Eolian deformation detection and modeling using airborne laser altimetry.

Roderik Lindenbergh and Ramon Hanssen
Mathematical Geodesy and Positioning
Delft University of Technology
Thijsseweg 11
2629 JA Delft, The Netherlands
Email: r.c.lindenbergh,hanssen@geo.tudelft.nl

***Abstract*— Monitoring of landscapes or sea bottoms by means of laser altimetry or multibeam results in huge amount of data covering the same area in different epochs. Often stable benchmarks are not available in the area covered. We propose a geodetic/geostatistical method to analyze possible deformations in such area out of time series of data. The method is used for a deformation analysis of six consecutive years of laser data covering a dune section on the south-west coast of the island of Texel, the Netherlands.**

I. INTRODUCTION

During the last decade Airborne Laser Altimetry has become available as a tool to obtain topographic elevation data. Due to its relatively low cost it is possible to perform repeated laser surveys on for example annual basis. This enables monitoring of dynamically deforming topographic features such as dune areas or tidal inlets. Often stable benchmarks like hard infrastructure doesn't exist in such areas implying that traditional ways of deformation analysis cannot be applied.

Every monitoring session produces a data set containing points consisting of three coordinates, representing the planar position and the altitude. During the years the (interpolated) altitude may change at certain positions. Such changes may have two possible causes.

1) Measurement and processing errors.
2) Actual surface deformation.

The main problem that we address is how to separate the actual deformation from the errors. We develop models which allow to draw conclusions such as 'at position (x, y) the altitude is changed by at least $\Delta_-(z)$ and at most $\Delta_+(z)$ with probability p.'

In order to solve this problem we set up a procedure that is easily implemented in a computer. During the setup we identify several parameters that have influence on both the conclusions and the quality of the conclusions. How these parameters should be chosen to obtain optimal results is subject of further research. Basically, we link up data from distinct years in order to obtain a 4D spatio-temporal data set. For every single position in the area we try to find an appropriate test corresponding to a model describing the change in altitude as a function of time. The testing is done by increasing dimension: first we test for stability using a $q = 1$ test. If this test fails we continue with $q = 2$ tests.

The procedure is used for analyzing annual laser data from 1996 to 2001 covering a dune section on the south-west coast of the island of Texel.

II. TIME SERIES ANALYSIS

Suppose we are given a series of data sets $D(t_i)$ representing the xyz coordinates of an area A at epochs $t_1, \ldots, t_m$. The following steps are performed in the deformation analysis process: a) Matching, b) Subdivision and Interpolation, c) Premodeling, d)Testing and (re)modeling, and e) Presentation of conclusions.

A. Matching

Especially when analyzing irregular landscapes, like dunes, matching of the data sets of the distinct epochs is important. If the matching step is left out, one could conclude that the topography of the landscape is changing, although in reality just the method of obtaining the data sets might have changed throughout the epochs.

As in [1], we distinguish two simple types of mismatches: planar mismatches, that are shifts in the xy-coordinates of the data sets, and altitude mismatches, shifts in the z-coordinate. These two mismatches can be combined in profile matching. In the following we assume that no reference data or stable objects are available

Altitude mismatches can be overcome by computing the average heights $\overline{z}_i$ of the total area in the distinct epochs. The quantity $\overline{z}_i - \overline{z}_1$ gives a possible correction value for the altitude offset.

Position mismatches can be found by local matching of landscape features. Divide the area in suited subareas and try per subarea to derive line elements like altitude lines or lines of maximal curvature. If one does that for every epoch, the line elements can be matched again.

Profile matching gives a tool for checking the results of applied positions- and/or altitude shifts. For fixed y-value, say, the altitude is plotted as function of the x-coordinate in the different epochs. The resulting profiles can be compared by visual inspection or by means of their relative Hausdorff distance, compare [2].

