PATTERNS IN WATER QUALITY PRODUCTS OF THE NORTH SEA: VARIOGRAM ANALYSES OF SINGLE AND COMPOUND SEAWIFS CHL & SPM GRIDS

M.A. Eleveld *, H.J. van der Woerd

Institute for Environmental Studies (IVM), Faculty of Earth and Life Sciences (FALW), Vrije Universiteit (VU) Amsterdam,
De Boelelaan1987, NL 1081 HV Amsterdam, the Netherlands

KEY WORDS: Hyperspectral, Ocean Colour, Water Quality Products, Phytoplankton, Total Suspended Matter (TSM), Patterns, Geostatistics, Semi-Variogram

ABSTRACT:

In Europe, increasing political interest drives the investigation into the spatial aspects of water quality measurements. Additionally, unprecedented information on patterns in chlorophyll (CHL) and suspended particulate matter (SPM, also known as total suspended matter TSM) concentrations in coastal and marine ecosystems can be discerned from long-term hyperspectral data from ocean colour sensors. Variograms extracted from a dataset of 287 TSM and 287 CHL Sea-viewing Wide Field-of-view Sensor (SeaWiFS) products from September 1997 until December 2004 were analysed to characterise CHL and TSM patterns in the North Sea. In this paper, the spatial variability in optical characteristics of the North Sea is demonstrated by examples of CHL and TSM quicklooks and variograms for both individual images and seasonal composites, which were made to capture persistent patterns and overcome cloud cover. The range and the form of the variograms vary substantially per parameter (CHL or TSM) and per image. A seasonal differentiation in particularly TSM and to a lesser extent CHL can be identified from the variograms of the composites. This information should be incorporated in the assessment of in situ sampling schemes.

1. INTRODUCTION

In Europe, information on water quality is lately not only scientifically, but also politically important (DGEnv, 2000, 2004a, 2004b). When optically active substances are involved, water quality parameters can be derived from ocean colour remote sensing, notably chlorophyll (CHL), and suspended particulate matter (SPM, also known as total suspended matter TSM). When CHL in Dutch coastal waters exceeds a background concentration of 10 μg/l, concerns increase over potential eutrophication. Attention for TSM increases in the case of human interaction with the sediment, e.g., during dredging, or the Mainport Development of Rotterdam (Tweede Maasvlakte), because of, a.o., changes in transparency and possible consequences for the marine ecosystem (DGEnv, 2004a & b).

The project ToRSMoN (Towards Remote Sensing based Monitoring of the North Sea) aims at investigating the possibilities of remote sensing (RS) for monitoring, as an addition to regular in situ sampling efforts, i.e. the Dutch MWTL measurements (Rijkswaterstaat, 2006a). This paper aims to explore the use of geostatistics to describe patterns in parameters (CHL & TSM) derived from remote sensing. Similar work has been performed by Curran (1988). However, his work was mostly based on a limited number of land cover maps from only a few RS images, whereas in this study we deal with multiple grids of water quality parameters from many RS images. In addition, we expect an inherently large spatio-temporal variability in CHL and TSM concentrations. The North Sea is a highly dynamic coastal sea where large-scale circulation, tidal currents and riverine fresh water inputs mix. The patterns will also be disturbed (masked) by clouds. Therefore a new methodology had to be developed to study these patterns in CHL and TSM concentrations from remote sensing through the study of variograms. But first a short background to the derivation of water quality parameters from remote sensing and an introduction to variograms are given in the next section.

2. THEORETICAL BACKGROUND

2.1 Deriving water quality parameters from remote sensing

Gordon et al.'s (1975) approximation of the radiative transfer model predicts subsurface irradiance reflectance *R(0-)* (that can be derived a.o. from remote sensing) as a function of the inherent optical properties (IOPs) absorption (a) and backscatter (b_b). The coefficient f can vary due to solar angle, scattering at a certain angle relative to total scattering (scattering phase function), and viewing geometry.

