Hamburger Beiträge zur Physischen Geographie und Landschaftsökologie

Heft 19

Jürgen Böhner, Thomas Blaschke & Luca Montanarella (Eds.)

SAGA – Seconds Out

Universität Hamburg
Institut für Geographie

2008

SAGA – SECONDS OUT

Edited by:

Prof. Dr. Jürgen Böhner
Universität Hamburg - Institut für Geographie
Bundesstraße 55, 20146 Hamburg (Germany)
Tel.: +49 (0) 40 42838 - 4960
E-Mail: boehner@geowiss.uni-hamburg.de

Prof. Dr. Thomas Blaschke
Research Studios – Studio iSpace
Bräuhausstraße 4b, 5020 Salzburg (Austria)
Tel.: +43 (0) 662 834 602 - 216.
E-Mail: thomas.blaschke@researchstudio.at

Dr. Luca Montanarella
European Commission - Environment Directorate
General Joint Research Centre
Via E. Fermi 1, 21020 Ispra (Italy)
Tel.: +39 03 3278 9111
E-Mail: luca.montanarella@jrc.it

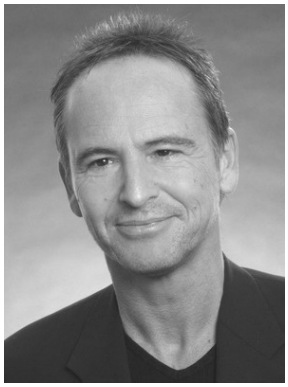
Herausgeber der Reihe:

Prof. Dr. Udo Schickhoff
Universität Hamburg - Institut für Geographie
Bundesstraße 55, 20146 Hamburg (Germany)
Tel.: +49 (0) 40 42838 - 4911
E-Mail: schickhoff@geowiss.uni-hamburg.de

Prof. Dr. Jürgen Böhner
Universität Hamburg - Institut für Geographie
Bundesstraße 55, 20146 Hamburg
Tel.: +49 (0) 40 42838 - 4960
E-Mail: boehner@geowiss.uni-hamburg.de

ISSN: 1866-170X

EDITORIAL



"Seconds Out", the famous second Live Album of GENESIS was taken as the title for this second SAGA volume. Not that we, the SAGA developer team want to compare ourselves with GENESIS, however, the shift from Göttingen to Hamburg and the reorganization of the working group since November 2006 was much of a live performance: loud, fast and sweaty. Two years after the first SAGA volume, launched at the AGIT symposium in 2006, we are pleased to present this second collection of scientific articles, dealing with analysis and modeling applications of the System for Automated Geoscientific Analyses (SAGA). Although the addressed topics and problems are hardly representative for the whole range of possible SAGA applications, we assume that the collection mirrors some advancements in the design and implementation of SAGA. However, you are to judge. Please taste: "*supper's ready*".

Hamburg, June 2008

Jürgen Böhner

The old days of GIS are over, sadly or fortunately, depending on your point of view. Unfortunate for the insiders is the loss of exclusiveness. Positive to most actors is a recent trend towards "free and open GIS" packages", sometimes called FOSS GIS. Of course, software development has to be paid at one point but business models are different. Commercial GIS products tend to be available as modular products. Users purchase the modules they require. In tendency, they acquire GIS much as an interoperable suite of software packages. This book emphasizes one highlight in this emerging field. SAGA developed from a GIS/DEM tool to a powerful but lean software suite. It comprises a range of packages serving the needs for scientists and practitioners. SAGA as well as FOSS GIS in general is pretty much at a stage where a powerful and effective GIS suite can be provided to users who don't have to acquire various pieces of code from various places and assemble them on their systems to make a functional GIS suite. I appreciate that FOSS GIS are undergoing frantic development right now. They are adding to the potential power of GIS.

Salzburg, June 2008

Thomas Blaschke

EDITORIAL

The importance of GIS in Environmental Research has been constantly growing over the past years. The recognition that environmental parameters and processes are spatially explicit has paved the way to a complete new wave of environmental research combined with sophisticated GIS tools and spatial distribution models that have changed many of our previous assumptions concerning the assessment of the environment and the impacts of human activities on it. For all major environmental compartments, like air, water and soil, there are now many georeferenced data sets available that allow for complex modeling of many of the parameters that are of interest for the regular reporting of the status of the European environment. The recent adoption of a legal framework for establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) has further facilitated the harmonization of existing data in the EU Member states and will allow in the future to dramatically improve the analysis of spatial data of relevance to the environmental reporting process. The establishment of a European Spatial Data Infrastructure and the definition of the relevant data centers for the various environmentally relevant themes will put in the hands of policymakers and scientists an enormous wealth of data and information. Good tools and models will be needed in order to make best use of these data. SAGA can certainly contribute to this, as well illustrated in this collection of possible applications of SAGA in Europe as reported in this volume.

Ispra, June 2008

Luca Montanarella

TABLE OF CONTENT

SEGMENTATION FOR OBJECT EXTRACTION OF TREES USING MATLAB AND SAGA

Benjamin Bechtel [1]*, Andre Ringeler* [1] *& Jürgen Böhner* [1]

[1]Institute of Geography – Section Physical Geography – University of Hamburg, Bundesstr. 55, D-20146 Hamburg (Germany)

Abstract: This paper compares alternative region growing algorithms with respect to their suitability for automated extraction of trees from aerial imagery. Two of the algorithms were especially designed for this purpose and were implemented in MATLAB while the third is an all-round region growing implemented in SAGA. Both were combined with a straightforward radiometric model to evaluate the accuracy and tested with an aerial image showing a pine plantation. Besides that special attention was drawn to the question of portability of algorithms. The simple model was found to be adequate in terms of complexity and performance. Regarding the different segmentation algorithms, the more general method in SAGA seamed less suitable for the sophisticated task of object extraction, since it stresses less emphasis on the geometric condition of the extracted regions. In contrast most objects have certain geometric features like compactness and simply connectedness which already should be utilised in the segmentation step. But wherever model assumptions like these are introduced in an algorithm it should be marked clearly to address the difficulty of portability of algorithms under different conditions. Therefore additional parameters were introduced to the specialised algorithms which were found to be useful to adapt to the given dataset.

1 INTRODUCTION

In the last few decades, the growing importance of Geoinformatics at the interface of Geosciences and Computer Sciences has considerably increased the need for high quality and spatial high resolution geodata. One answer to the problem of how to provide vast amounts of geodata at commensurate scales is the automated extraction of objects from remotely sensed data. Therefore, two knowledge based concepts are utilised: the formalisation of knowledge in order to be processed by computers in model-generation and the application of this knowledge for extraction (HINZ 2003).

Particularly trees are a challenging matter of object extraction given their heterogeneity in appearance and scale which has been addressed by several disciplines (including forestry science, remote sensing, geo sciences and computer sciences) for more than two decades. In forestry the main aim is to support forest management and reduce the high costs of conventional forest inventory. Due to the important role of forests in biodiversity, national and international environmental policies require a high demand for environmental monitoring of forest resources.

To date, a wide range of techniques to extract single trees from both, optical (e.g. PINZ 1989, GOUGEON 1995, POLLOCK 1996, BRANDBERG & WALTER 1998, LARSEN & RUDEMO 1997, ERIKSON 2004a) and LIDAR-vector data (HYYPPÄ *et al.* 2001, PERSSON 2001, PERSSON *et al.* 2002, HEURICH *et al.* 2004, MORSDORF *et al.* 2004) or a combination of both (STRAUB 2003) have been proposed, all with promising results but mostly limited to simple forest conditions. Most algorithms, designed and optimised for specific datasets, are distinctly less suitable for performing with different data sets or forest conditions.

Moreover the specific preconditions for the application of certain algorithms and the basic assumptions about the image material are not properly defined. Against this background, ERIKSON *et al.* (2006: 1) stated: "It is most likely that none of today's methods are alone able to handle all types of forests. Comparative studies of different segmentation methods with different types of forests are therefore of importance in order to clarify how much a method is reliable with respect to a certain type of forest. This

knowledge can, for instance, be used to build up an expert system which is supposed to be able to extract individual tree crowns in any type of forests". Therefore, both comparative studies evaluating existing algorithms for different conditions and a more conceptual analysis of the problem of "portability of algorithms" are required. Only if the underlying assumptions are explicitly defined, an expert system could select the best available algorithm for any given image. Furthermore existing algorithms should be improved to address a wider range of problems by introducing suitable control parameters.

In this paper several region-growing-algorithms for image segmentation including the SAGA region growing (BÖHNER *et al.* 2006) are tested with a new aerial image scene of a pine plantation in South Africa. For the extraction of the tree hypothesis a very simple but suitable model is introduced.

2 MATERIAL AND METHODS

Study site and data: The analysed dataset is an aerial image of a pine plantation close to Franschhoek in the Western-Cape-Province of South Africa. The scene shows a part of the La Motte Plantation of the South African Forestry Company Limited Company, which is situated north of the Hottentotten nature reserve at about 33.9268°S/19.0664°E. The image was taken in March 2006 with a digital reflex camera from a microlight aircraft in low flight altitude of about 300m. It was georeferenced and pre-processed with the Leica Photogrammetry Suite and has a ground resolution of about 15cm.

The available image is jpeg-compressed but was chosen as an example of lower radiometric quality to test the adaptability of the algorithms. The contrast is rather low and shows an interesting characteristic. In the section LAMOTTE1 shown in Fig. 2 about 8.9% of the pixels in the red channel have less then 15% of the maximum intensity while it are only less then 0.5% of the pixels in the green and blue channel respectively. Since ideally the shadows should be black throughout the spectrum, the path radiance should be little due to the small flight height and the scattering is potentially higher in the red channel (the mean reflectance in the red channel is higher than in the blue channel due to the red soil) it can be concluded that the radiometric sensitivity of the camera is different in the three channels.

Fig. 1: The analysed scene: pine plantation close to Franschhoek, South Africa.

The image shows about 12 year old *Pinus radiata* trees. The stand structure is rather simple since all trees are of cohort age and separated in equal distances. The crowns are quite circular und compact. They have the appearance of a brighter spot (so called blob) but are only partly separated by shadows. Since the sun is directly reflected from the surface, there are also bright soil-blobs in the image, which pose an additional challenge for the extraction. Therefore the utilised model should at least be able to distinguish trees from ground patches.

From the scene two sections where selected. LAMOTTE1 is from the centre, while LAMOTTE2 is from the upper right and geometrically distorted. Both are shown in Fig. 2 together with the according reference regions, which were extracted manually from the aerial image by an operator.

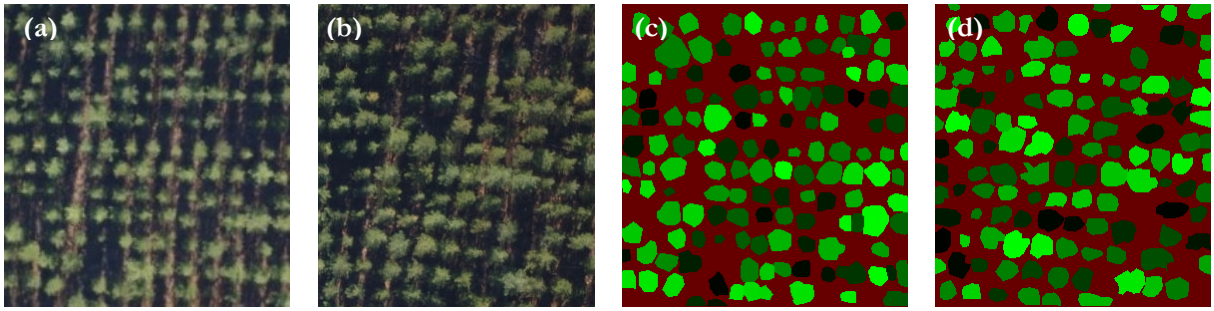


Fig. 2: Sections of the used scene: LAMOTTE1 (a) and LAMOTTE2 (b) and reference regions for LAMOTTE1 (c) and LAMOTTE2 (d).

Strategies in object extraction: Generally two approaches are differentiated in image analysis. Top-down approaches like template matching (POLLOCK 1996, LARSEN & RUDEMO 1997) start at a high level of representation and try to project and decompose a 3-D model into the image. Bottom-up analysis starts by pre-processing the image, finding and grouping features like edges (BRANDBERG & WALTER 1998) or regions (PERSSON *et al.* 2002, STRAUB 2003 and DIEDERSHAGEN *et al.* 2003) and trying to match those with the higher level model in a classification step. Thereby the segmentation of the image and the choice of suitable features are the key tasks, while the classification itself is a standard problem in pattern recognition, covered by a wide range of algorithms.

Region Growing is a class of segmentation procedures in image processing where – starting with single pixels (so called seeds) – image regions are expanded as long as they fulfil a certain homogeneity or similarity criterion, which can contain both radiometric and spatial features. The result is highly dependent of a suitable similarity measure as well as the quality of initial seeds which implies that seed-generation is an important part of the overall segmentation procedure. In the following, three specific region growing algorithms are presented and tested with the above dataset.

Basic Region Growing: The Basic Region Growing (BRG) algorithm is based on the work of ERIKSON (2003, 2004a, ERIKSON & OLOFSSON 2005). Starting from the seeds all neighbour-pixel that meet the similarity-condition expressed in [1] are merged with the adjacent regions:

$$\exp\left(-\frac{1}{2}\left(\frac{(\mathbf{f}_p-\mathbf{f}_s)^t(\mathbf{f}_p-\mathbf{f}_s)}{\sigma_1}+\frac{(\mathbf{r}_p-\mathbf{r}_s)^t(\mathbf{r}_p-\mathbf{r}_s)}{\sigma_2}\right)\right)>\alpha \qquad [1]$$

where $\mathbf{r}_p$ is the position of the candidate pixel in position-space, $\mathbf{f}_p$ is the position in colour-space and $\mathbf{r}_s$ and $\mathbf{f}_s$ are seed positions in position-space and colour-space respectively. The superscript ^{t} indicates vector transposition, σ_1 und σ_2 are variances in colour and position space respectively and α is the threshold for the similarity-criterion. Since the distributions don't need to be normalised α can be chosen freely and is set to 0.15 in ERIKSON & OLOFSSON (2005). For a given threshold the variances can be determined from

formula [1] in the following way: σ_1 controls the maximum allowed distance in colour-space. For a maximum distance of 40% of the peak intensity (which is normalised to 1 in a pre-processing step) in each of the three colour channels in the image, the variance is:

[2] $$\sigma_1 = \sqrt{\frac{3 * 0.4^2}{-2\ln(\alpha)}} = 0.3557$$

Accordingly σ_1 controls the maximum spatial distance δ which is the largest crown radius in the case of spherical-shaped crowns:

[3] $$\sigma_2 = \sqrt{\frac{\delta^2}{-2\ln(\alpha)}}$$

Segmentation via Brownian motion: The segmentation via Brownian motion (BROWNMO) is also based on ERIKSON (2004a) but extended by additional control-parameters. The basic idea is the simulation of the motion of a virtual particle in the image according to the random movement of real particles in liquids or gas. Therefore, starting at the seed point, normalised random vectors are added to the particle's virtual position and the number of visits at every pixel and from every single seed-point is recorded in a Matrix called `numPass`. Since the particle should preferably not leave the crown, the step distance is modified in the following way: The darker the pixel that would be reached without modification the smaller the step. The position r_i after $\boldsymbol{i}$ steps is given by:

[4] $$\mathbf{r_i} = \mathbf{r_{i-1}} + w(\mathbf{I}(\mathbf{r_{i-1}} + \mathbf{z}))\mathbf{z}$$

where I is the image, z a random vector and *w* the weight function.

In this algorithm several assumptions about the radiometric properties of trees in the image are drawn, e.g. that trees are brighter than the background, that the centre of the tree is the brightness peak and that there are dark valleys between the trees existing. Each run stops after `NUM_STEPS` steps and is repeated `NUM_RUN` times. This was introduced as a new feature to cut the impact of accidentally running into adjacent crowns. The effects of the parameters can be seen in an idealised experiment in a one-dimensional homogenous white image in Fig. 4. The number of runs controls the size of the visited region, the number of steps the certainty of the resulting distribution and therefore the quality of the result. Furthermore the probability density is isotropic which means it is just a function of the distance from the seed point, not of the direction. Consequently it is assumed, that the sought object is circularly shaped and that the seed point represents the centre.

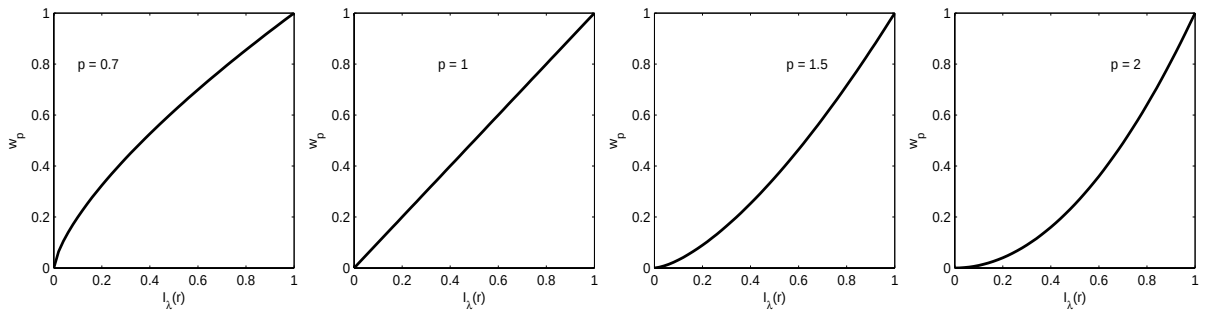


Fig. 3: Weight functions w_p for p = 0.7, 1, 1.5 and 2.