B. Subdivision and interpolation

In general, simultaneous processing of all data representing the whole area in the distinct epochs will be almost impossible. Moreover, data will probably be delivered per epoch, although we want to compare data from distinct epochs covering the same area. A third problem could be that data in distinct epochs were interpolated to different grids or were not interpolated at all. Therefore we have to perform some subdivision and interpolation steps.

A first approach is to divide the whole area into pairwise non-intersecting squares of, say, $100m \times 100m$. Fix some square S. From every epoch we select those data with positions in and close to S for an interpolation to a regular grid within S that is the same for every epoch. Here we have to choose a grid distance and an interpolation method. As a result we have altitude data available for every epoch for every position in the regular grid.

As an alternative we could allow some gaps in the altitude data for a certain epoch if not enough positions close to the grid point in that epoch are known. Another adaption can be made for the purpose of multi position modeling, where we model and analyze the change in altitude at a group of positions: here a moving window with certain overlap is probably more desirable.

C. Premodeling

The premodeling step consists basically of the obtaining and processing of geophysical and civil-technical information. It can be very helpful to know on forehand what deformation is to be expected in the observed area. One should think about the parameters that can be used to model such expected deformation and about the number of these parameters related to the number of observations in both the temporal and spatial domain. Sometimes it may be convenient to design a a simplified model that still describes the expected deformation in an accurate way.

III. Geostatistical modeling, single position

The mathematical system that we will use to deduce conclusions on type and amount of deformation is that of geostatistics. This means that we will formulate a number of models, of which we think that they might describe the actual deformation. These models will be tested using the available data while incorporating the uncertainties in the data. In this section we will only discuss models that describe the altitude at one position as function of time. More on testing and adjustment theory can be found in [3], [4].

A. Adjusting and testing observations.

After the interpolation step we assume that for every position (x, y) that we consider, a height $h_{t_i}(x, y)$ is given in all epochs t_i for $i = 1, \ldots m$. We restrict ourselves to linear models, that is, the observed variables, in our case the heights, are a linear combination of the model parameters. Moreover, we require that the number of model parameters, n, does not exceed the number of observations, m. This means that we can always describe such model by a rank n model matrix $A \in \mathcal{M}(m, n)$, where $\mathcal{M}(m, n)$ denotes the family of $m \times n$ matrices. Due to the stochastic nature of the data, there will be an error in the observed variables that is modeled by a covariance matrix $Q_h \in \mathcal{M}(m, m)$. Summarized, we have

$$E\{\underline{h}\} = Ax, \qquad D(\underline{h}) = Q_h.$$

where E denotes the expectation and D the dispersion.

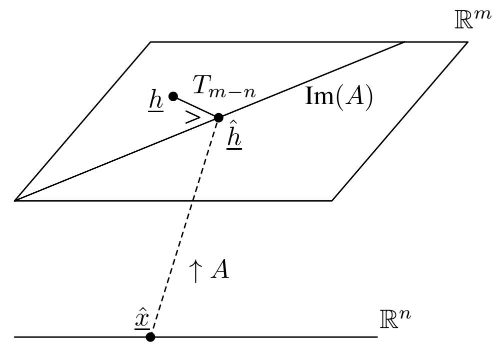

Fig. 1. Adjusting an observation vector.

Now $\text{Im}(A) = \{Ax \colon x \in \mathbb{R}^n\}$ defines a n-dimensional subspace of $\mathbb{R}^m$ that in general will not contain the vector of observed heights $\underline{h}$. A sketch of the situation is given in Figure 1. In order to fit the observation vector $\underline{h}$ in the model it is projected on that point $\hat{\underline{h}} \in$ Im(A) that is most close to $\underline{h}$ with respect to the weighted distance defined by Q_h. As $\hat{\underline{h}} \in$ Im(A) it can be written as $\hat{\underline{h}} = A\hat{\underline{x}}$ for some $\underline{x} \in \mathbb{R}^n$. It's not difficult to verify that