Absorption and backscatter of natural water can be related to the optical properties of water (w) and its optically active constituents chlorophyll (*CHL*), Total Suspended matter (*TSM*), and Coloured Dissolved Organic Matter (*CDOM*). Absorption and backscattering are linear functions of the concentrations of the constituents, which allows defining Specific Inherent Optical Properties (a^* and b_b^*). This can also be written as:

$$R(0^-) = f \cdot \frac{b_{b,w} + b^*_{b,TSM} TSM}{a_w + b_{b,w} + a^*_{CHL} CHL + (a^*_{TSM} + b^*_{b,TSM}) TSM + a^*_{CDOM} CDOM} \quad (1)$$

* Corresponding author: marieke.eleveld@ivm.vu.nl

The SeaWiFS ocean colour sensor covers the spectrum in nine narrow optical bands between 400 and 900 nm. From this information a reliable atmospheric correction and air-water interface correction can be derived to provide R(0-) (equation 1) in 7 spectral bands. CHL, TSM and CDOM can be retrieved by inversion of Equation 1 if the wavelength-dependent specific absorption (a^*) and backscatter (b_b^*) are known from in-situ measurements.

The derived POWERS TSM algorithm was validated with MWTL measurements in Pasterkamp et al., 2005.

2.2 Variogram analysis

A semivariogram (in short also called variogram) is a description of spatial variance. Semivariance is based on the common notion that the value of two points closer to each other, are likely to be more similar than when further apart. Following Webster (1985) an estimate of semivariance can formally be described as:

$$\hat{\gamma}(\vec{h}) = \frac{1}{2n(\vec{h})} \sum_{i=1}^{n(\vec{h})} \{z(\vec{x}_i) - z(\vec{x}_i + \vec{h})\}^2 \quad (2)$$

$\hat{\gamma}$ is the estimated semivariance, , $\vec{h}$ is the lag distance, n is the number of observations, z is the value, $\vec{x}_i$ is location.

The variogram is the plot of semivariance against distance between point pairs (lag distance) (Figure 1). A variogram can contain the following information through its main components or features:

- Nugget variance: a non-zero value for γ when h approaches zero. This is caused by various sources of unexplained error (e.g. measurement error);
- Sill: for large values of h the variogram levels out; there no longer is any correlation between data points. This should be equal to the variance of the data set;
- ·Range: the value of h where the sill occurs (or 95% of the value of the sill).

Other characteristics of a variogram are:

- In general, 30 or more pairs per point are needed to generate a reasonable sample variogram;
- The most important part of a variogram is its shape near the origin.

For more information some standard textbooks such as Isaaks & Srivastava (1989) and Davis (2002) are recommended.

3. DATA AND METHOD

3.1 Data

We were provided with 574 WADI XML files (Rijkswaterstaat, 2006b) containing either CHL (in µg/l) or TSM (in mg/l) values plus meta-information. Grid cells representing clouds or land are present within the CHL and TSM data as the value -9999. The CHL and TSM files had resulted from previous processing of 287 SeaWiFS images with the ARGOSS empirical CHL algorithm, and IVM's POWERS TSM algorithm (Van der Woerd, & Pasterkamp, 2004), respectively.

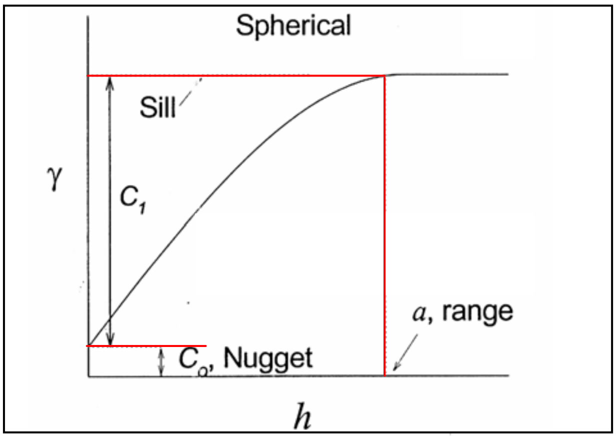

Figure 1. Schematic semivariogram

3.2 Method

Matlab (The Mathworks, 2006) programs were created to convert these XML files to mat files. Quicklooks were produced for fast visualisation, and random sampling was used to extract CHL and TSM values from the North Sea area. In the sampling, 3000 random X- and 3000 random Y-values were generated, masks were created to exclude land, clouds, other seas, and Lake IJssel (IJsselmeer), and finally the corresponding CHL and TSM values were extracted in comma separated (.csv) files.