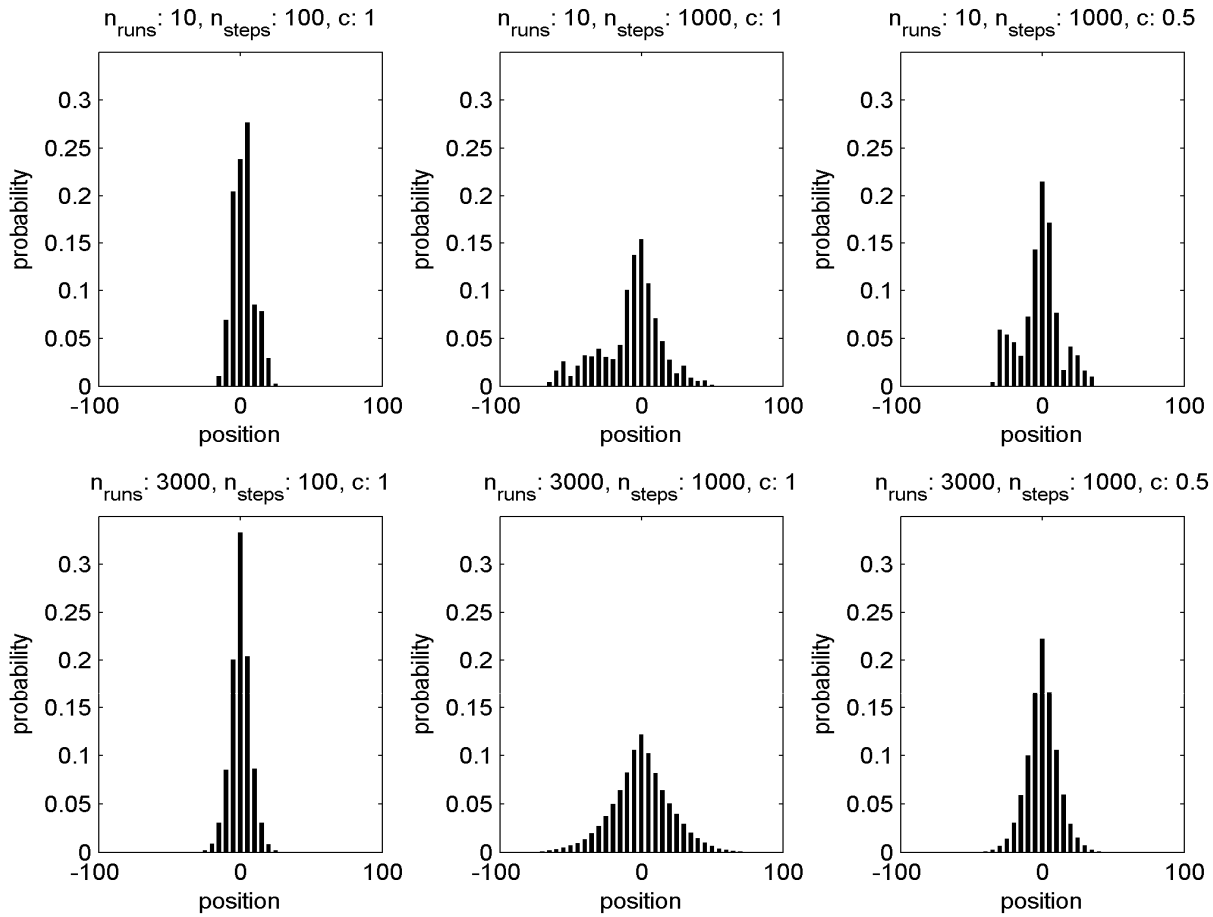


Fig. 4: Relative probabilities of finding a particle at a certain position in an idealised experiment (1D homogenous white image) with different parameter settings. Top: 10 runs, bottom: 3000; Left: 100 steps, middle and right: 1000; left and middle: constant weight factor c=1, right: c= 0.5.

Another way to modify the result in order to increase the portability is the use of different weight functions. For example if due to atmospheric scattering the shadows are not as dark as assumed, a modified weight function could prevent the particle from entering shadow-areas. ERIKSON (2004a) himself uses the grey value of the target pixel (which are normalised to the range [0..1]) multiplied by a constant factor *c* as a weight function:

$$w(\mathbf{I}(\mathbf{r})) = c\, I_\lambda(\mathbf{r}) \qquad [5]$$

Fig. 4 shows the effect of a constant factor of 0.5 at the right side compared to 1 in the middle. But generally any monotonically increasing function of the type:

$$w : \mathbf{I}(\mathbf{r}) \mapsto [0:1] \qquad [6]$$

would also be suitable as a weight function. Several different weight functions where tested, e.g. linear combinations of different colour channels:

$$w(\mathbf{I}(\mathbf{r})) = \frac{1}{n}\sum_{\lambda=1}^{\mathbf{n}} a_\lambda I_\lambda(\mathbf{r}), \text{ with } \sum_\lambda a_\lambda = 1 \qquad [7]$$

(which includes the arithmetic mean of the colour channels red, green and blue for $a_R = a_G = a_B = 1/3$). Furthermore different powers of the functions were tested. For a single channel the according weight function is:

[8] $$w_p(\mathbf{I}(\mathbf{r})) = (I_\lambda(\mathbf{r}))^p$$

The effect is shown in Fig. 3: for a higher power it becomes less likely, that the particle runs into darker areas, but at the same time the mean step width (as well as the segmented area) becomes smaller similar to the constant factor in Fig. 4.

The actual region growing is performed in a second step after the generation of `numPass`. The region of every seed is initialised with the largest circle of pixels that have been visited in the `numPass` image. Then all neighbouring pixels (according to 4-neighbourhood) are administrated in a list. The list is sorted and processed in order of the numPass values so that the pixel with the most visits is added to the region first. Afterwards the list is updated with the new neighbour pixels which are not member of a region so far and have a sufficient `numPass` value. The algorithm terminates when the candidate list is empty.

SAGA Seeded Region Growing (SRG): For segmentation of multispectral sensor data in feature space (also property space) a fast seeded simultaneous region growing algorithm is implemented in SAGA. The starting seeds for the algorithm can be defined manually, or determined automatically by the SAGA maximum representativeness analysis module as described in BÖHNER *et al.* (2006).

The recursive region growing process is presented schematically in Fig. 5. Starting with the seeds as an initial set of clusters, a list of border pixels $B_1 \cdots B_n$ containing their 4-neighbours, the pixel values in all dimensions of feature space $x_1 \cdots x_n$ as well as their adjoining clusters is built. For every border point B_i and every adjacent cluster the maximum Euclidean distance in feature space to any other pixel in the cluster d_i is computed and the list is sorted by increasing distances. This complete linkage strategy guarantees spectrally homogenous clusters and avoids the risk of centroid linking, where the clusters radiometric features can alter during processing. In every step the first pixel in the list is associated with the cluster; the boundary pixel list is updated and resorted. For a more detailed description see BÖHNER *et al.* (2006).

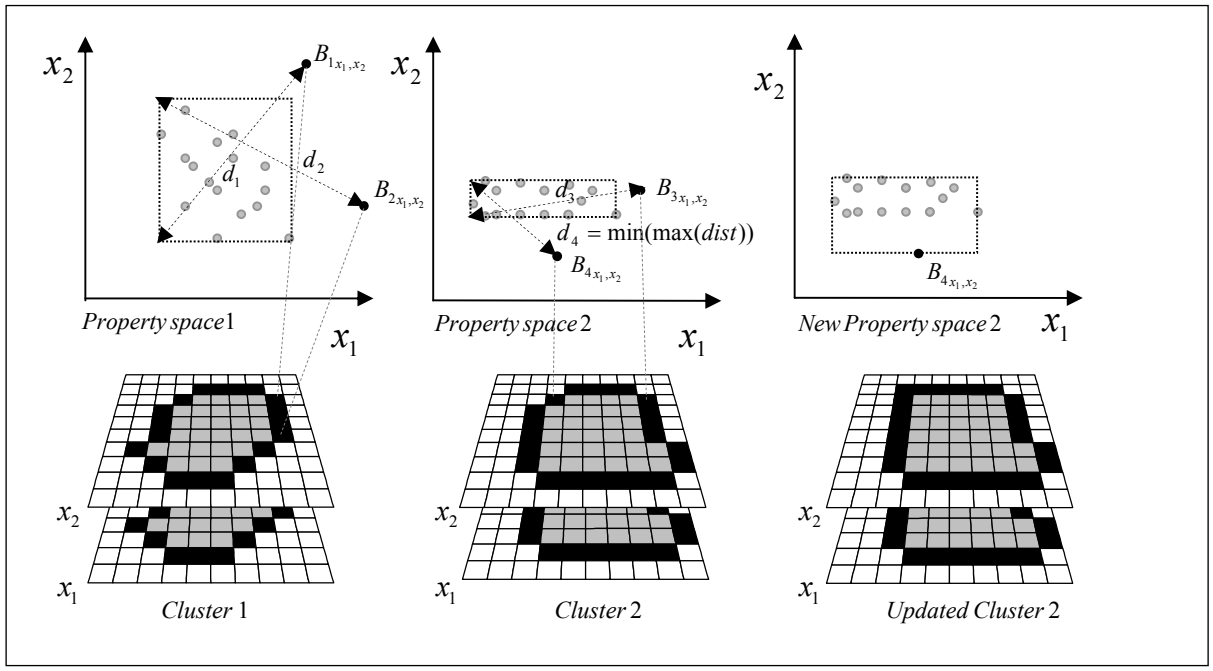


Fig. 5: Schematic description of the linkage in SAGA Seeded Region Growing.

Since a region merging algorithm for post processing is not implemented yet, the results of the region growing process tend to over-segmentation. While the algorithm is designed to produce compact regions in feature space it doesn't necessarily produce compact regions in position space, which can be connected in any possible way.

Evaluation: to compare the results certain regions are selected as a valid tree hypothesis in the following way. Using the module `bbFeatureExtraction`, in a first step, several features of the regions are extracted, including the geometric features AREA, SOLIDITY, CIRCLENESS and RADIUS, which refer to the whole region and the radiometric features, which are mean values of all (BRIGHT and DOA) or a subselection (BOUND_BRIGHT) of the pixels of a region (for detailed description of the features see BECHTEL 2007: 90).

The model should be appropriate for the complexness of the given situation which means basically it must be able to distinguish crowns from ground patches and shadows. Since the geometric features are partly determined by the selected segmentation algorithms (particularly in case of BRG and BROWNMO), radiometric features seem most suitable. The degree of artificiality (DOA – like NDVI with green and red channel) is a measure, to distinguish between ground patches and crowns whilst the brightness is considered, to eliminate shadow-regions. Consequently, [DOA > 0 && BRIGHT > mean(BRIGHT)] forms a straightforward model with *a priori* fixed decision boundaries and is implemented as 'DOA & Bright' in the module `bbHypoSelect`.

Secondly the accuracy is assessed by comparing the selected regions with the reference polygons. For matching, an overlap factor according to STRAUB & HEIPKE (2004) is defined:

$$OF_{AB} = \frac{A \cap B}{\min\left(|A|, |B|\right)} \qquad [9]$$

An $OF > 0.5$ is considered as a match and written to a `matchlist`. In the case of several overlaps fulfilling the criterion for a single region, the one with the biggest overlap is chosen.

3 RESULTS

Fig. 6 shows the result of BRG on LAMOTTE1. The seeds in the first figure were generated with a special technique giving an estimated size of the corresponding tree (for details see BECHTEL 2007, ERIKSON 2004b) which is used to initialise the variance in [3] if the parameter `AUTO_TREE_SIZE` is set. Consequently the segmented regions in the second figure are allowed to differ in size. Beside trees single ground patches are segmented, which are excluded by the hypothesis selection in the third figure. This is possible because the BRG segmentation doesn't presume high radiometric intensity but just relative homogeneity. The last figure shows the result of the matching with the reference data. The completeness is 86% and the correctness 98%, but all false positives are half trees at the image boundary which means it is even better. The area of the trees was underestimated systematically and is 74% on average. Since the darker areas at the crowns' boundaries are not part of the segmented regions in some cases, the ratio of the convex hull of the segmented regions to the reference was chosen as a second measure for accuracy and is 83% on average for this run. The mean positional error is 2.2 pixels corresponding to 32cm, which can be assessed as good.

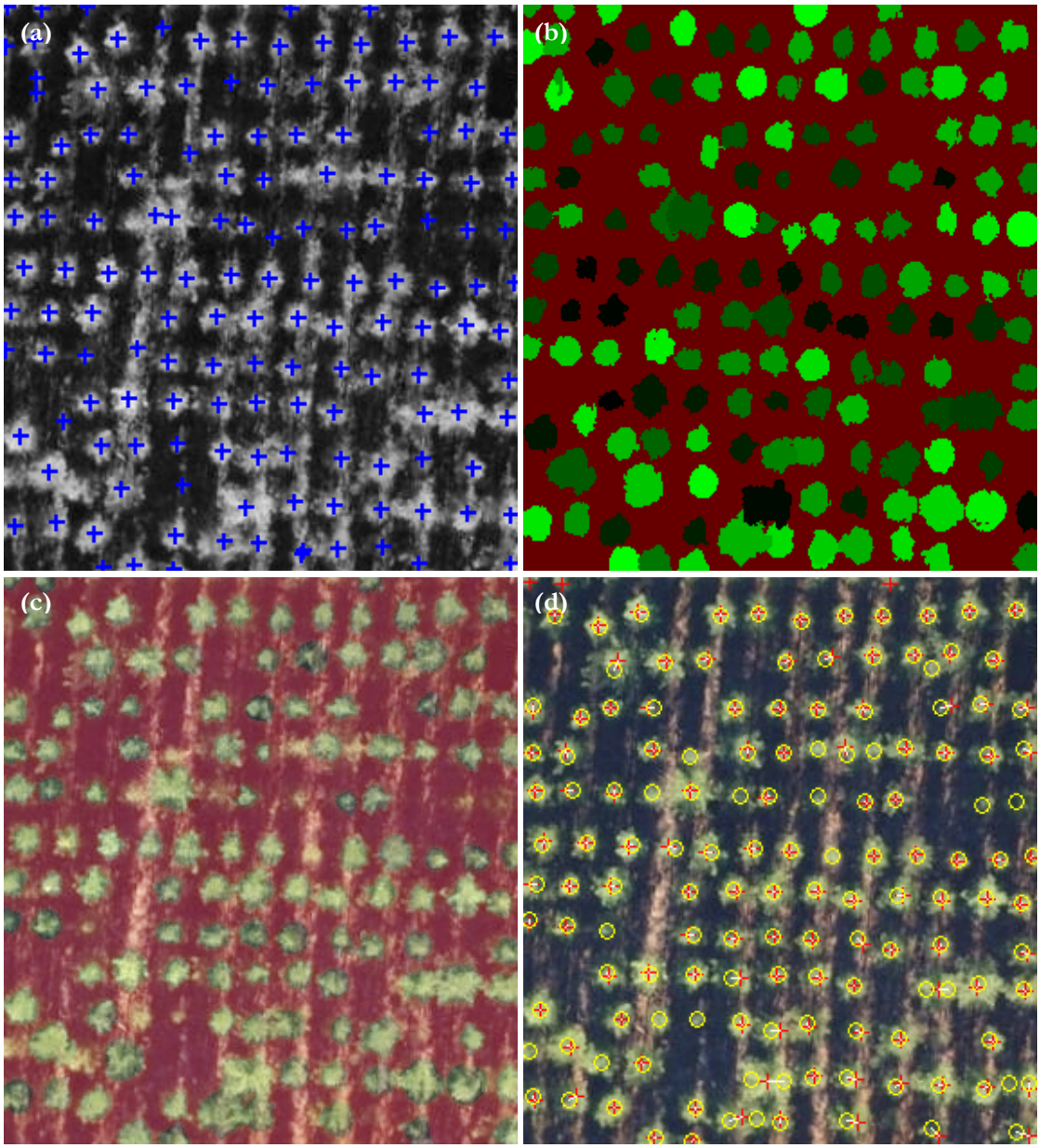


Fig. 6: Results of BRG with LAMOTTE1: seed-points (a), segmentation result (b), selected tree hypothesis (c) and matching with reference trees (d); red crosses indicate extracted trees, yellow circles indicate reference trees and the white lines indicate the matches.

Fig. 7 shows the result of BROWNMO with LAMOTTE2. As weight function the mean of the three colour channels with a weight power of 1.5 was chosen. The segmented regions in the first figure show a different characteristic since the virtual particle accidentally ran into adjacent ground patches. In the second figure it is recognisable that the model mistakenly discarded some actual trees due to their radiometric properties. Some of them might be infected by wood warps and thus have altered their colour but partly it might also be owed to the radiometric quality of the image, since all those regions have a degree of artificiality of less than zero. The third figure shows the matching results and the last plot the ratio between extracted and actual size (red crosses) and the same for the convex hull (green crosses) respectively. The completeness of the run is 84% and the correctness is 96%. The mean positional error is 38cm and the area is 80% of the reference on average.