$$\hat{\underline{x}} = (A^T Q_h^{-1} A)^{-1} A^T Q_h^{-1} \underline{h}$$

The weighted distance T_{m-n} between $\underline{h}$ and $\hat{\underline{h}}$ is called the test statistic and is a measure for the adequacy of the used model: if T_{m-n} is small, it is more likely that the model described by A corresponds with the physical reality that makes the heights change (or not). T_{m-n} can be computed in terms of the adjustment vector $\hat{\underline{e}} = \underline{h} - \hat{\underline{h}}$. That is,

$$T_{m-n} = \hat{\underline{e}}^T Q_y^{-1} \hat{\underline{e}}$$

The null hypothesis assumes that the model as described by A gives a good approximation of reality. This is tested by comparing the test statistic to a critical value κ_α: if $T_{m-n} \leq \kappa_\alpha$ the model is said to be accepted, otherwise it is rejected. This critical value depends on a preset level of significance α and on the number $m-n$ of degrees of freedom in the model. The test statistic T_{m-n} has a $\chi^2(m-n, 0)$ distribution if the null hypothesis is true. Now, the critical value is defined as the solution for x of the equation

$$1 - \text{cdf}(\chi^2(m-n, 0), x) = \alpha$$

where cdf is the continuous distribution function. The level of insignificance α indicates the chance that the null hypothesis is rejected although it is true. But if α is made too small, the chance that the null hypothesis is accepted although it is false becomes too big.

B. A $q = 1$ stability test

We discuss some tests now. Throughout the tests, $q = m - n$ indicates the number of parameters in the model corresponding to the test. We start with the most simple test, the stability test $T(h)$, that tests whether the altitude at a given position (x, y) has changed at all. It is a one parameter test, $q = 1$, as the height h is the only parameter in the model.

$$\underline{h} = \begin{pmatrix} h_{t_1} \\ h_{t_2} \\ \vdots \\ h_{t_m} \end{pmatrix}, \quad A = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad x = \begin{pmatrix} h \end{pmatrix}.$$

One could also interpret this test as an outlier test: if one height observation differs a lot from the observations in the other epochs, the test will fail.

C. A $q = 2$ constant velocity test

We can extend the $q = 1$ stability test to a $q = 2$ test by adding a parameter v for the velocity. If we put $v = 0$ we are back at the stability test. The $q = 2$ constant velocity test $T(v, h)$, tests whether the altitude at position (x, y) changes at constant speed, by fitting a line through the (t_i, h_{t_i}) points and computing the weighted least-squares error. The parameter v gives the slope of the line while h gives its offset provided we assume that $t_1 = 0$.

$$\underline{h} = \begin{pmatrix} h_{t_1} \\ h_{t_2} \\ \vdots \\ h_{t_m} \end{pmatrix}, \quad A = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix}, \quad x = \begin{pmatrix} v \\ h \end{pmatrix}.$$

D. A $q = 2$ instantaneous deformation test

We can also extend the $q = 1$ stability test to a $q = 2$ suppletion test. In this case we test whether a suppletion of size s has been added or removed after epoch t_i but before epoch t_{i+1}. So there are in fact $m - 1$ distinct suppletion tests $T^i(h, s)$, one for every $i \in 1, \dots, m - 1$. Note that we get back the stability test in case $s = 0$.

$$\underline{h} = \begin{pmatrix} h_{t_1} \\ \vdots \\ h_{t_i} \\ h_{t_{i+1}} \\ \vdots \\ h_{t_m} \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} h \\ s \end{pmatrix}.$$

E. Other tests

It is of course not difficult to think of other $q = 2$ tests or even tests with more parameters. One could extend for example the $T^2(h, s)$ test that tests for one single suppletion after the second epoch to a test $T^{2,j}(h, s, t)$ that tests for an additional suppletion somewhere after the second epoch. In fact any choice of $A \in \mathcal{M}(m, n)$ for $n \leq m$ defines a test. This shows the importance of the premodeling step: test only for expected deformations.