After exploring the individual results, and based on Campbell's (1995) observations of a log normal distribution of bio-optical variability in the sea, a program was made to calculate bi-monthly geometric means. For CHL, these covered the period of adverse phytoplankton growing conditions in December (of the year before) and January, and blooms in April and May. For TSM, images from the months of quiet conditions (March and April) and with frequent stormy periods (September and October) were pooled (Eleveld et al., 2004). Subsequently, our program to generate quicklooks and sample the North Sea was re-applied to these grids of bi-monthly geometric means.

From the samples, several experimental variograms were produced using Surfer (Golden Software, 2002). To standardise the variograms, the following settings were chosen: a maximum lag distance of 300 km, and a lag width of 10 km. These values were within the system boundaries for the southern North Sea system, more or less cover the Netherlands Continental Shelf (NCS), reach beyond the most seaward MWTL monitoring station (at 235 km offshore of Terschelling) (Figure 2), and do not exceed or conflict with the standard settings calculated by the Surfer Software which vary around 320 km lag distance, and a 12 km lag width. Despite these standardisation efforts, clouds are an external factor that can cause differences in spread and number of samples, particularly in individual images.

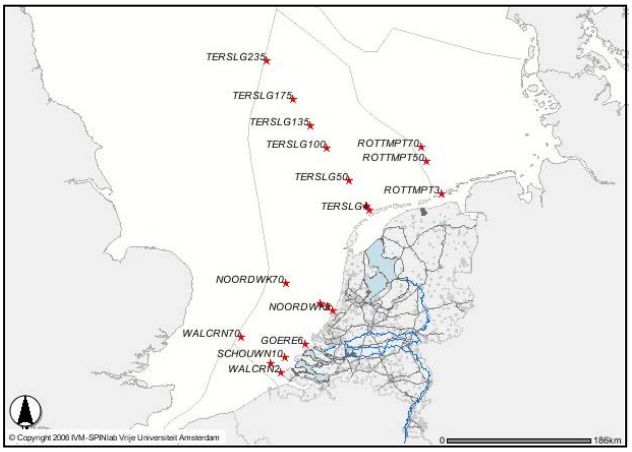

Figure 2. Distances on the North Sea (scale bar on the lower right indicates 186 km), and MWTL monitoring stations in the Netherlands Continental Shelf (NCP)
(source: IVM-SPINlab, 2006)

4. RESULTS

4.1 Results from an analysis of several individual images CHL & TSM

4.1.1 SeaWiFS image of 25 September 1997

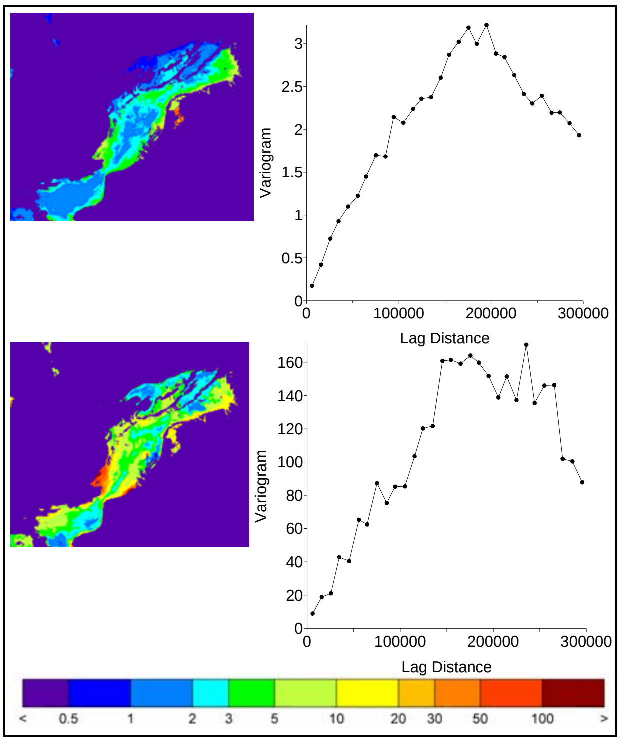

Figure 3. SeaWiFS image of 25 September 1997 (S1997268121448).
(a) Quicklook and (b) experimental variogram of S1997268121448_CHL_empirical.
(c) Quicklook and (d) experimental variogram of S1997268121448_TSM_POWERS.
(a) indicates upper left, (b) upper right, (c) lower left, and (d) lower right.
The scale bar shows values for CHL in µg/l, for TSM in mg/l.