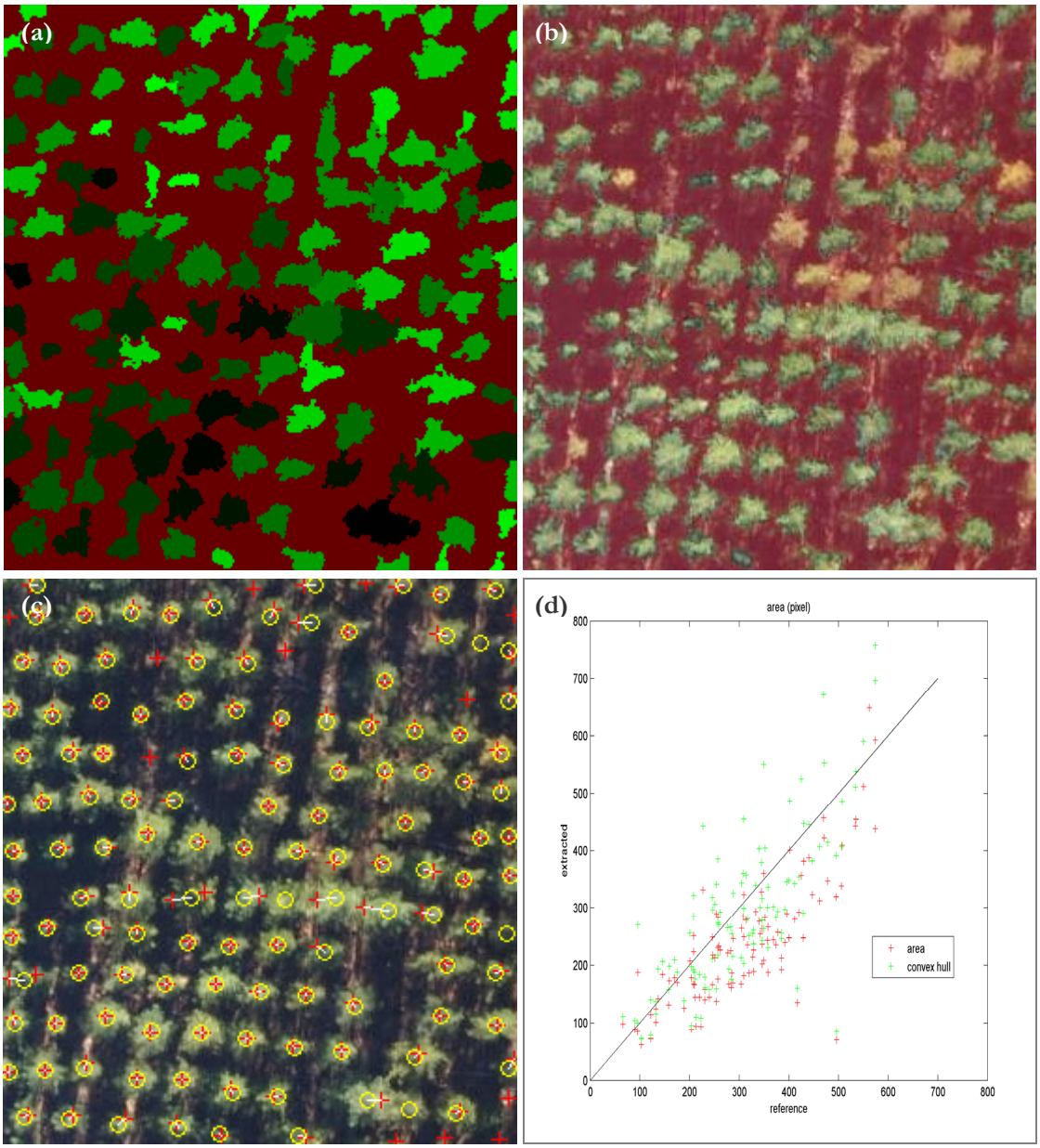


Fig. 7: Results of BROWNMO with LAMOTTE2: segmented regions (a), selected tree hypothesis (b), matching with reference trees (c), comparison of extracted and reference area of the matched regions (d).

The results of SRG with LAMOTTE2 in Fig. 8 illustrate, that the applied radiometric model is also very suitable for hypothesis selection with the complete segmentation of SRG, even though the extracted regions have rather specific radiometric than geometric conditions. The third figure reveals a certain over-segmentation of the image. Especially the crown borders tend to be segmented into "midtone" regions, which mostly comprise several crowns like in the highlighted example. This is a result from the high emphasis on radiometric homogeneity while the only geometric requirement is 4-neighbour connection and leads to outliners in the matching of the extracted regions.

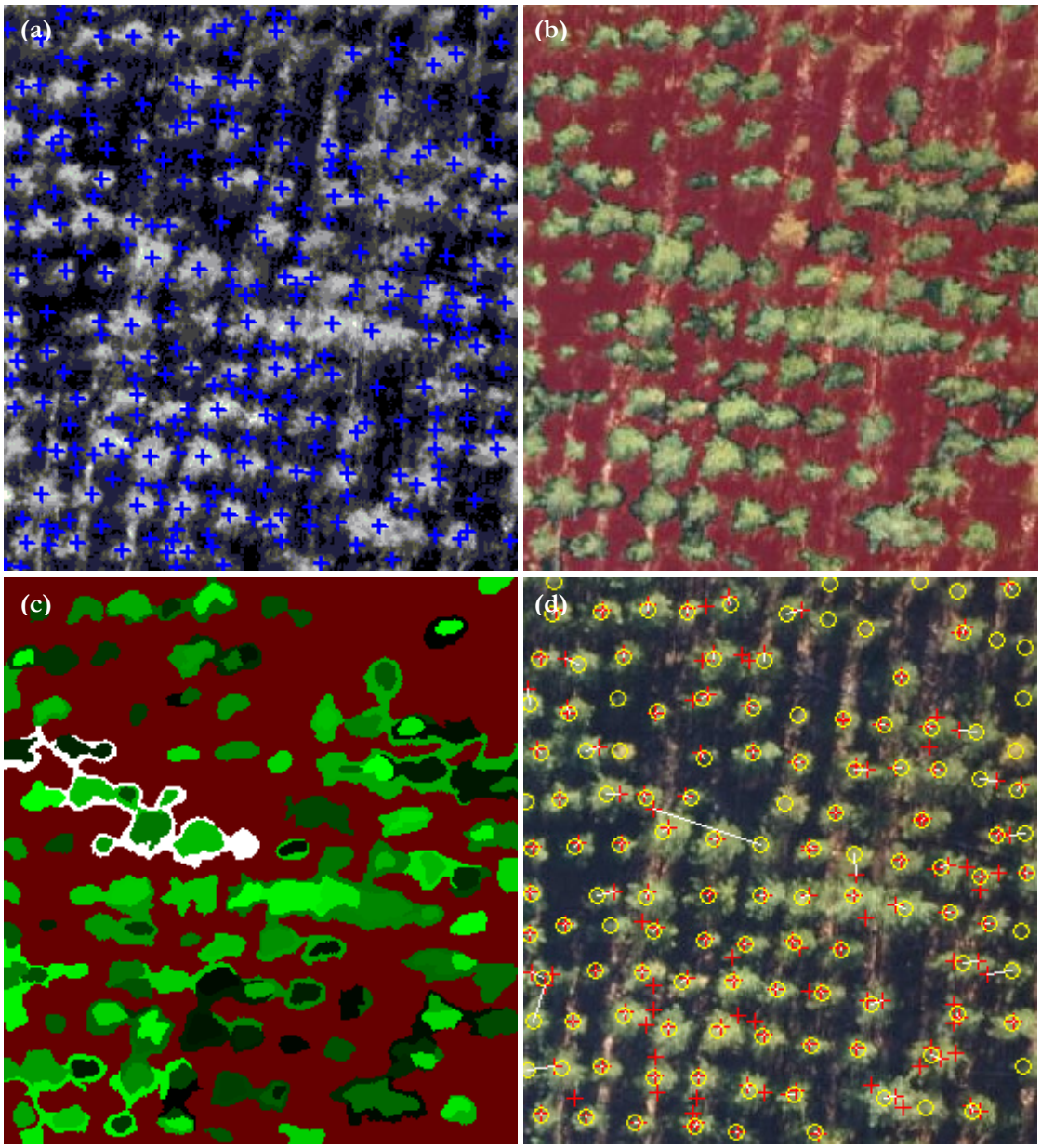


Fig. 8: Results of SRG with LAMOTTE2: seed points (a), selected tree hypothesis as masked image (b) and corresponding regions (c), matching (d).

4 Discussion and Conclusions

For BRG and BROWNMO the completeness is between 80 and 90% and the correctness is 95% or even more, comparable with better values achieved in other studies for similar forest stands (STRAUB 2003), which is a good result considering the simplicity of the utilised model. The accuracy of the extracted parameters (especially position and crown area), however, requires a more differentiated assessment.

With BRG the crown area was systematically underestimated, but according to ERIKSON "it is important that the segmentation result preserves the contour of the tree crown since it is used for classifying […]" (2004b: 471). Particularly shaded parts of the crown are not added to the segment but exactly this feature can be used to discriminate different species. Furthermore for most forestry applications the relevant size is the crown diameter instead of the crown area which should be less affected by the shadowing effect. Anyway it has to be stated that the reference data was manually extracted from the image and may therefore underestimate the crown size as well or even completely hide small shaded trees (see PERSSON *et al.* 2002).

Tab. 1: Accuracy of the extraction result for the datasets LAMOTTE1 (LM1) and LAMOTTE2 (LM2) with different segmentation algorithms and the simple model 'DOA & Bright'; wf ='G': only green channel, 'RGB_mean': all colour channels, wp: power of weight function.

	BRG		**BROWNMO**				**SRG**	
Dataset	LM1	LM2	LM1	LM1	LM2	LM2	LM1	LM2
Run identifier	BRG1	BRG2	BM1	BM2	BM3	BM4	SRG1	SRG2
Parameters	AUTO_ SIZE	AUTO_ SIZE	wf:'G'; wp: 1.5	wf:'RGB_mean'; wp: 1.5	wf:'RGB_mean'; wp: 1.5	wf:'RGB_mean'; wp: 2		
Model	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright
Completeness (%)	86	87	83	81	84	87	88	84
Correctness (%)	98	98	96	95	96	96	69	69
Mean positional error (m)	0.32	0.35	0.34	0.35	0.38	0.41	0.47	0.58
Ratio extracted area / reference region (%)	74	72	100	95	80	51	63	72
Ratio convex complement / Reference (%)	83	79	121	117	99	63	82	120

For BROWNMO the estimated area seems to be more accurate, but this is partly since the virtual particles mistakenly ran into ground patches. This assumption is stressed by the fact that the mean *OF* is lower for BROWNMO than for BRG (89% for BM1 compared to 93% for BRG1). BM4 was included to show that the completeness and correctness, which are rather dependent from the seed points, are quite robust against the chosen power of weight function parameter, whilst the size of the extracted crowns, instead, is very sensitive to the weight function parameter. With additional runs it was examined whether different weight functions are a suitable parameter to control segmentation output. It could be shown, that higher powers at the weight function produce brighter and more homogenous regions. Since higher powers also lead to smaller regions, this is partly an effect of the texture-specific semi variance, but it still was found useful to adapt to the poor radiometric quality of the given image. In contrast the use of weight functions with different colour channels did not derive any additional benefit.

Though SRG performs well the result shows that for the complex task of object extraction more specialised algorithms which indirectly comprehend model-knowledge are preferable. In particular for the purpose of object extraction more emphasis should be aspired on the compactness and simple connectedness of the extracted regions since most objects tend to have those characteristics. As aforesaid, the crowns are partly decomposed into several regions which mostly show a circular characteristic. Accordingly, the radiometric homogeneity appears to be weighted heavier than the compactness of the region – another model feature, since it usually can be assumed, that trees are compact and simply connected. The over-segmentation results in a high completeness and a little correctness. For the given run (SRG2) 143 regions are selected as hypothesis but after removing double matches (especially two regions matching with one reference tree) only 105 are left. Accordingly the area ratio is rather low. In contrast fringe regions like the highlighted in Fig. 8 form outliners, resulting in a huge ratio of the convex hull which is obviously no suitable measure for non- simply connected regions.

The mean positional area is small (about 30-40cm for BRG and BROWNMO and a bit higher for SRG due to outliners) which is partly an effect of the rather strict matching rules. If the underlying concept of matching identity was loosened, the positional error would increase but the error of omission would decrease. Against this background and in order to make SAGA a more valuable tool for object extraction, we propose to introduce more sophisticated segmentation algorithms including parameters for the compactness in position space.

REFERENCES

BECHTEL, B. (2007): Objektextraktion von Bäumen aus Luftbildern. Vergleich und Steuerung von Segmentierungsverfahren zur Vorbereitung eines Expertensystems. – Diploma thesis at the University of Hamburg.

BÖHNER, J., SELIGE, T. & A. RINGELER (2006): Image segmentation using representativenesss analysis and region growing. – In: BÖHNER, J. MCCLOY, K.R. & J. STROBL (Eds.): Göttinger Geographische Abhandlungen 115: 29-38.

BRANDBERG, T. & F. WALTER (1998): Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. – Machine Vision and Applications 11: 64-73.

DIEDERSHAGEN, O., KOCH, B. & H. WEINACKER (2003): Automatic Estimation Of Forest Inventory Parameters Based On Lidar, Multi-Spectral And FOGIS Data. – Proceedings Optical 3-D Measurement Techniques VI II: 259-269.

ERIKSON, M. & K. OLOFSSON (2005): Comparison of three individual tree crown detection methods. – In: Machine Vision and Applications 16(4): 258-265.

ERIKSON, M. (2003): Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules. – Canadian Journal of Forest Research 33: 1557-1563.

ERIKSON, M. (2004a): Segmentation and Classification of Individual Tree Crowns in High Spatial Resolution Aerial Images. – Uppsala.

ERIKSON, M. (2004b): Species classification of individually segmented tree crowns in high-resolution aerial images using radiometric and morphologic image measures. – Remote Sensing of Env. 91: 469-477.

ERIKSON, M., PERRIN, G., DESCOMBES, X. & J. ZERUBIA (2006): A comparative study of three methods for identifying individual tree crowns in aerial images covering different types of forests. – In: Proc. Int. Society for Photogrammetry and Remote Sensing (ISPRS), Marne La Vallee, France, July 2006.

GOUGEON, F.A. (1995): A crown-following approach to the automatic delineation of individual tree crowns in high resolution aerial images. – Canadian Journal of Remote Sensing 21(3): 274-284.

HEURICH, M., PERSSON, Å., HOLMGREN, J. & E. KENNEL (2004): Detetcting and measuring individual trees with laser scanning in mixed mountain forest of central Europe using and algorithm developed for Swedish boreal forest conditions. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 - 8/W2.

HINZ, S. (2003): Automatische Extraktion urbaner Straßennetze aus Luftbildern. – München.

HYYPPÄ, J., SCHARDT, M., HAGGRÉN, H., KOCH, B., LOHR, U., PAANANEN, R., SCHERRER, H.U., LUUKKONEN, H., ZIEGLER, M., HYYPPÄ, H., PYYSALO, U., FRIEDLÄNDER, H., UUTTERA, J., WAGNER, S., INKINEN, M., WIMMER, A., KUKKO, A., AHOKAS, E. & M. KARJALAINEN (2001): HIGH SCAN: The first European-wide Attempt to Derive Single-Tree information from Laserscanner Data. – The Photogrammetric Journal of Finland 2(17): 58-68.

LARSEN, M. & M. RUDEMO (1997): Estimation of Tree Positions from Aerial Photos. – In: LINDEBERG, T. (Ed.): Proceedings of the 1997 Swedish Symposium on Image Analysis, pp. 130-134.

MORSDORF, F., MEIER, E., ALLGÖWER, B. & D. NÜESCH (2004): Cluster-Analyse von LIDAR-Vektordaten zur Erfassung geometrischer Eigenschaften von einzelnen Bäumen. – Photogrammetrie, Fernerkundung, Geoinformation 4/2004: 259-268.

PERSSON, A. (2001): Extraction of Individual Trees using Laser Radar Data. – Linköping.

PERSSON, A., HOLMGREN J. & U. SÖDERMAN (2002): Detecting and Measuring Individual Trees Using an Airborne Laser Scanner. – Photogrammetric Engineering & Remote Sensing 9 (68): 925-932.

PINZ, A. (1989): Final results of the Vision Expert System VES: finding trees in aerail Photographs. – In: PINZ, A. (Ed.): Wissensbasierte Mustererkennung. Wien.

POLLOCK, R.J. (1996): The Automatic recognition of Individual trees in Aerial Images of Forests based on a Synthetic Tree Crown Image Model. – Vancouver.

STRAUB, B.M. & C. HEIPKE (2004): Concepts for Internal and External Evaluation of Automatically Delineated Tree Tops. – In: IntArchPhRS XXXVi 8/W2, pp. 62-65. Freiburg.

STRAUB, B.M. (2003): Automatische Extraktion von Bäumen aus Fernerkundungsdaten. – In: Wiss. Arbeiten, Fachrichtung Vermessungswesen 249. Hannover.