F. Testing by the number of parameters.

The way we propose to setup and test single position models is illustrated in Figure 2. We start testing at $q = 1$. All positions that pass the stability test are considered to have constant height. The adjusted height from the test is stored. We continue testing only with the positions that failed the test. There are several $q = 2$ tests in the Figure 2. For every $q = 2$ test that we consider, we compute the test statistic. If a position fails every $q = 2$ test, it continues with $q = 3$ tests. In the other case the model which produces the lowest test statistic is chosen as the appropriate model for position (x, y).

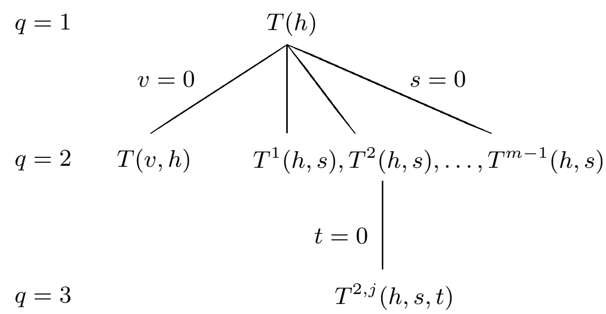

Fig. 2. Testing procedure.

IV. Geostatistical modeling, multi position

Until now we have only considered tests that test deformation at a single position. Spatial correlation of the deformation can be taken into account by testing groups of nearby positions simultaneously. By considering groups of positions, outliers are better detected and deformation can be better classified even if only a small number of epochs is available. Therefore multi positions modeling will be one of our key interests in future.

V. Analyzing Texel data

A. On the data

We will apply some of the methods discussed above on a set of airborne laser altimetry data covering an area in the south-west of Texel, compare [5]. We have data sets of every year between 1996 and 2001, which means that we have six epochs. The quality of the data sets is very diverse. An extensive study of the data is described in [1]. It should be noticed that the data we have worked with are not the original laserdata: in some years third party corrections or interpolations were already made before data delivery. All data sets were manipulated by the Survey Department of Rijkswaterstaat to obtain a good matching with two footpaths for which reference data were available.

B. Processing the Texel data

We consider a subarea of $500m \times 700m$. This subarea is divided into 35 squares of $100m \times 100m$. Any such square S

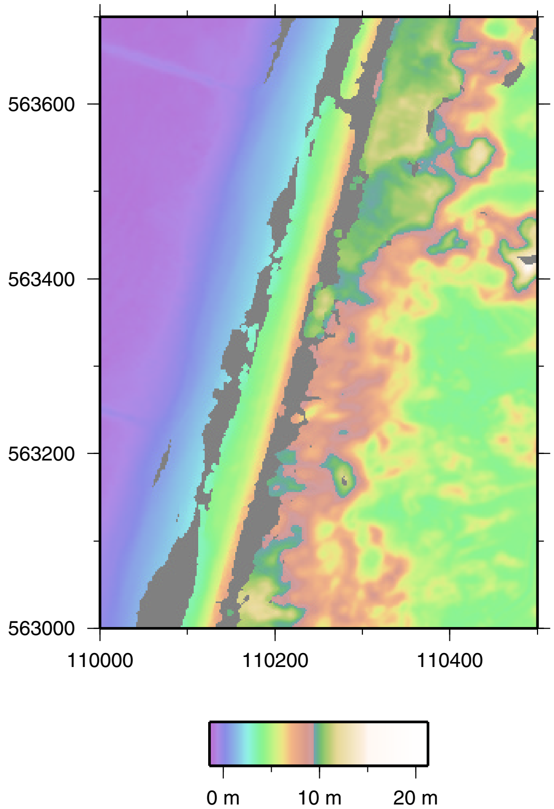

Fig. 3. Altitude at stable positions in meters.