Figure 3a shows a quicklook of a SeaWiFS CHL product S1997268121448_CHL_empirical, where S indicates SeaWiFS, year is 1997, day number is 268, time is 12:14:48 UTC, parameter is CHL and algorithm is ARGOSS-empirical). The image shows two gradients in CHL on both the UK and mainland side, divided by a NE-SW directed central axis with low concentrations. The sampling consisted of 552 points, and the NW part of the Southern North Sea was under-sampled because of clouds (their value of -9999, causes their appearance in the lowest class (< 0.5 µg/l) indicated in purple). Figure 3b gives the matching variogram: if the lag is less than ca 175 km, the values of the points can be used to estimate the others within this range. If the points are more than 175 km apart, they are independent.

Figure 3c and d give results for TSM, obtained with IVM's POWERS algorithm. The image shows steep gradients over small distances, i.e., from class >100 to class 2-3 mg/l TSM over ca 100 km, but also contains small scale patterns. Sampling (number and distribution of points) was the same as for CHL. The variogram shows a range of correlation (spatial dependence) of ca. 150 km. Observations that were more than 150 km apart are spatially independent. (Variance (sill) in Figure 3d sééms much higher than in 3b, but the units CHL in µg/l and TSM in mg/l should be kept in mind.)

This image shows a typical pattern for the North Sea: CHL high along the coastlines (where nutrients are available) and decreases seawards, TSM is high along shallow areas (along the coastlines) and low along deeper areas (seawards).

4.1.2 SeaWiFS image of 15 May 1998

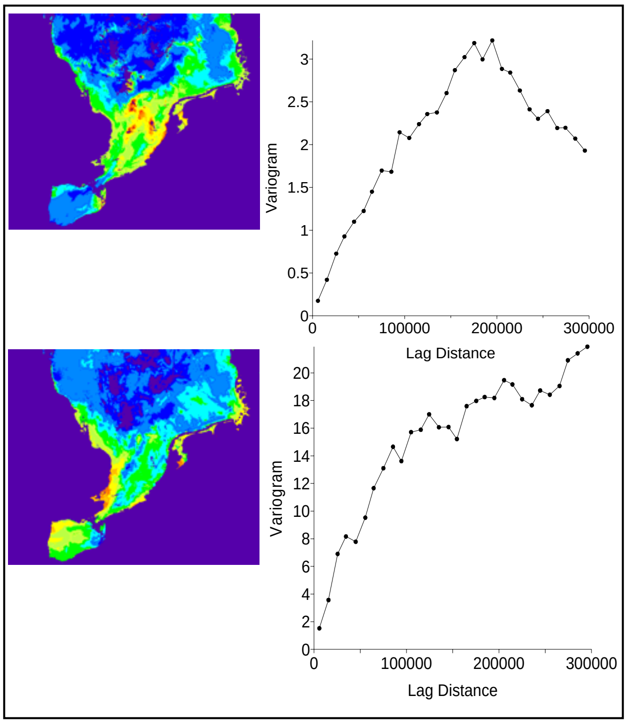

Figure 4. SeaWiFS image of 15 May 1998 (S1998135123234).
(a) Quicklook, and (b) experimental variogram of S1998135123234_CHL_empirical.
(c) Quicklook and (d) experimental variogram of S1998135123234 _TSM_POWERS.

(See Figure 3 for scale bar.)

In Figure 4a multiple algal blooms can be perceived off the UK East Anglian and Dutch North-Holland coast, including in the central part. The sampling consisted of 962 points; there were clouds in the centre of the Central North Sea. There is a strong correlation at lags < 50 km. Figure 4c shows a steeper gradient on the UK coast than on the continental margin. Range of correlation (spatial dependence) at lags <175 km.

These results show lots of activity going on in the North Sea. In May we typically expect algal spring blooms in the North Sea, but TSM also is relatively high. Usually TSM concentrations are low because of low resuspsension under moderate wind conditions (Eleveld et al, 2004).