PREDICTING THE DEPTH OF HYDROMORPHIC SOIL CHARACTERISTICS INFLUENCED BY GROUND WATER

Michael Bock [1,2] *& Rüdiger Köthe* [2]

[1]Institute of Geography – Section Physical Geography – University of Hamburg, Bundesstr. 55, D-20146 Hamburg (Germany)

[2]scilands GmbH – scientific landscapes, Goetheallee 11, D-37073 Göttingen (Germany)

Abstract: In this paper a method is provided to derive plausible data of depth of soil characteristics influenced by groundwater. This model aims for usage of the data in the Soil Information System of the Federal Institute for Geosciences and Natural Resources (BGR). It contributes as one out of several data layers to support the Digital Soil Mapping process at scales 1:50,000 to 1:200,000. As input data a Digital Terrain Model, digital stream network and point data from groundwater measurements are used. The main focus lies on how the input data has to be prepared so that a Regression model that produces good results can be applied. All the data handling is done with SAGA GIS.

1 INTRODUCTION

The method for predicting the depth of hydromorphic soil characteristics determined by groundwater, which will be introduced here, is a model aiming for the initialization of digital data for the spatial data base of the Federal Institute for Geosciences and Natural Resources (BGR) Soil Information System. The BGR Soil Information System itself consists of a relational data base administering the attributes of digital features, most of all the Soil Map of Germany at scale 1:200,000 (KRUG *et al* 2003). The BGR Soil Information System impersonates a very special kind of digital data model the reduction attribute (comp. STACHOWIAK 1973) of which has to be considered, both concerning its spatial validity and its profoundness in content. The goal behind the presented model has to be stressed as there have again and again been requirements which cannot be claimed of such a model.

These requirements refer to modelling of form and location of a certain groundwater surface at a point in time. This, however, is subject to hydrologic models, which take into consideration – among other things – the hydraulic characteristics of the substratum. These have not been included into the model of regionalization of soil properties influenced by groundwater.

In the model presented here it is postulated that the vertical distance of a position to the drainage channel has a connection with the characteristics of hydromorphic soil properties. It uses a digital stream network, a digital terrain model (DTM) and groundwater stations as input data, which contain the long-term average of groundwater height close to the surface as data. The model can be described as follows according to the main features of a model (STACHOWIAK 1973).

Feature of representation: the model is a digital representation of reality, i.e. it displays, in areas with certain features, in what depth hydromorphic soil characteristics will be likely to turn up. This representation is given in form of a digital grid data set.

Feature of reduction: the model is reduced (i.e. simplified) concerning its size (compulsory in geography) and its complexity compared to reality; i.e. it idealizes reality, as various factors, which in nature influence the specificity of hydromorphic soil characteristics caused by groundwater do not enter the model. This goes for e.g. hydraulic characteristics of the soil or the climate.

Feature of pragmatism: the model describes its being goal orientated; i.e. it has been created for a certain field of appliance. This orientation has to be stressed especially here: the presented model is to serve the soil mapping process for scales 1:50,000 to 1:200,000 and the results will become part of the BGR Soil Information System. The cartographies in those scales actually represent models, too or make use of the

quickly grown number of models (because of the rapid development in GIS within the last two decades, (BÖHNER 2006)) concerning the distribution of different soil characteristics. Within the process of generating soil maps models for the distribution of soil types are created with the help of various approaches from different existing levels of data (measuring and model data) and with the help of methods of gathering new data (BEHRENS & SCHOLTEN 2006). Hydromorphology, i.e. the features of a soil influenced by groundwater or backwater, is one of many data levels defining the characteristics of a certain type of soil.

The feature of pragmatism of the model presented here must be stressed, as there are similar models, but with an orientation to different applications in order to avoid mistakes. These others are the various groundwater models, mainly of hydrologic branches of study. Our approach of regionalization of groundwater characteristics of soils, however, does neither comprise hydraulic features of the soil nor does the specific depth of the surface of a groundwater body, which could e.g. be modelled out of reasons of protection of the water, designate the features of the model.

2 TESTING SITES AND MATERIAL

Testing sites: Several areas in the lowlands of the federal state of Lower Saxony have been chosen as testing sites. This is due to two characteristic: On the one hand the model was to reach plausible results in spite of its reduction. Thus solid rock areas, which influence the attitude of groundwater close to the surface significantly by changing geological layers, were excluded right away. Moreover, the required data on their spread and mechanisms of control have not been available for operation. The assumption that the differentiation of quaternary unsolidated sediments of the lowlands of Northern Germany could be ignored when dealing with the distribution of groundwater features is a simplification and can – in particular cases – lead to misleading results. However, the amount of control by varying unsolidated sediments is commonly regarded as less significant than the changes in solid rocks. The question as to what amount solid rock areas with varying petrology influence the depth and form of groundwater features of soils could be the issue of another expansion of the model.

A second argument for choosing areas of the lowlands of Northern Germany as testing areas is the homogeneity of data records of the groundwater stations. A homogenization of groundwater stations of different origins was carried through in this state in a particular project (HYDOR 2003), so that there were no problems at all concerning the data preparation. The question whether this is true for other German states or European countries cannot be answered at the moment.

Data sources and preparation: The selection of data input highly influences the quality of the result of the model. In this study importance was attached to using digital geographical data which are available ubiquitary in several federal German states nowadays. It is a fact generally known that this is not the case in particular cases and that the data are presented in different levels of quality. However, up to now, the quality of data concerning geographical data has improved in Germany so far that – theoretically, at least – there should be data of a quality comparable to that presented in this paper in each federal state. Nevertheless, the spatial resolution of a DTM (in meters) or the density of a stream network (in m/km^2) or a network of groundwater stations (in n/km^2) is not the only criterion relevant for the quality of a data base and so also for a result of regionalization which should be applied. Therefore the following paragraphs will concentrate on the question which qualifying criteria are relevant for particular data layers.

Digital Terrain Model: As DTM the countrywide available "DGM5" of Lower Saxony was chosen. This data set was constructed within ten years from various data sources (digitized contour lines, photogrammetry, laser scanning) and therefore is rather heterogeneous. Tests have proved that for the task presented here the comparatively high resolution (12.5m grid cell size) of "DGM5" is by no means necessary. Using a grid cell size of 25m can also render results of the same quality. So a spatial resolution is demanded which is available nationwide in all federal German states nowadays.

Certain characteristics of the "DGM5" of Lower Saxony apart from spatial resolution, however, pose a problem for the chosen method. For modelling groundwater levels it is absolutely necessary to use a DTM without anthropogenic artefacts those being surfaces within the DTM which are not in accordance with the natural earth's surface. This is the case with e.g. artificial embankments or cuts caused by traffic routes (motorways or railway tracks), acclivities, excavations a.s.o. The problem of anthropogenic artefacts in DTM tends to turn up again and again as DTM from laser recordings becomes more common. The term „artefact" as a product or phenomenon created by human or technical influence opposed to the not influenced or natural phenomenon is valid in this context, even if it is used more often as a synonym for signal interferences. These, however, are not of interest here the DTM being interpolated from completely correct measurements thus embodying a model which normally corresponds to the real earth's surface. For modelling natural soil phenomena this anthropogenically shaped surface is not relevant. Soil characteristics have been developing in their natural terrain within long periods of time. A model building up coherency of these features with the form of the earth's surface must therefore use DTM containing the natural earth's surface. Using certain algorithms it is possible to detect most anthropogenic forms in DTM and to reconstruct something like a natural model of the earth's surface.

Actual mistakes, which more than never occur in nationwide DTM of the federal states, should be eliminated. These are above all tile structures and tile ridges, regular ridges, noise and triangle facettes. Using algorithms to eliminate such mistakes does not mean that genuinely missing altitude values can be reconstructed. However, it is possible to interpolate a plausible relief in such cases (KÖTHE & BOCK 2006). If, on the other hand, such structures remain in the DTM, they can trace as far as the modelling results.

Stream network: Data concerning the digital stream network bear special relevance in modelling the medial depth of groundwater close to the surface. Within the given model it is assumed that in natural systems the altitude of an open body of water coincides with that of the highest groundwater level. It is known that this is not always the case[1], but it is nevertheless adopted as simplifying assumption within the model. According to that a consistent digital stream network forms a vital source of data for the interpolation of a model-like surface corresponding to the depth of soil characteristics influenced by groundwater. Another possible data base are e.g. the locations of other inland waters, e.g. lakes. However, due to a multitude of artificial, gravel pits a.s.o. the altitude level of the waterbody can more often than not not be related to the altitude level of a groundwater body close to the surface. So this data of inland waters were neglected. Fens, however, are a useful supplementation, as groundwater always turns up naturally at the surface here. Nevertheless, the core modelling refers to running waters and criteria have to be considered according to which the ingoing data is evaluated and maybe adapted: Artificial channels of an altitude level above the natural earth's surface have to be eliminated from the data set. For such waterways the assumption mentioned above is not applied according to which the surface of a stream coincides with the surface of the ground water level close to the surface. For the areal sections shown here the elimination of said data objects could be undergone by hand. Once larger areas are the case this is no longer reasonable or the choice of sections cannot definitely be ascertained by the progression of the object.

One possibility of automated detection is the selection of data with the help of information from the attribute table. However, this works only under the condition that the information "artificial waterway" or "natural stream" were placed in the data set before. Since this is not always the case, a possibility is to be worked out to both extent and correct the stream-dataset with objective criteria. For the project presented here the stream network from Lower Saxony Water Management, Coastal Defence and Nature Conservation Agency (NLWKN) derived from Digital Base Landscape Model (DLM) of Lower Saxony could be used. This basic DLM is part of the ATKIS and should be built up in Germany by several

[1] Colmation (clogging by sedimentation of suspended matter) leads to an interruption of the otherwise corresponding systems flowing waters - groundwater mostly in water bodies heavily strained with suspended matters thus causing the highest groundwater level of allochton streams (i.e. rivers receiving water from remote areas of drainage) to lie considerably beneath the altitude level of the flowing water.

agencies for geo data of the federal states according to uniform criteria. On average the data set consists of vector data for 758 meters of stream per km². Seeing it from the view of modelling, however, the data set cannot be considered as complete as branches of stream are missing, which exist on the topographic map and aerial photographs. Such limitations must also be expected in other federal states. Another limitation is due to the fact that the information whether a stream holds water perennially, or just periodically, or even intermittently cannot be read out of the data set. This, however, would be vital for the regionalization of long-term lows or highs of the groundwater contact of the soils. Using the same stream network for the model for the medial, highest and lowest groundwater level can e.g. in places where streams run dry during the summer lead to systematical errors which have to be taken into consideration.

Measurement data of the groundwater level by HYDOR (2003): A data collective could be used as measurement data which in 2003 had been researched for in a separate project by Hydor Consult GmbH, Berlin, deriving from different data sources for the area of Lower Saxony. Apart from groundwater stations (data source: NLWKN) the data collective includes data from public utilities for drinking water, stream gauges (data source: NLWKN) and auxiliary points from different maps and other sources. A data network of a density of 0.9 points/km² has come to existence by this project. The data of mainly two categories were not suitable for the purpose of regionalization of the groundwater level presented here. For the testing sites were eliminated:

- Stream gauges form a dense network along streams and rivers. In the introduced method a certain terrain parameter is derived from a digital stream network so the altitude of the stream locations is taken into consideration at a different point anyway. The usage of this data category in a regression model would have meant an overweighting of points from locations were the dependent variable is known. In contrast the number of measurement points at positions were the groundwater level does not reach the surface would have been very low so the stream gauges have been resigned.
- Auxiliary points from intersection with DTM: This category of „groundwater data“ was introduced by HYDOR (2003) to achieve an aggregation of the point data to perform an interpolation. To use this category would have meant to anticipate parts of the introduced method already for some of the input data. Additionally the DTM that was used by HYDOR (2003) was of less quality compared to the “DGM5”. So these points were removed out of the data collective.

These categories were removed for the testing sites. All the remaining points represented measured point data. If more categories should be removed when other areas of Lower Saxony are processed must be kept open. Figure 1 shows the locations of the testing sites and the spatial distribution of the input data.

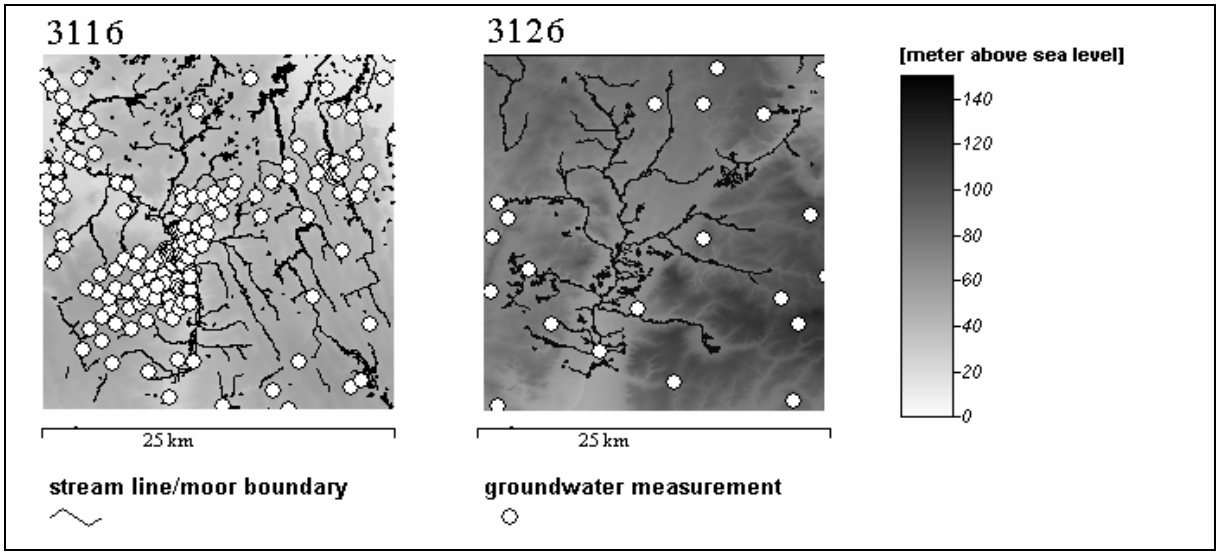


Fig. 1: The two testing sites in the lowlands of Lower Saxony and the equipment with input data.

3 METHODS

Terrain analysis: The open source GIS Software SAGA provides numerous free modules to work on DTM. The Terrain Parameters that can be produced (and even further more that are not available in the free modules) were tested if they could fit to the worked out model. The parameter „*vertical distance to channel network*“ soon was extracted to obtain the highest correlation coefficients in the regression model with the groundwater stations. This parameter proved to be so dominant, so it was decided to give a description of its functionality here.

The module „*vertical distance to channel network*“ is part of the module library *Terrain Analysis - Channels* in SAGA. It calculates a grid data set, which contains a value for the vertical distance above the level of the channel lines[2] for each grid cell. Frequently the parameter is used in Geomorphographic Maps to support the soil mapping process. Also as auxiliary variable in Geostatistics the parameter „ *vertical distance to channel network* “ is suitable for predicting e.g. soil texture. The SAGA module that does the calculation needs a DTM and a grid data set with locations of channel lines as input. Instead of a theoretical channel network in this study more generally locations of groundwater discharge are transferred to the module. These are the processed stream network and boundary lines of fen areas additionally. It does not derogate the calculation at all that this is not a line network at all.

In general it could be said that the implemented algorithm „*vertical distance to channel network*“ is a interpolation method[3], that does an iterative aggregation of values – the number of steps depends on the size of the grid data set. It interpolates between the positions of grid cells with channel lines, until all grid cells between these channel lines have got a z-value. The grid cells with channel lines were assigned with z-values of the DTM at these locations. The result is a grid data set that constitutes a theoretical surface between the channel lines and is called „base level“. The functionality surpassing a pure interpolation routine is that the values of the calculated grid data set are prevented to get higher than the according values in the DTM. This function can optionally turn off by the user.

In most cases the user does not need the base level and the module does not even produce it as an output data because this is optional, too. The final result of the algorithm, the grid data set ”*vertical distance to channel network*“, is the result of a simple subtraction. The DTM minus base level leaves ”*vertical distance to channel network*“ and this is returned to the user. Here the base level is essential for the prediction of the depth of soils influenced by groundwater. In a regression model this parameter explains mostly more than 90% of the distribution of the measured values.

The functionality of the SAGA module ”*Vertical distance to channel network*“ is in C++ code. Like all other modules under GNU public licence it is free to use and copy and the source code is open (CONRAD 2006). The algorithm was written by CONRAD (2002) and reflects developments of BÖHNER & Köthe (2003). Concededly the terrain parameter ”*Vertical distance to channel network*“ was introduced as one out of several, so a detailed description of its functionality could not be given there.

Beside numerous routines or functions of the SAGA Programming Interface (SAGA API), that are used in the module and are not explained here, the module uses own routines to control the sequences. The routine `On_Execute`, is obligatory for performing of all SAGA modules because the SAGA environment automatically calls this routine during execution. Further functions are defined outside `On_Execute` and are called on demand by the main function of the module.

---

[2] Channel lines are theoretical pathways of the artificial of a DTM. These channel lines must exist before the execution of the module in form of a DTM according grid data set, in matters of grid cell size and grid dimension. The generation of plausible channel networks is not an issue of this study.