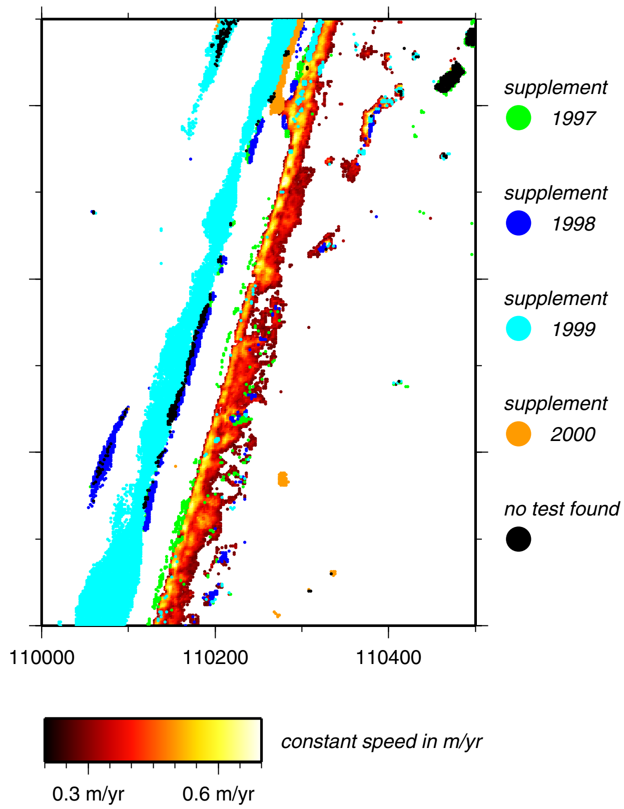

Fig. 4. Detected deformation, or non stable positions.

is analyzed as follows. For every year all available position-altitude data are selected in a $140m \times 140m$ square containing S in the middle. These data are interpolated to a regular $2m$ grid covering S by means of the surface defined by the Delaunay triangulation of the available position data. Now it is possible to collect the altitudes in the distinct years belonging to a single position in the grid. We assume that every point altitude has a variance of $0.3m$ and that no covariance exists between different years. This variance is supposed to be a safe upper bound, compare [1]. As the level of significance for all tests we have chosen $\alpha = 0.005$ giving a critical value of $\kappa_\alpha(q = 1) = 14.8603$ and $\kappa_\alpha(q = 2) = 16.7496$. The processing is finished by going through the $q = 1$ and $q = 2$ tests in Figure 2. For the constant velocity test we silently assume that the data are acquired at regular intervals, that is that the period between two consecutive data acquisitions is approximately 12 months.

C. On the results

For some positions no testing is possible due to lack of data in some year. The adjusted altitudes for positions tested stable are plotted in Figure 3. All other positions are shown in Figure 4. Black positions didn't test positive on any of the $q = 2$ tests. The top of the first row of dunes seems to be increasing as many positions pass the velocity test reporting a positive velocity indicated by the bar at the bottom of Figure 3. In orange a suppletion on the beach between the acquisition of the 2000 and 2001 data is indicated. Some more local changes seem to have occurred at other moments.

VI. Conclusion

In this paper we have presented a framework for the deformation analysis of time series of data covering areas without stable benchmarks. Within this framework, many options are still open: e.g. choices of parameters, models of deformation, methods of interpolation and the number of positions to be tested at once.

Acknowledgment

The authors would like to thank the Survey Department of the Dutch Ministry of Transport, Public Works and Water Management for providing them with the Texel data and for fruitful discussion on the analysis of the data.

References

[1] A. Bollweg, A. Brügelmann, R. De Lange, and V. E., "Correctie laserdata texel," Meetkundige Dienst, Tech. Rep. MD-GAR-2002-7, 2002.
[2] G. Cain, *Introduction to General Topology*. Addison-Wesley, 1994.
[3] P. J. G. Teunissen, *Adjustment theory; an introduction*, 1st ed. Delft: Delft University Press, 2000.
[4] ——, *Testing theory; an introduction*, 1st ed. Delft: Delft University Press, 2000.
[5] An., *Topografische kaart van Nederland, 1:25000, 9 B Texel*. Topografische Dienst., 1998.