4.1.3 SeaWiFS image of 24 September 2003

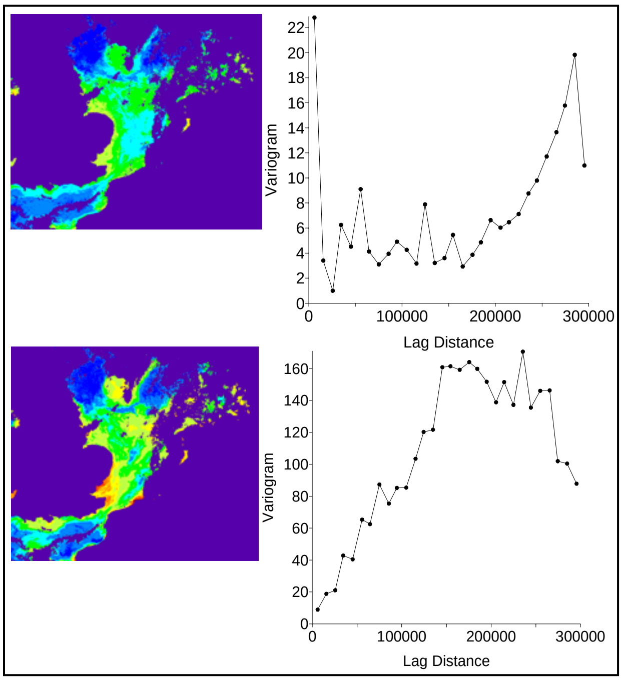

Figure 5. SeaWiFS image of 24 September 2003 (S2003267124041)
(a) Quicklook and (b) experimental variogram of S2003267124041_CHL_empirical.
(c) Quicklook and (d) experimental variogram of S2003267124041_TSM_POWERS.
(See Figure 3 for scale bar.)

Figure 5a shows a semi-circular pattern, with a weak gradient from the Thames and East-Anglia westward. The sampling consisted of 660 points with the cloudy NE relatively under-sampled. The variogram shows no spatial dependence, and contains some relatively high values for semivariance when compared to other variograms of CHL. The latter could be caused by disturbance from cloud edges or high aerosol loading that caused the atmospheric correction in the pre-processing to fail. The nugget effect indicates noise (random variance) caused by measurement error and small-scale spatial variability. Figure 5c is also partly erratic with clouds possibly disturbing large-scale patterns. The variogram increases gradually with lag distance; measurements are related even when far apart, up to ca 150 km.

In this image TSM has more outspoken large-scale patterns.

4.2 Results from an analysis of various CHL & TSM composites

4.2.1 Bi-monthly geometric mean CHL during inhibited algal growth and bloom conditions in 2003

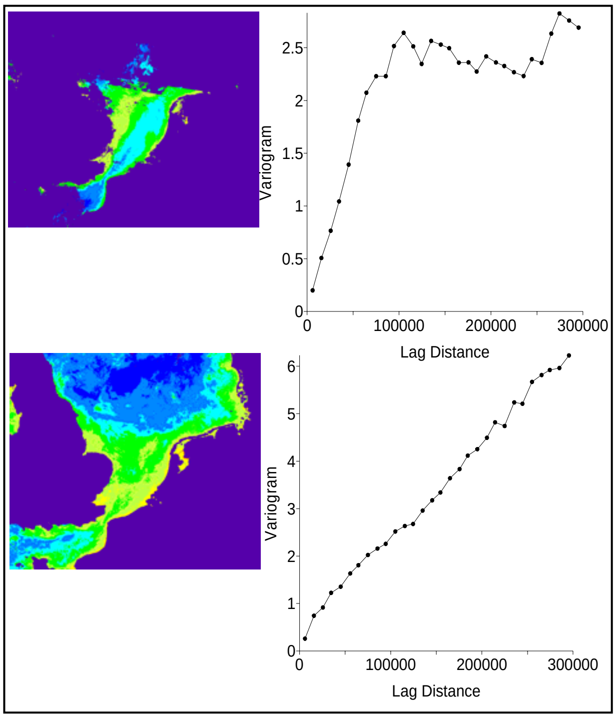

Figure 6. Bi-monthly geometric mean CHL.
(a) Quicklook and (b) experimental variogram during winter (Dec 2002-Jan 2003), a period of inhibited algal growth.
(c) Quicklook and (d) experimental variogram during spring (Apr-May 2003), a period of bloom conditions.
(See Figure 3 for scale bar.)