[3] A similar interpolation method, going back to LEE (1997), is „*Multilevel B-Splines Interpolation*“. It was implemented in SAGA by CONRAD (2006). The regression model with a surface interpolated by this module achieved similar statistical results. Though from the geo scientific point of view the results are less plausible because in valley bottoms often values are interpolated that exceed the values of the DTM. The derived groundwater distance then gets strongly negative. Here the ‚brake' in „*vertical distance to channel network*“ that prevents values higher than the surface of the DTM stands the test.

At the beginning of the execution the number of loops the calculation has to go through is determined depending on the size of the DTM. It is the maximum exponentiation of 2 that is still smaller than the longer side (X or Y direction) of the used grid data set. For example for a grid data set of 1000 to 1000 grid cells the number of loops thus is 9 ($2^9 = 512$, $2^{10} = 1024$ is to large).

Each loop starts with call of the function `Initialize_Surface`. It assigns a temporary grid data set (empty in loop 1) with z-values of the DTM at the locations of the channel lines. After that the grid cells of the temporary grid data are processed at the beginning in large distances, in later loops in decreased distances until in the last loop every grid cell (distance 1) is treated. The surrounding area of the actual grid cell is searched for grid cells that have already values. Does the function find grid cells it calculates the arithmetic mean and assigns it to the actual grid cell. If no values are found the function assigns the value of the DTM of that position to the temporary grid data set.

The function `Set_Surface` effects a further aggregation. It uses another function called `Get_Changed` that considers the distance of the centre of the grid cell as a weighting factor. Existing values of neighbouring grid cells again in decreasing distances according to the actual loop number are weighted with factor 1 in cardinal direction and with factor square root of 2 in diagonal direction. The weighted arithmetic mean of the values of that found out grid cells is compared with the value of the DTM at this position and `Get_Changed` returns the smaller value to the calling function `Set_Surface`.

This procedure of value aggregation continues until the last loop. Hereafter every grid cell between the locations of channel lines is assigned with a z-value. An interpolated surface comes into existence, the so called Base level. The Terrain parameter "*vertical distance to channel network*" results in a subtraction of DTM and Base level afterwards. Figure 2 shows the procedure of value aggregation exemplarily for situations of the temporary grid data set after four different loops.

Linear regression model: For the regression model to predict the depth of soil parameters influenced by groundwater the free SAGA module "*Regression Analysis (Grids/Points)*" from the module library *Geostatistics - Grids* was used. The source code was written by CONRAD (2004).

Before a stepwise multiple Regression analysis with numerous terrain parameters was performed. The results for the terrain parameter *Base level* of the module "*vertical distance to channel network*" were so convincing, that the decision to use the parameter with the help of a Simple Linear Regression for predicting the depth of soils influenced by groundwater was done. In Tab. 1 the statistical parameters for the analyzed data sets are listed. The following analysis of the residuals was not able to find any spatial autocorrelation so the geostatistic interpolation and the adjustment of the model results with interpolated residual data could be skipped.

The number of groundwater stations is very low in 3126 after removing the not suitable ones (cf chapter 2). Regardless the model resulted in plausible values. The site 3116 with at least 141 measuring points is equipped very well. Though in this site the situation regarding the parent material of soil is very heterogenic so it is likely that only one regression model maybe is not enough here.

The testing sites are characterized by very high values of R^2. These high values are quite unusual in regression analysis of soil properties. The relatively high standard error relativizes the high coefficients of determination. The consequences of that standard error of the regression model for predicting values for concrete X shows Figure 3.

The span of the standard error is visualized for a confidence interval of 90% exemplarily for the medial groundwater depth in 3116. Naturally the span is low near the arithmetic mean of the *Base Level* but is increasing towards the borders of the distribution. But even more than 2 m span standard error in the area of the arithmetic mean seems to be a high value. Nevertheless the model is able to predict plausible values (see chapter 4).

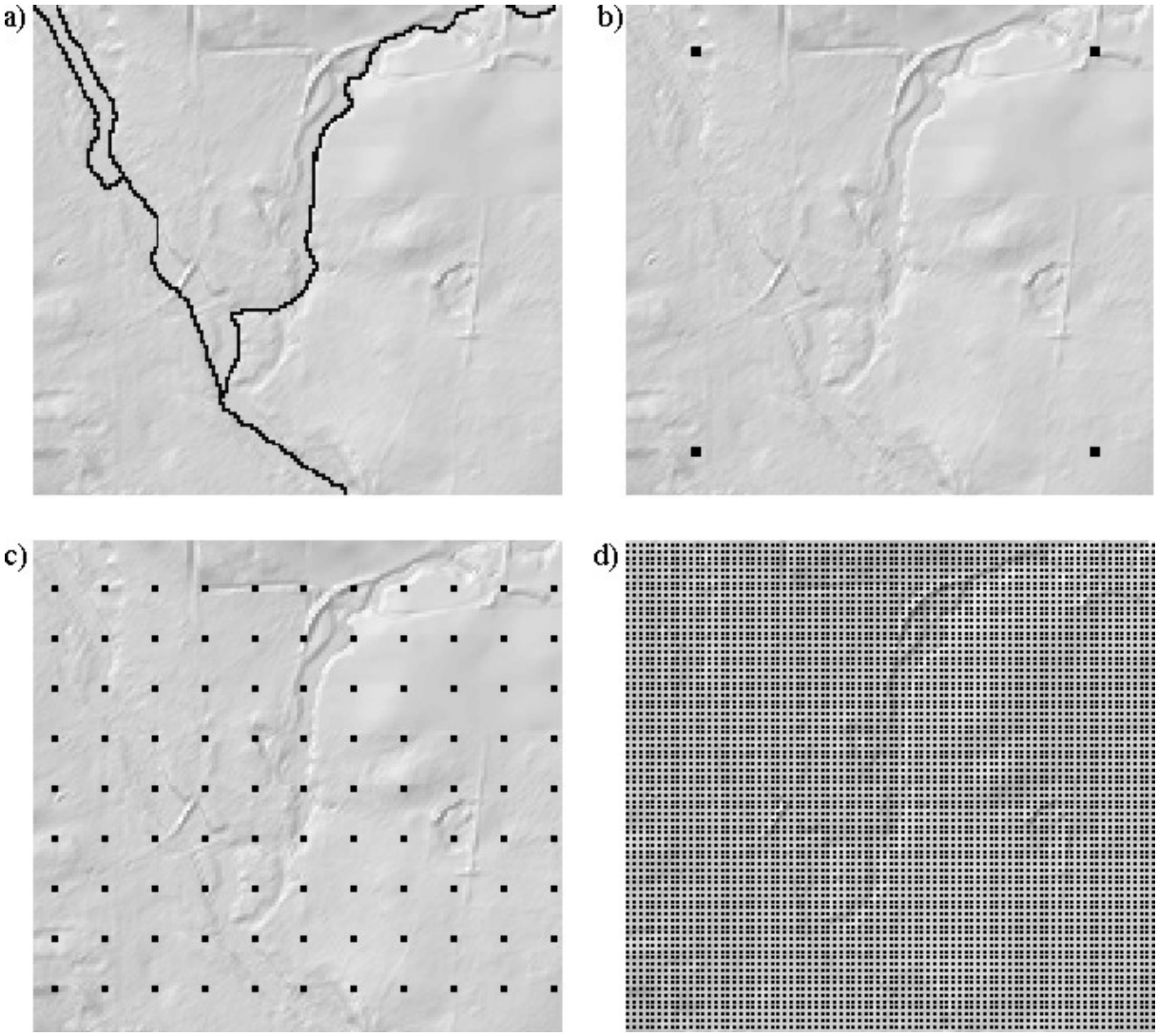


Fig. 2: The procedure of aggregation of values for testing site 3126. a) locations of grid cells with stream lines, b) the temporary grid data set after loop 5, c) after loop 8, d) after loop 11. In the next loop every grid cell will be assigned with a new value.

Tab 1: Statistical parameters in the testing sites.

		Number of values	Coefficient of det. (R^2)	Regression constant	Regression coefficient
3126	**Long term high value of groundwater**	14	0.91	-2.48	1.03
	Long term medial groundwater value	40	0.94	1.95	1.18
	Long term low value of groundwater	36	0.93	0.92	0,98
3116	**Long term high value of groundwater**	141	0.93	-1.53	1.03
	Long term medial groundwater value	167	0.93	-0.43	0.96
	Long term low value of groundwater	142	0.93	-1.13	0.98

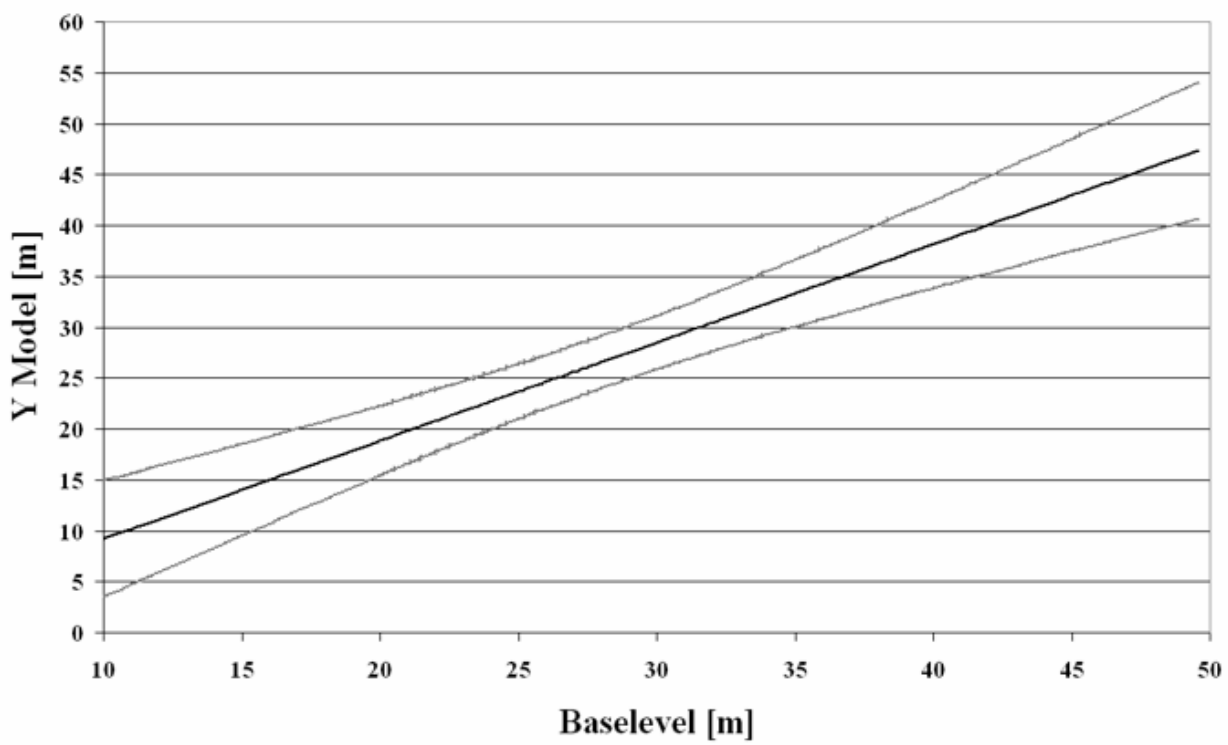


Fig. 3: Testing site 3116, medial groundwater level. The standard error for every predicted Y for a confidence interval of 90%.

In general, regression analysis in testing sites that lead to modelling functions producing plausible results cannot be transferred uncritically to other areas. A single regression function that represents differently equipped landscapes cannot be imagined. As soon as the influences from conditions left out of consideration change fundamentally, e.g. the climate or the parent material, as is the case in the two testing sites (sand content), the modelling function has to be adapted to the changing conditions. Therefore it is important to use such a model in landscapes with more or less homogenous conditions and completely independent from arbitrary map extends. Alternatively the model is provided with operationalized digital data of the changing conditions. This should be the field of further development.

4 RESULTS

In the following the results of applying the model in both testing sites are shown. Additionally a further step of processing helps this visualisation of the results: subtraction of DTM and the model result. This calculation leads to a data set that can be interpreted as the model-like groundwater distance. This data set can be used in the soil mapping process directly. To present further results like the derived amplitude as well as the comparison of this model results with existing models to predict the medial depth of a groundwater table (contour map of the groundwater surface) would go beyond the scope of this article and will be presented elsewhere soon.

Both in 3116 and in 3126 the model predicts plausible medial depths of groundwater table resp. soils influenced by groundwater in general. This affects the floodplains of the streams but also areas outside the floodplains where the height of the earth's surface is accordingly low. Naturally in some cases the model predicts groundwater heights near the earth's surface that are wrong. Due to errors in the input data and to its feature of reduction this is inevitable.

Figure 4 shows both the predicted depth of the medial groundwater table as well as the derived groundwater distance. The first exemplary use of the model results in soil mapping showed, that it is a suitable contribution to homogenization and acceleration of the mapping process.

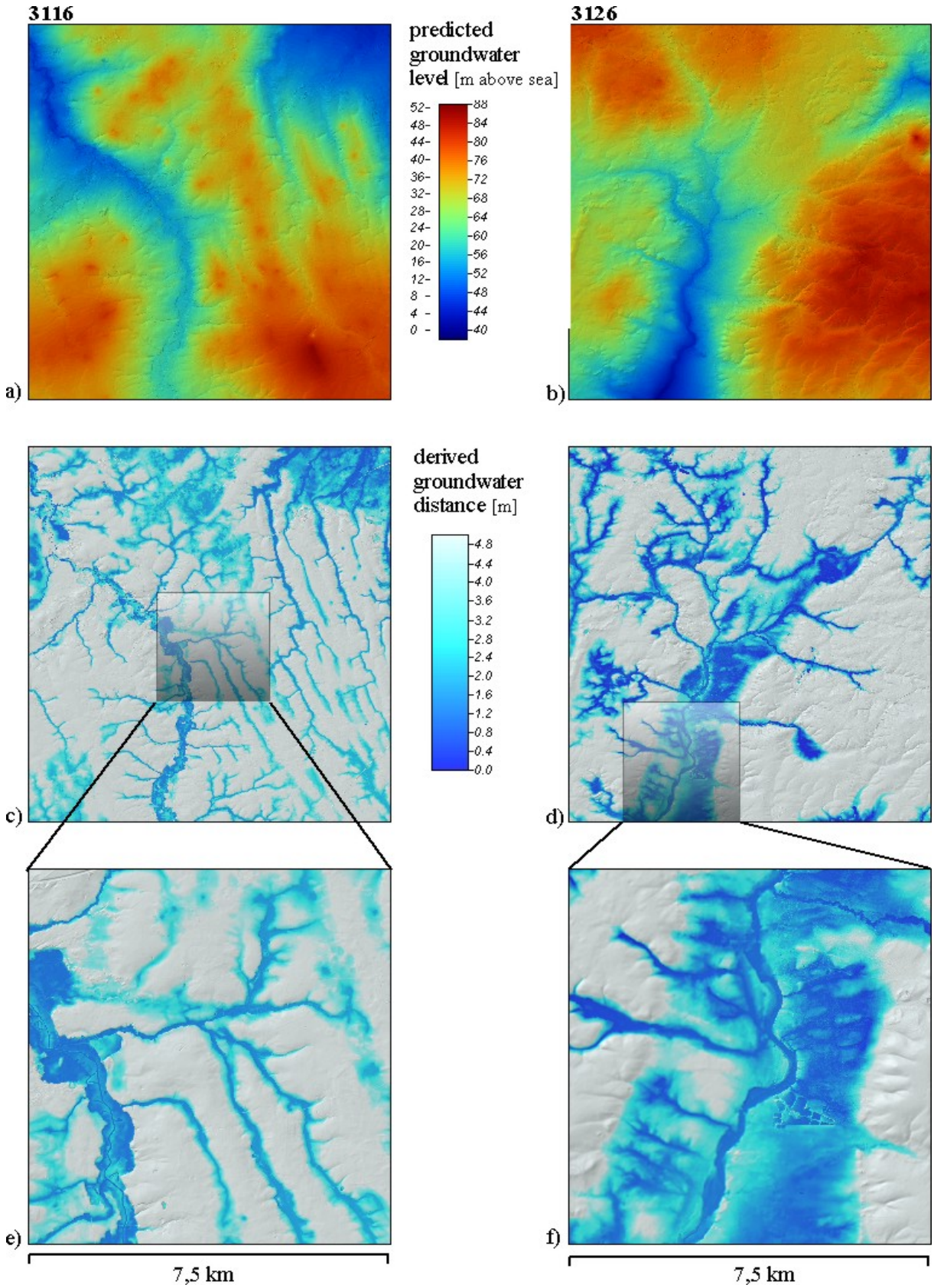


Fig. 4: a) shows the predicted medial groundwater level for site 3116, b) shows the predicted medial groundwater level for site 3126. As described DTM minus prediction leaves the medial groundwater distance from the surface (c and d). e) and f) are zoomed into the map. f) shows relatively high groundwater distances near the stream (indeed there are soils without groundwater influence) and lower values in a greater distance. This is quite plausible in this kind of landscape.

Acknowledgement: This study represents work of a research project that was funded by the Federal Institute for Geosciences and Natural Resources (BGR). This contribution is gratefully acknowledged.

