability of the geological model.

We used cross-validation and moving-window techniques to assess the magnitude of the uncertainties and derived satisfactory results.

## 5. REFERENCES

- Cressie, N and J. M. Ver Hoef, 2001. Multivariate Geostatistics for Precision Agriculture. [isi.cbs.nl/iamamember/cd2/pdf/295.PDF](http://isi.cbs.nl/iamamember/cd2/pdf/295.PDF)
- Cressie, N., 1993. Statistics for spatial data, Revised Edition. John Wiley & Sons, New York, 900 pp.
- Goovaerts, P., 2001. Geostatistical modeling for uncertainty in soil sciences. *Geoderma* 103, 3-26.