Figure 6a shows a quicklook of the geometric mean CHL during a period with inhibited algal growth, based on 5 products. The variogram (Figure 6b) is based on 350 samples. The solar angle is so low that only part of the scene is acquired. Consequently, the variogram only described the southern North Sea, Dover Strait and Channel. It approximates a spherical model, with a range of ca 90 km. Figure 6c shows a quicklook of the geometric mean of CHL during a period of spring bloom conditions, based on 12 products. In this case, the variogram is based on 1210 samples, and covers also the Central North Sea. Semi-variance increases linearly with lag distance.

4.2.2 Bi-monthly geometric mean TSM during quiet and stormy conditions in 2003

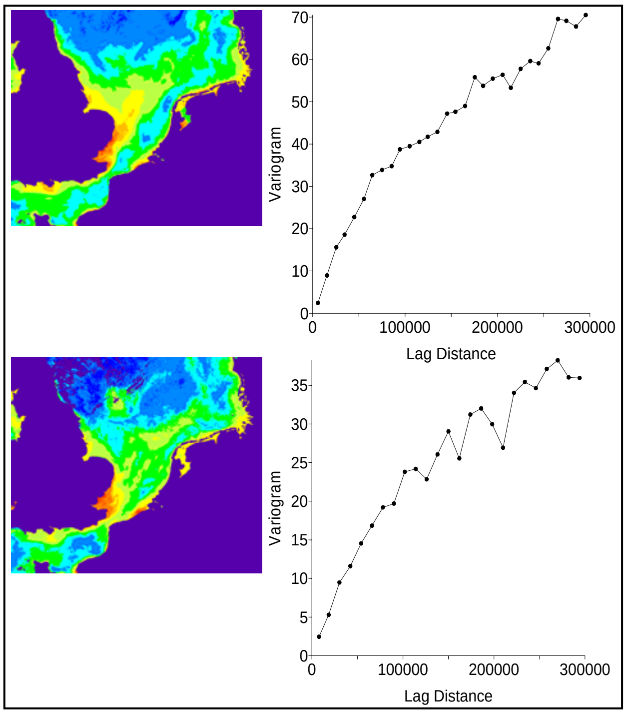

Figure 7. Bi-monthly geometric mean TSM.
(a) Quicklook and (b) experimental variogram for Mar-Apr, a period of quiet conditions.
(c) Quicklook and (d) experimental variogram for Sept-Oct, a period of stormy conditions.
(See Figure 3 for scale bar.)

Figure 7a show the bi-monthly geometric mean of TSM for a quiet spring period. The product is based on 13 TSM products. An area of high concentrations at the surface extends from the UK East Anglian coast to just off the Danish coast. Low TSM concentrations (< 3 mg/l) can be found in a strip off the French (Channel), Dutch and German coast, whereas the nearshore zone has higher TSM concentrations again. The matching variogram is shown in Figure 7b. It is based on 1203 samples. The variogram follows a logarithmic model: semivariance increases with lag distance, levelling out at larger distances (> 100 km).

Figure 7c shows the bi-monthly geometric mean of TSM for a stormy period (based on 11 single images). In this period, TSM concentrations are generally higher than 3 mg/l for the entire central North Sea area. In Figure 7d, the variogram is given, based on 1132 samples. The form of the variogram is, again, logarithmic, but semi-variance is smaller than that given in Figure 7b (both are in the same mg/l, units).

5. DISCUSSION

The presented results showed an independent statistical description of the data plus an interpretation focusing on the distance between monitoring points. This is an interpretation of the experimental variogram. The exact range (distance) depends on the model used to fit the variogram, which is a subjective (user interpretation) step that is taken when using the variogram for interpolation.

6. CONCLUSIONS & OUTLOOK

Geostatistics can be used to describe the patterns in the CHL and TSM products, which result from in spatial variability in optical properties (Van der Woerd et al, 2004). Large variation in spatial correlation was perceived between different parameters (CHL and TSM) of single images, and between single images of different dates. Composites exhibit a long range of correlation, and show clear differences between the seasons. These large-scale patterns can therefore be described with sampling points that are far apart. It could be worthwhile to investigate if results from UK and Dutch in situ sampling efforts could be compared.