REFERENCES

BEHRENS, T. & SCHOLTEN, T. (2006): Digital soil mapping in Germany - a review. – J. Plant Nutr. Soil Sci. 169: 434-443.

BÖHNER, J. (2006): Modelle und Modellierungen. – In: GEBHARD, H., GLASER, R., RADTKE, U. & REUBER, P. (Eds): Geographie: 533-538, Heidelberg.

BÖHNER, J. (2004): Climate Spatial Prediction and Environmental Modelling by Means of Terrain Analyses, Process Parameterisation and Remote Sensing. – Göttingen. Habilitation thesis at the Department of Geosciences and Geography, Georg August University's Göttingen.

BÖHNER, J. & KÖTHE, R. (2003): Bodenregionalsierung und Prozessmodellierung: Instrumente für den Bodenschutz. – Petermanns Geographische Mitteilungen 147(3): 72-82.

CONRAD, O. (2002): ChannelNetwork_Altitude.cpp, http://sourceforge.net/saga-gis: 24.05.2008.

CONRAD, O. (2004): GSGrid_Regression.cpp, http://sourceforge.net/saga-gis: 24.05.2008.

CONRAD, O. (2006): Gridding_Spline_MBA.cpp, http://sourceforge.net/saga-gis: 24.05.2008.

CONRAD, O. (2007): SAGA – Entwurf, Funktionsumfang und Anwendung eines Systems für Automatisierte Geowissenschaftliche Analysen. – Dissertation, Göttingen.

HYDOR (2003): Erarbeitung einer digitalen Übersicht zur Grundwasseroberfläche des Oberen Grundwasserleiters in Niedersachsen im Maßstab 1 : 200 000. – Research Report to State Authority for Mining, Energy and Geology (LBEG), unpublished, Hannover.

KÖTHE, R. & BOCK, M. (2006): Development and use in practice of SAGA modules for high quality analysis of geodata. – In: BÖHNER, J, MCCLOY, K.R. & STROBL, J. (Eds): SAGA-Analysis and Modelling Applications. – Göttinger Geographische Abhandlungen 115: 85-96, Göttingen.

KRUG, D., HARTWICH, R. & ECKELMANN, W. (2003): Harmonised Spatial Data for Soil Protection in Germany. – Proceedings 4th European Congress on Regional Geoscientific Cartography and Information Systems (17.-20.6.2003), 1: 140-141; Bologna.

MCBRATNEY, A.B., MENDOCA SANTOS, M.L. & MINASNY, B. (2003): On digital soil mapping. – Geoderma 117: 3-52.

LEE, S., WOLBERG, G. & SHIN, S.Y. (1997): Scattered Data Interpolation with Multilevel B-Splines'. – IEEE Transactions On Visualisation And Computer Graphics, 3(3): 228-244.

STACHOWIAK, H. (1973): Allgemeine Modelltheorie. – Wien.

STATISTICAL GEOCOMPUTING COMBINING R AND SAGA: THE EXAMPLE OF LANDSLIDE SUSCEPTIBILITY ANALYSIS WITH GENERALIZED ADDITIVE MODELS

Alexander Brenning [1]

[1]Department of Geography – University of Waterloo, Ontario (Canada)

Abstract: The integration of statistical software with geographical information systems is required to be able to efficiently combine the most powerful tools and techniques available in both environments. The RSAGA package, which provides access to SAGA GIS geoprocessing functions from within the R statistical data analysis environment, is a recent contribution to this endeavor. The present work gives an overview of the structure of the RSAGA package, and demonstrates its usefulness in the context of landslide susceptibility modeling with terrain attributes and generalized additive models (GAMs). The GAM is an extension of the generalized linear model (e.g. linear and logistic regression). It is able to model nonlinear relationships, but retains an interpretable additive structure. In the case study on landslide distribution in the Ecuadorian Andes, several local as well as catchment-related morphometric attributes are important, mostly nonlinear predictors of landslide occurrence. Other applications that can benefit from an integration of modern statistical computing techniques and GIS-based digital terrain analysis include pedometrics, Precision Agriculture, and species habitat studies.

1 INTRODUCTION

Spatial analysis using Geographical Information System (GIS) techniques on one side and the statistical analysis of environmental data on the other provide two different views of spatial data analysis problems that are too often separated by disciplinary and software-related barriers. However, both perspectives have much to contribute to spatial data analysis: GIS software provides a rich set of tools for spatial data manipulation, queries, and visualization, and statistical data analysis software can offer spatial and non-spatial techniques required for understanding data or applying predictive models. These models may range from traditional statistical ones such as linear regression to complex black-box models such as the support vector machine, a powerful, emerging machine-learning technique.

Given the fast progress in geographical information science and computational statistics, efficient ways of coupling GIS and data analysis software are required. This integration has been an important topic in recent years in the developer community of R (see e.g. BIVAND 2000, BRENNING & VAN DEN BOOGAART 2001, BIVAND *et al.* 2008), an open-source data analysis environment that is widely used in statistical sciences (IHAKA & GENTLEMAN 1996). A wide variety of free R extensions ("packages") is now available that support geodata formats (e.g. packages shapefiles, rgdal, maptools) or spatial statistical techniques (e.g. packages gstat, spdep, nlme).

The objective of the present work is to introduce the R package RSAGA (BRENNING 2007a), which integrates SAGA functionality into R, in a case study of landslide susceptibility modeling that also exemplifies the need for a seamless integration of terrain analysis and statistical models. A similar link is currently being established to access geoprocessing tools of ESRI's ArcGIS from within R using the RPyGeo package and a Python interface (BRENNING 2007b).

This work is divided into six sections: The first two sections introduce the landslide distribution data, and the terrain analysis and statistical techniques suitable for their analysis. The next two sections give an overview of the structure of the RSAGA package, and outline its application in constructing generalized additive models of landslide susceptibility. Finally, analytical and predictive model results are presented and discussed, and some general conclusions are drawn.

2 CASE STUDY: LANDSLIDE SUSCEPTIBILITY

Landslide susceptibility modeling on a regional scale can be considered a (soft) classification problem, where a measure of the likelihood of landslide occurrence has to be estimated. A variety of approaches has been applied for this purpose, mainly empirical ones (see the review by BRENNING 2005), but also physically based models (e.g. MONTGOMERY & DIETRICH 1994), and cellular-automata models (GUTHRIE *et al.* 2008). Empirical models require a training data set – a landslide inventory – to be provided for model fitting, and especially for the assessment of flexible machine learning techniques it is also required to use spatially or temporally independent test data sets or a spatial cross-validation (BRENNING 2005). Among the predictors that are most widely used in landslide susceptibility modeling are terrain attributes (especially slope angle and contributing area), land use, and soil or lithological properties.

The present study uses a multitemporal landslide inventory from the Andes of Southern Ecuador that was created by STOYAN (2000). Landslides and their importance in controlling vegetation succession in the tropical mountain rainforests of the area were studied by STOYAN (2000), WILCKE *et al.* (2003), BUSSMANN (2004) and LOZANO *et al.* (2005). BRENNING (2005) compared different statistical and machine-learning techniques for landslide prediction using a different subset of STOYAN's inventory, however with an extended set of predictor variables.

The study area comprises 11.2 km², 0.94 km² of which were classified as mass movements in the transition between shallow translational landslides and debris flows (Fig. 1). Landslide inventories were compiled for 1962, 1969, 1976, 1989, and 1998. Deforestation and subsequent grazing take place along the road from Loja to Zamora; this as well as the direct impact of the road on slope stability lead to increased landslide activity near the road. Distance from deforestation and distance from road were therefore used as predictor variables in addition to terrain attributes (see next section) and distance to past landslide locations. Note that all distance values above a certain cutoff value (100 m for distance from past landslides, 300 m otherwise) were reduced to that cutoff value to avoid overfitting.

In the present study, two landslide susceptibility models are fitted to STOYAN'S (2000) landslide inventory of 1976, and evaluated based on the inventories of 1989 and 1998. This evaluation using "future" inventories is intended to avoid reporting overoptimistic error measures that would be obtained by measuring the performance on the training data itself (cf. BRENNING 2005).

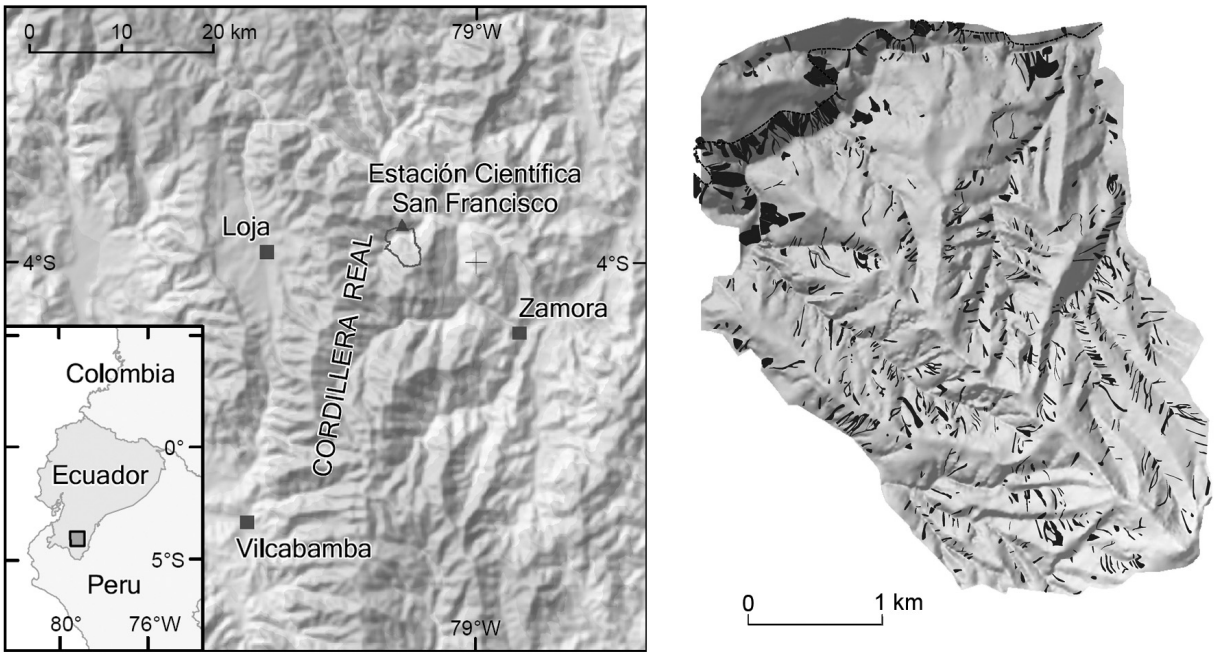


Fig.1: Location of the study area (left) and spatial distribution of all landslides observed between 1962 and 1998 (right).

3 TERRAIN ANALYSIS AND STATISTICAL METHODS

In addition to soil and geological controls on landslide susceptibility, for which accurate data is unavailable in many mountain areas, terrain attributes derived from digital elevation models (DEMs) have been identified as important influences on the spatial distribution of landslides (e.g. MONTGOMERY & DIETRICH 1994, ATKINSON & MASSARI 1998, OHLMACHER & DAVIS 2003). The importance of terrain attributes is related to the physical controls on slope stability (slope angle, upslope contributing area; MONTGOMERY & DIETRICH 1994), but also to the fact that terrain attributes are proxies for a variety of processes and substrate properties (e.g. soil erosion and accumulation; soil moisture content, soil texture). The present work estimates landslide susceptibility based on selected terrain attributes (slope angle, plan and profile curvature, size and slope of the upslope contributing area) and land-cover related variables (distance to road, to deforestation, to past landslide locations). Terrain attributes are derived from a DEM created by STOYAN (2000) from topographic map data enhanced by air photo evaluation.

Since substrate properties – including variables that vary in time such as soil wetness – are usually not known at all or not in detail, the combination of a variety of terrain attributes in empirical-statistical models is often more promising than the application of physically based model with insufficient data. Logistic regression is among the most widely used methods in landslide susceptibility modeling and many other applications where a "soft" discrimination between the presence and absence of a phenomenon is desired (e.g. ATKINSON & MASSARI 1998, OHLMACHER & DAVIS 2003).

While logistic regression has proven to be competitive with some of the highly flexible machine-learning techniques that tend to overfit to spatially dependent data (BRENNING 2005), it is unable to model nonlinear relationships that are often found in landslide patterns. The generalized additive model (GAM: HASTIE & TIBSHIRANI 1990; implemented in the R package gam: HASTIE 2006) extends the logistic and linear regression models by introducing nonlinear transformations of the predictor variables while maintaining an interpretable additive structure, i.e. the effects of different predictors are still added up. The study of BRENNING *et al.* (2007) is a recent example of the use of the GAM in geomorphological research based on terrain attributes as predictors.

The receiver operating characteristic (ROC) curve is suitable for assessing the predictive performance of soft classifiers independently of a specific decision threshold that could be used to delineate unstable areas. The ROC curve represents all possible combinations of sensitivities (fraction of true positive predictions) and specificities (fraction of true negative predictions) that can be achieved by a given soft classifier. The area under the ROC curve (AUROC) takes a value between 0.5 (no discrimination) to 1.0 (perfect discrimination), which summarizes the ability of a classifier to separate landslide and non-landslide locations.

4 THE RSAGA PACKAGE

The RSAGA package provides R geocomputing functions that make use of the command line interface of SAGA GIS, saga_cmd.exe, to execute SAGA GIS modules. RSAGA is currently limited to the Windows platform. The RSAGA package consists of the following three general components.

1. The **RSAGA core** provides low-level access to the SAGA command line program through the `rsaga.geoprocessor` function, which converts R arguments into a command that is executed by the operating system. It also accesses SAGA module libraries and help files to give information on available modules (function `rsaga.get.modules`) and command line options (`rsaga.get.usage`). Most RSAGA users will however find the information they need in the R help files, unless they wish to use a SAGA module that has not yet been integrated in R.
2. The **RSAGA modules** part of the package implements a series of interface functions that allow the R user to run SAGA modules from within the accustomed R environment. Examples are functions such as `rsaga.local.morphometry` (or the derived `rsaga.slope`) and `rsaga.flow.accumulation`, which will be applied later in this work. These functions are often more convenient than the corresponding SAGA modules; as an example, flow accumulation algorithms can be chosen by their

name or acronym (e.g. "mfd" for the multiple flow direction algorithm) as an alternative to the integer code required by SAGA itself.

3. The **RSAGA grid tools** are a compilation of additional useful functions for geocomputing and data input/output. This package section implements, for example, a generic function for applying user-defined functions on moving windows (`focal.function`). This allows the user to implement application-specific terrain attributes (such as the wind shelter index for snow distribution modeling; PLATTNER *et al.*, 2004) or experiment with new filters. The family of "pick" functions (e.g. `pick.from.ascii.grid`, `pick.from.shapefile`) furthermore allows to quickly add data from a variety of sources as new variables to a data frame in R. (It should be noted that these functions are currently rather inefficient for large data sets because they are fully implemented in R, which is an interpreter language.)

5 LANDSLIDE SUSCEPTIBILITY MODELING USING R AND RSAGA

Key parts of the R implementation of the present analysis are outlined in this section. The starting point is a set of ASCII grids representing the DEM, distance to road, the five landslide inventories, and distance to deforestation at the five time points. (The RPyGeo package of BRENNING (2007b) could have been used to derive the distance grids from vector data, and to export the original ESRI rasters to ASCII format, but integration with ArcGIS is not the focus of this work.)

After converting all ASCII grids to SAGA format with RSAGA by calling

```
rsaga.esri.to.sgrd(dir(pattern = glob2rx("*.asc")))
```

local morphometric parameters are calculated with the method of ZEVENBERGEN & THORNE (1987), sinks are filled, and contributing area size and slope are determined with the multiple flow direction (MFD) algorithm (QUINN *et al.* 1991):

```
rsaga.local.morphometry("dem", out.slope = "slope",
      out.hcurv = "hcurv", out.vcurv = "vcurv",
      method = "poly2zevenbergen")
rsaga.sink.removal("dem", out.dem = "sdem", method = "fill")
rsaga.parallel.processing("sdem", out.carea = "carea",
      out.cslope = "cslope", method = "mfd")
```

The results are then converted back to ASCII format with the `rsaga.sgrd.to.esri` function.

To obtain a compact training sample with a sufficient number of landslide pixels, random samples of 500 landslide and 500 non-landslide pixels are drawn from each of the landslide inventory data sets. Without going into the details of the sampling, the terrain attributes are added to the table (data frame `d`) with the sampling locations using commands such as

```
d = pick.from.ascii.grid(d, "slope")
```

This uses nearest neighbor interpolation, which is just fine since the sample locations coincide with grid nodes. Similar "pick" functions are available in RSAGA to conveniently match data from, for example, point shapefiles to R data frames using either nearest neighbor or kriging interpolation. The time-dependent grids (landslide inventories and deforestation distances) are matched to the samples in a similar way.

Spinograms (as produced by the R function `spineplot`) are a simple means for simultaneously plotting the empirical frequency of landslide occurrence conditional on a predictor variable, and the empirical frequency of that variable (see Fig. 2).

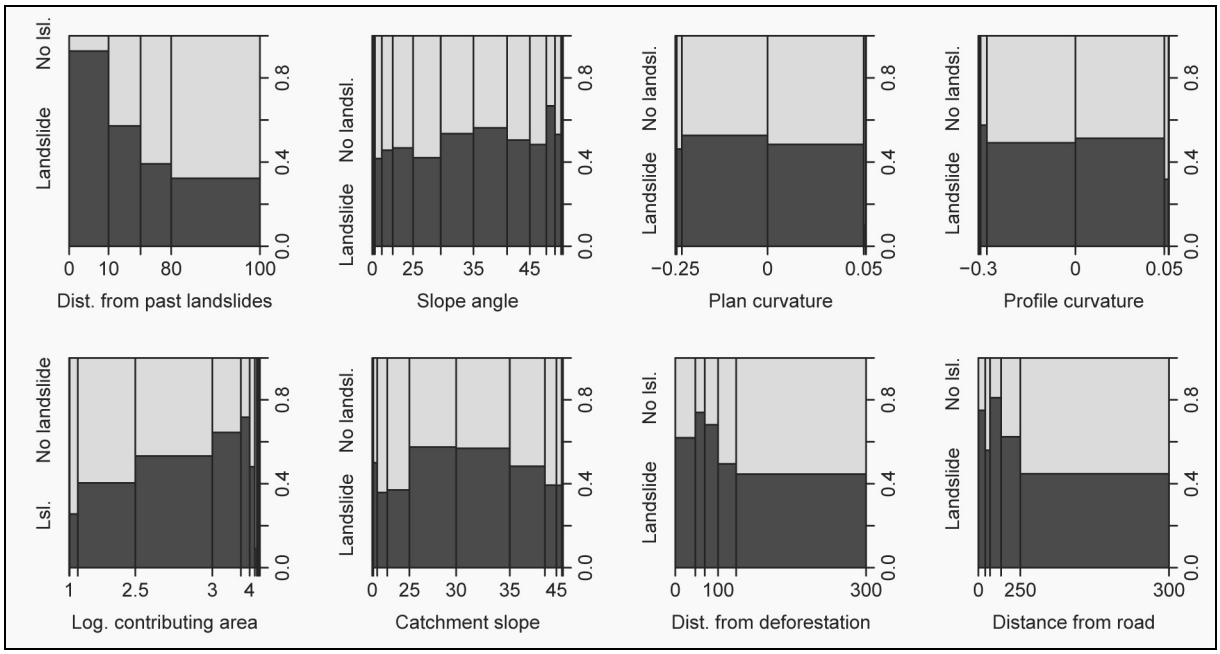


Fig. 2: Empirical frequencies of landslide occurrence in the 1976 training data conditional on the predictor variables. Bar widths in these spinograms are proportional to the empirical frequency of the given interval of predictor values.

The generalized additive model (GAM) is built using a stepwise variable selection starting from the null model and using the Akaike Information Criterion (AIC) as selection criterion. The scope of the model is set up to allow the GAM to choose in the variable selection between (a) including a variable as a nonparametrically transformed predictor consuming two equivalent degrees of freedom, (b) including it as a linear predictor, or (c) omitting it. The AIC penalizes the model size and therefore guarantees that the final model is compact. The R implementation is as follows:

```
library(gam)
gam.scope = list(
      ~ 1 + distroad + s(distroad,2),
      ~ 1 + I(distdeforest > 0) + distdeforest + s(distdeforest,2),
      ~ 1 + log.carea + s(log.carea,2),
      ~ 1 + cslope + s(cslope,2),
      ~ 1 + hcurv + s(hcurv,2),
      ~ 1 + vcurv + s(vcurv,2),
      ~ 1 + slope + s(slope,2) )
fit0 = gam(slides ~ 1, data = train, family = binomial)
fit = step.gam(fit0, scope = gam.scope,
      direction = "both", trace = TRUE)
pred = predict(fit, newdata = test, type = "response")
```

In addition to this model, which will be referred to as GAM1, a second model GAM2 is constructed that is offered an additional variable, the distance from past landslides (i.e. from the previous inventory) as a candidate predictor in stepwise variable selection. The nonlinear transformations both GAM models are shown in Figures 3 (GAM1) and 4 (GAM2).

In R, fitted statistical models can conveniently be applied to data available as R tables ("data frames"). A simple approach for applying fitted models to stacks of raster files representing the predictor variables is however not available. RSAGA provides an interface for applying focal functions on moving windows to raster data sets. This approach is currently being extended to be able to deal with model predictions on grids in the next RSAGA release (version 0.9-1).

The following R statement outlines how this will be resolved (implementation details may still change); the resulting prediction map is shown in Figure 6:

```
focal.function(
    in.grids = c("slope", "hcurv", "vcurv", "carea",
                 "cslope", "distroad", "distdeforest"),
    out.varnames = "pred", sep=""),
    fun = grid.predict, control.predict = list(type = "response"),
    fit = fit, trafo = my.trafo )
```

In detail, the `grid.predict` function passed on to `focal.function` is responsible for calling the `fit` object's actual `predict` method; since the raw variables are often modified before going into a model as predictors (example: log-transformation of the contributing area variable), the `trafo` argument allows for user-defined variable transformations as defined in the `my.trafo` function that will be automatically applied to the grid values taken from the `in.grids`.

6 RESULTS AND DISCUSSION

Both models, GAM1 and GAM2 (Figures 3 and 4, respectively), share the $\log_{10}$-transformed upslope contributing area (`log.carea`) and the profile curvature (`vcurv`) as nonlinear terms; the transformations are practically the same in both models, with an obvious nonlinearity in `log.carea` and a nearly linear behavior for `vcurv`. GAM1 (Figure 3) furthermore uses linear terms for distance from road and slope angle; as expected, steep slopes and areas close to the road are more likely to be unstable in GAM1, and this is especially true where the upslope contributing area is "not too small" (at least a few hundred square meters) and where the profile curvature is convex (i.e. positive).

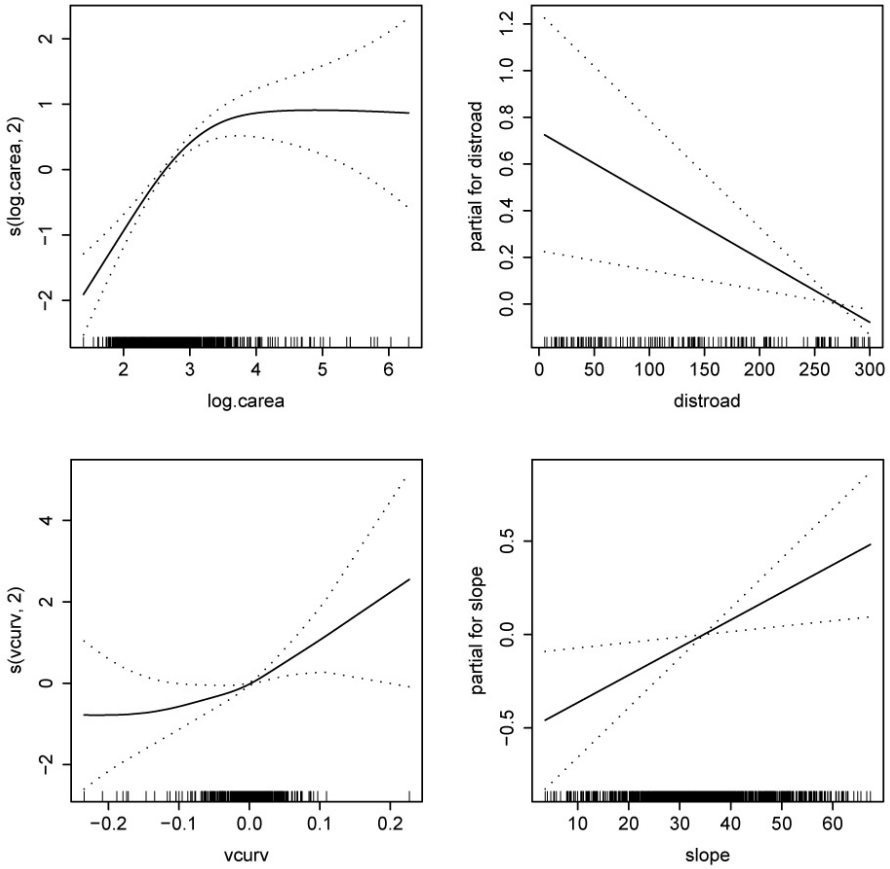


Fig. 3: Transformations of predictor variables in the generalized additive model that *does not use* distance to past landslide locations as a predictor variable.

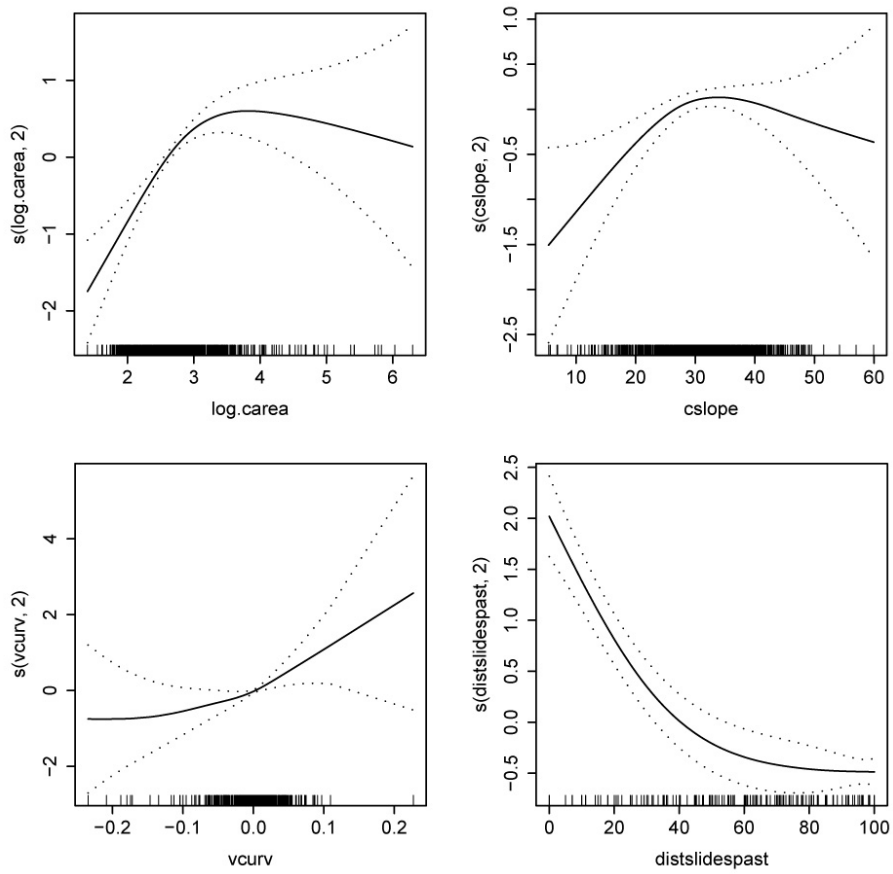


Fig. 4: Transformations of predictor variables in the generalized additive model that *uses* distance to past landslide locations as a predictor variable.

The second model (GAM2; Fig. 4) indicates that landslides are very likely to occur at or near past landslide locations (within <30 m distance), which relates to the reactivation or lateral growth of existing landslide scarps. Since landslide density was always highest near the road and near deforestation areas, the past-landslide variable makes it unnecessary to use the road and deforestation related variables. A steep upslope contributing area (>20° slope angle) is an additional risk factor with a nonlinear influence, as well as a "not too small" contributing area and a convex profile curvature.

The predictive performance of both models on the 1989 and 1998 data sets, i.e. in the prediction of "future" landslides, is visualized in Figure 5 as ROC curves. AUROC values range between 0.67 and 0.77, which is reasonable for landslide susceptibility models evaluated on independent test data sets (compare BRENNING 2005). The GAM2 model, which uses the distance to past landslide locations as a predictor variable, achieves higher AUROC values than GAM1. This is due to the relatively high sensitivities (true positive rates) achieved by GAM2 at low false positive rates (high specificities), which is related to reactivated landslides that can easily be detected based on distance from past landslides.

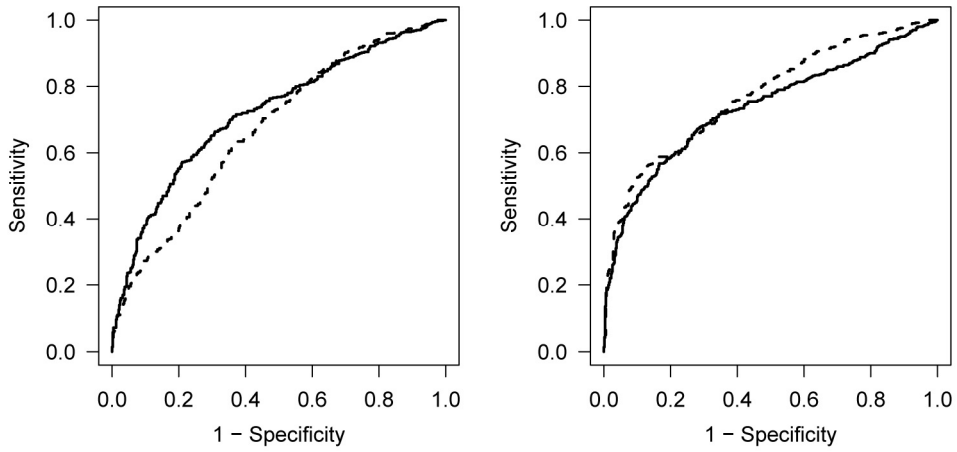


Fig. 5: ROC curves of landslide susceptibility predictions generated by the generalized additive models trained on the 1976 landslide inventory (left: GAM1, not using distance to past landslides as predictor variable; right: GAM2, including this variable). ROC curves correspond to predictions for 1989 (solid line; AUROC = 0.72 for GAM1 and 0.74 for GAM2), and 1998 (dashed line; AUROC = 0.67 for GAM1 and 0.77 for GAM2). See also Fig. 6.

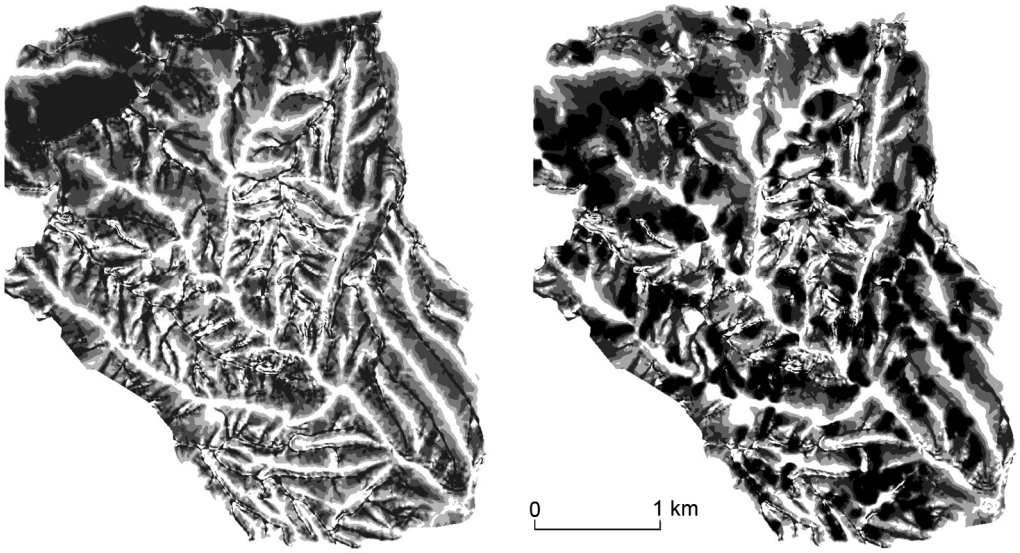


Fig. 6: Landslide susceptibility map for 1989 predicted by the generalized additive models GAM1 (left, *not* using distance to past landslides as predictor variable) trained on the 1976 landslide inventory, and GAM2 (right, using the distance variable). See also Fig. 5.

7 CONCLUSIONS

The present application of terrain analysis methods and generalized additive models to landslide susceptibility analysis provides an example of how GIS-based tools can be combined with powerful statistical models. The access to SAGA GIS modules through its command line interface facilitates the integration with R as implemented in the RSAGA package (BRENNING 2007a). Drawbacks of the present implementation are the need for frequent file conversions between SAGA and ASCII grid formats and the platform dependence (currently only Windows) because of the direct call to the command line interface program. It will be necessary to synchronize RSAGA with major steps in SAGA development, and to provide some level of forward and backward compatibility with SAGA GIS in the implementation and documentation of RSAGA.

In an applied context, terrain attributes are widely used not only in geomorphology and glaciology (e.g. ATKINSON & MASSARI 1998, PLATTNER *et al.* 2004, BRENNING *et al.* 2007, and the present study), but also in pedometrics (e.g. BÖHNER & SELIGE 2006) and Precision Agriculture (e.g. BRENNING *et al.* 2008). These are some of the fields of research and application of terrain analysis and environmental statistics that may benefit from a better integration of GIS and statistical software, or from an integration of SAGA GIS and R in particular. The RSAGA package is a contribution to this integration that is hopefully useful to many spatial modelers.

Acknowledgements: The author thanks R. STOYAN, Erlangen, for providing the landslide inventory and additional thematic data from the ECSF area, and the SAGA GIS developer team for creating and maintaining this excellent geocomputing environment.