Omnidirectional variograms were used in this study, but a first analysis has shown that the geography of the North Sea causes variability to differ most in certain directions. Variogram modelling seems to indicate an anisotropy ratio of 2 and angle of 135º. To validate the descriptive value of the experimental variograms, variogram modelling followed by kriging will be applied. The resulting maps will be compared with the original quicklooks.

REFERENCES

Campbell, J.W., 1995. The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research 100(C7)*, pp. 13,237-13,254.

Davis, JC., 2002. Statistics and data analysis in geology. 3rd ed. Wiley, New York.

DG Env, 2000. Water Framework directive (2000/60/EC) *Official Journal of the European Communities Official Journal* (OJ L 327) http://europa.eu.int/comm/environment/water/water-framework/index_en.html

(Last accessed 5 April 2006)

DG Env, 2004a. Birds Directive (79/409/EEC) CONSLEG: 1979L0409 — 01/05/2004. Office for Official Publications of the European Communities http://europa.eu.int/comm/environment/nature/nature_conservation/eu_nature_legislation/birds_directive/index_en.htm (Last accessed 5 April 2006)

DG Env, 2004b. Habitats Directive (92/43/EEC). CONSLEG: 1992L0043 — 01/05/2004. Office for Official Publications of the European Communities http://europa.eu.int/comm/environment/nature/nature_conservation/eu_nature_legislation/habitats_directive/index_en.htm (Last accessed 5 April 2006)

Eleveld, M.A., Pasterkamp, R. & Van der Woerd, H.J. (2004). A survey of total suspended matter in the southern North Sea based on the 2001 SeaWiFS data. *EARSeL eProceedings*, 3(2), 166-178. CD & URL: http://las.physik.uni-oldenburg.de/eProceedings.

Golden Software, Inc.,2002. Surfer 8. (Software)

Gordon, H.R., Brown, O.B. & Jacobs, M.M., 1975. Computed relationships between the inherent and apparent optical properties of a flat homogeneous ocean. *Applied Optics 14* (2), pp. 417-427.

Isaaks, E. H. & Srivastava, R. M., 1989. Applied geostatistics: An introduction. *Oxford University Press*, New York.

IVM-SPINlab, 2006. WATeRS: A portal for water quality information products from operational remote sensing. http://ivm10.ivm.vu.nl/mapserver/waters/ (Last accessed 5 April 2006)

Pasterkamp, R., Eleveld, M.A. & Van der Woerd, H.J. (2005). Design of single-band sediment algorithms: wavelength considerations session: suspended sediment. Halifax, *8th Conference on Remote Sensing for Marine and Coastal Environments.*

Rijkswaterstaat, 2006a. WADI. http://www.wadi.nl/ (Last accessed 5 April 2006)

Rijkswaterstaat, 2006b. Waterbase. http://www.waterbase.nl (Last accessed 5 April 2006)

The MathWorks, Inc., 2004. MATLAB R2006a. (Software)

Webster, R. 1985: Quantitative spatial analysis of soil in the field. Advances in Soil Science 3, pp. 1-70.

Van der Woerd, H. & Pasterkamp, R., 2004. Mapping of the North Sea turbid coastal waters using SeaWiFS data. Can. J. Remote Sensing 30(1), pp. 44-53.

Van der Woerd, H.J., Pasterkamp, R., Peters, S.W.M. & Eleveld, M.A. (2004). How to deal with spatial variability in bio-optical properties in coastal waters: a case study of CHL-retrieval for the North Sea. Aus. Ocean Optics XVII (OOXVII-2-184). *Proceedings Ocean Optics XVII*, Fremantle, (25-29 Oct. 2004).

ACKNOWLEDGEMENTS

The SeaWiFS CHL & TSM XML files were provided by Rijkswaterstaat - AGI (Geo-information and ICT Department). With this paper, we did not intend to make a statement on the quality of these data-sets, we did intend to illustrate the (geo-) statistical method. We thank the SeaWiFS Project (Code 970.2) and the Distributed Active Archive Center (Code 902) at the Goddard Space Flight Center, Greenbelt, MD 20771, for the production and distribution of the original SeaWiFS data. This study was partly financed by the Dutch national space budget through the National User Support Programme 2001-2005 (NUSP) project Towards Remote Sensing Supported Monitoring of the North Sea (ToRSMoN).