REFERENCES

ATKINSON, P.M. & R. MASSARI (1998): Generalised linear modeling of susceptibility to landsliding in the Central Apennines. – Italy, Computers & Geosciences 24: 373-385.

BIVAND, R.S. (2000): Using the R statistical data analysis language on GRASS 5.0 GIS data base files. – Computers & Geosciences, 26: 1043-1052.

BIVAND, R.S., PEBESMA, E. J. & V. GÓMEZ-RUBIO (2008): Applied spatial data analysis with R. – Springer.

BÖHNER, J. & T. SELIGE (2006): Spatial prediction of soil attributes using terrain analysis and climate regionalisation. – In: BÖHNER, J., MCCLOY, K.R & J. STROBL (Eds.): SAGA – Analyses and Modelling Applications. – Göttinger Geographische Abhandlungen 115: 13-28 and 118-120.

BRENNING, A. (2005): Spatial prediction models for landslide hazards: review, comparison and evaluation. – Natural Hazards and Earth System Sciences 5(6): 853-862.

BRENNING, A., GRASSER, M. & D. FRIEND (2007): Statistical estimation and generalized additive modeling of rock glacier distribution in the San Juan Mountains, Colorado, USA. – Journal of Geophysical Research 112: F02S15.

BRENNING, A. (2007a): RSAGA: SAGA geoprocessing and terrain analysis in R. – R package version 0.9.

BRENNING, A. (2007b): RPyGeo: ArcGIS geoprocessing in R via Python. – R package version 0.9-0.

BRENNING, A., PIOTRASCHKE, H. & P. LEITHOLD (2008): Geostatistical analysis of on-farm trials in Precision Agriculture. – Proceedings, GEOSTATS 2008, Eighth International Geostatistics Congress, December 1-5, 2008, Santiago, Chile, 6 pp. (accepted).

GUTHRIE, R.H., DEADMAN, P.J., CABRERA, A.R. & S.G. EVANS (2008): Exploring the magnitude-frequency distribution: a cellular automata model for landslides. – Landslides 5: 151-159.

HASTIE, T. (2006): gam: generalized additive models. – R package version 0.98.

HASTIE, T. & R. TIBSHIRANI (1990): Generalized additive models. – CRC Press, Boca Raton, Florida.

IHAKA, R. & R. GENTLEMAN (1996): R: a language for data analysis and graphics. – Journal of Computational and Graphical Statistics 5: 299-314.

LOZANO, P., BUSSMANN, R.W. & M. KÜPPERS (2005): Landslides as ecosystem disturbance – their implications and importance in South Ecuador. – Lyonia - A Journal of Ecology and Application 8(1): 67-72.

MONTGOMERY, D.R. & W.E. DIETRICH (1994): A physically based model for the topographic control on shallow landsliding. – Water Resources Research 30(4): 1153-1171.

OHLMACHER, G.C. & J.C. DAVIS (2003): Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. – Engineering Geology 69: 331-343.

PLATTNER, C., BRAUN, L.N. & A. BRENNING (2004): The spatial variability of snow accumulation on Vernagtferner, Austrian Alps, in Winter 2003/2004. – Zeitschrift für Gletscherkunde und Glazialgeologie 39: 43-57.

QUINN, P.F., BEVEN, K.J., CHEVALLIER, P. & O. PLANCHON (1991): The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. – Hydrological Processes 5: 59-79.

R DEVELOPMENT CORE TEAM (2008): R: a language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

WILCKE, W., VALLADARAZ, H., STOYAN, R., YASIN, S., VALAREZO, C. & W. ZECH (2003): Soil properties on a chronosequence of landslides in montane rain forest, Ecuador. – Catena 53: 79-95.

ZEVENBERGEN, L.W. & C.R. THORNE (1987): Quantitative analysis of land surface topography. – Earth Surface Processes and Landforms 12: 47-56.

THE USE OF SAGA AS A MOBILE FIELD-TOOL IN THE ENVIRONMENTAL GEOCHEMISTRY

Wolfgang Czegka [1] & *Frank W. Junge* [1]

[1]Sächsische Akademie der Wissenschaften Leipzig, Karl-Tauchnitz-Straße 1, D–04107 Leipzig (Germany)

Abstract: As represented in CZEGKA (2007b) the working group Pollutant dynamics of the Saxon Academy of Sciences evaluated to what extent open source or Low Cost GIS software is able to substitute proprietary commercial GIS products in a mobile use on a notebook /sub notebook in field work. We decided to use the scientific software SAGA (Versions 1.2 – 2.02). With three short examples (Navigation on Lakes, in situ sampling preview, and GPS/Echo sound bathymetric sampling) we will present in which way SAGA was used as a GIS tool in mobile operation. As result we can say that in contradiction to the conclusions made for complex use cases (cf. ZORNIG & KIENBERGER 2006), in use cases with an clear and stringent application profile the use of a low cost GIS like SAGA is practicable and eligible.

1 INTRODUCTION – WHY USING A LOW-COST GIS

The field of work of the multidisciplinary working group "pollutant dynamics" of the Saxon academy of the sciences at Leipzig is the long-term examination of ecological damages in terrestrial and aquatic systems in the Central German region. A special focus is the quantification of the temporal and spatial alteration of the heavy metal load and to recognize principles of the transport and sedimentation and the discharge of heavy metals as well as the rise of the New Central German Lake District (NCGLD).

Our area of interest is Central Germany, especially the Saale Elster catchment area with the focus onto the New Central German Lake District (Mitteldeutsches Neuseenland) (compare Hanisch et al. 2005, Czegka et al. 2008). GeoInformation Systems (GIS) are used in general by the work group simple as documentation and analysing tool.

As represented in CZEGKA (2007b) the working group evaluated to what extent open source or Low Cost GIS software is able to substitute proprietary commercial GIS products in a mobile use on a notebook /sub notebook in field work. In primary use not complex GIS analyses are in the foreground, (compare ZORNIG & KIENBERGER 2006, CZEGKA 2007a, 2007b) but the use of a slim Open Source software at the portable data acquisition and visualization in the field. As background of this purpose the price collapse at Notebook/Sub notebooks can be seen . The use of notebooks enables more flexibly than the use of PDA or handhelds (THOMAS & GARBE 2003). One of the conclusion described in CZEGKA (2007b) was the use of SAGA (1.2, 2.0 RC3, 2.02) chosen out of a bunch of "similar" GIS programs (PELCH 2007) for our normal field work.

In the following we will present three examples (Navigation on Lakes, in situ sampling preview, and GPS/Echo sound bathymetric sampling) in which SAGA was used as a GIS tool in mobile operation.

2 SAGA APPLICATIONS IN MOBILE OPERATIONS

Navigation on lakes: To outline the problem is short: the main task is finding a defined sampling area/ point on a lake. Background is the "Sedigas project" where methane gas sampling is done at lakes which were former lignite open pits. These "open pit"- lakes are stamped, in contradiction to "normal lakes, by a small scaled bathymetric profile in which depth steps drastically change in the 5-10 m range. The gas sampling is carried out at deep steps of zero, 2 , 5.. m. in different defined depths.

One the other side the gas production is dependent on the nutrient supply which depends on the general flows. The sampling area/point has to be selected and decides on situ on the lake within the predefined sampling areas or a discrete circle about a sampling-point. A direct to point navigation which can easily done with simple GPS trailers is in this case less favourable. On lake, without direct orientation points, this task can be done with a electronic bathymetric map a GPS mouse coupled with a notebook and SAGA (via GPS Babel) .

In situ Sampling preview: Normally gas sampling on lakes is due to his slowliness a more than a one day campaign so it is more favourable creating a sampling preview during the field campaign to give an overview. Not only the simple "sampling-point" but also easy to measure parameters like degassing rate (cm^3/h) temperature, conductivity, chlorophyll –α etc. This "first view" gives us a possibility for first estimation and coordination of the sampling strategy.

GPS/echo sound bathymetry of lakes: One of the basic data for sampling on lakes is the bathymetric map. Normally at the new rising lakes these data were drawn from mine surveying maps of the former lignite mine. But in the NCGLD there are several lakes from older provenience. Samples for these type of lakes are the Lake Rust from Osternienburg area or the Lake Muldestausee. Lake Rust developed himself by a collapse from a subterranean lignite mine between 1920- 1950, Lake Muldestausee was a planned flooding of an former lignite mine with the river Mulde in 1975. The solution was to generate bathymetric maps by using a WAAS/EGNOS GPS trailer and echo sound (Magellan /Furuno echo Sounder FE 606). The x,y-data were generated via GPS-Trailer, SAGA and GPS Babel – the deepness "z"-data manual integrated from a manual readout from the Furano echo sounder. These xyz data determined by the GPS echo sounding coupling were interpolated to a about an "inverse distance" procedure (Fig. 1).

The sea ground relief has a special influence on stratification and circulation of the depth water. As we can see the bottom of lake Rust is still 60 years after his creation moulded by the former support columns of the former subterranean lignite mining. The hollows between the these pillars sank after mining was given up and formed in slow process triggered by halokinetic formed surrounding the outscape of the later lake. The columns rested still arise and can still be detected.

3 CONCLUSION

In contradiction to the conclusions made for complex use cases drawn in ZORNIG & KIENBERGER (2006), in use cases with an clear and stringent application profile the use of a low cost GIS like SAGA is possible and eligible. With FreeGIS applications coming out of scientific institutions ("scientific software") as SAGA a custom designed and competitive solution is practicable. Because scientific software addresses primary to informed users by all a balancing act between user friendliness and costs has to be made. In the case "field use" SAGA as shown in upper use cases can be used in daily field work application easily.

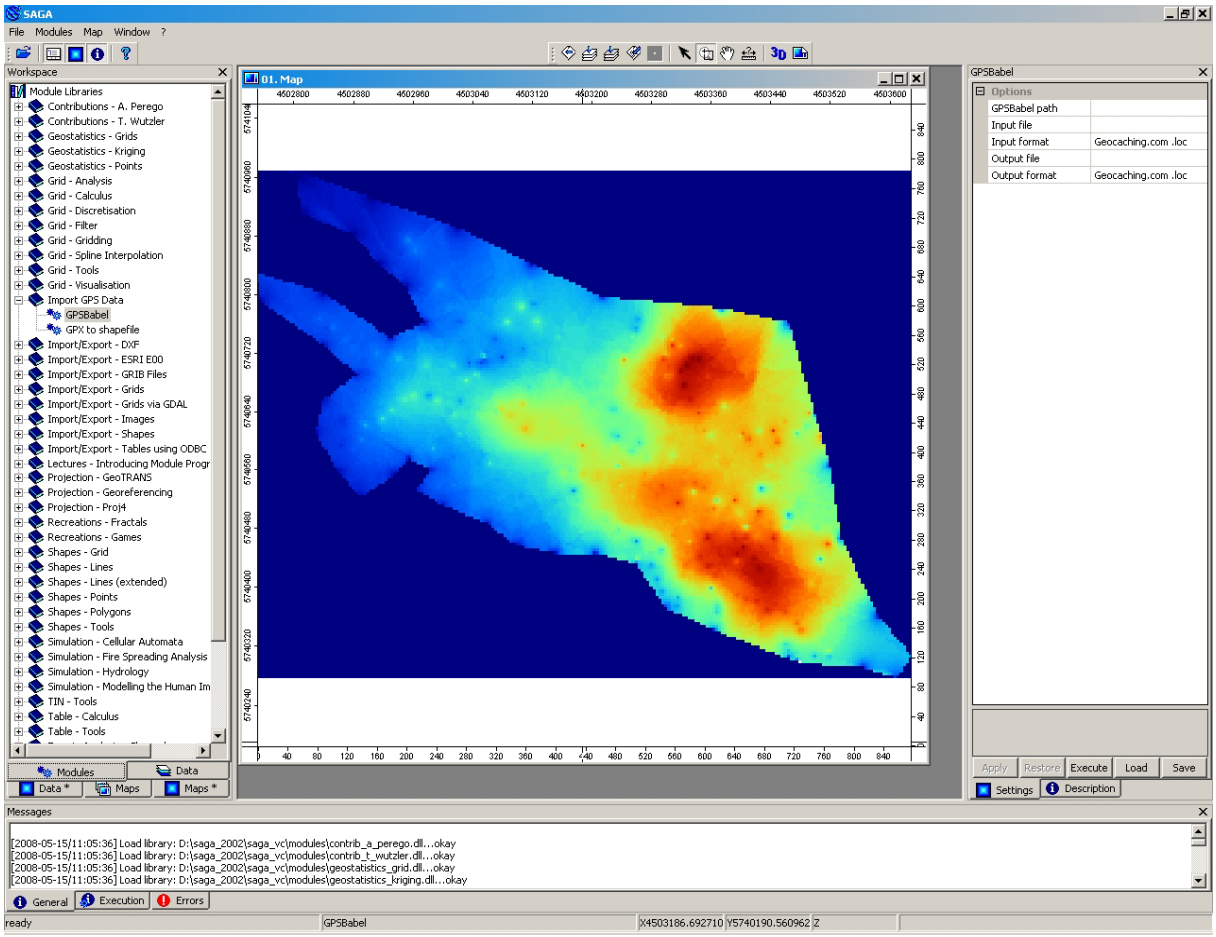


Fig.1: Bathymetric map of lake Rust (Osternienburg, Saxony-Anhalt) sampled with GPS and echo sound. The discolour presentation is for a use in field work better than the normally used "natural gradation".

REFERENCES

CZEGKA, W. (2007a): Das Open Source GIS Programm SAGA und seine Anwendungsmöglichkeiten in den Geowissenschaften. – Aufschluss 58: 153-159.

CZEGKA, W. (2007b): Open Source und Low Cost GISysteme im portabelen Einsatz in der Umweltchemie. – In: STROBL J., BLASCHKE T. & G. GRIESEBNER (Eds.): Angewandte Geoinformatik 2007, Wichmann-Heidelberg: 137-143.

CZEGKA, W., JUNGE, F.W., HAUSMANN, J., KUCHENBUCH, V. & R. WENNRICH (2008): Hydrochemical und geochemical parameters of new developing lakes of the Neuseenland (New Central German lake district), an overview. – ZDGG 159: 141-154.

HANISCH, C., ZERLING, L., JUNGE, F.W. & W. CZEGKA (2005): Verlagerung, Verdünnung und Austrag von schwermetallbelasteten Flusssedimenten im Einzugsgebiet der Saale. Ein Beitrag zur Abwägung von Gefährdungspotenzial und Selbstreinigungsvermögen unserer Fließgewässer. – Abh. d. Sächs. Akad. d. Wiss. zu Leipzig, Math.-nat. Kl. 64, H. 1: 135 p.

PELCH, L. (2007): Low Cost GIS Software Options. – VCGI Newsletter 02/2007. http://www.vcgi.org/commres/publications/newsl_feb2007.pdf (date accessed : 15.04.07).

THOMAS, P. & F. GARBE (2003): Suchen und Kartieren: Erfahrungen mit GPS-Geräten in der Naturschutzpraxis. – http://www.fht-stuttgart.de/fbv/fbvweb/veranstaltungen/GPS_fuer_GIS/pdf_Tagungsband/Low/Bericht_Thomas.pdf (date accessed : 15.04.07)

ZORNIG, J. & S. KIENBERGER (2006): Low-Cost GIS im Produktveinsatz – HIVpathuis GIS als Monitoringtool von AIDS-Patienten in Chibabava, Mosambique. – STROBL J., BLASCHKE T. & G. GRIESEBNER (Eds.): Angewandte Geoinformatik 2006. Wichmann Heidelberg: 793-802.

COLD AIR PRODUCTION AND FLOW IN A LOW MOUNTAIN RANGE LANDSCAPE IN HESSIA (GERMANY)

Helge Dietrich [1] *& Jürgen Böhner* [1]

[1]Institute of Geography – Section Physical Geography – University of Hamburg, Bundesstr. 55, D-20146 Hamburg (Germany)

Abstract: This paper presents methods for analysing the impacts of different terrain parameters on cold air production, flow and cold air accumulation. These methods have been tested using a Digital Terrain Model (DTM) in combination with point source nocturnal temperature data observed during autochthonous calm and clear weather conditions. The gently undulated test site Melsungen in Northern Hessia features 130 single observation points representing a range of different morphological situations. Assuming the terrain reveals a seasonal varying influence on nocturnal air temperature patterns near ground, a comparative study was carried out between the meteorological transitional period showing active vegetation cover and the colder season without any active vegetation cover. Using stepwise multiple linear regression analysis, different terrain parameters have been tested. The results reveal changing influences of terrain attributes on nocturnal air temperature pattern depending on seasonal variations of vegetation cover. During the transitional period, a maximum variance of 55 % is explained by a combination of different terrain parameters and in the colder season even up to 68 %.

1 INTRODUCTION

Local climate is the result of complex interactive processes between different influencing factors. Small scale climate variability often is of great importance and relevance to human-bioclimatologic questions as for example with regard to thermal comfort and aeration conditions in built-up or residential areas. The most pronounced local climate effects occur during autochthonous weather conditions. These conditions are characterised by low wind velocities and a low relative humidity, a low-level cloud cover and increased radiation fluxes between the earth´s surface and the atmosphere.

However, nocturnal cold air production and flow require further fundamental influencing factors, which are substantially derived from land use and land cover, topography and topology. Determining surface parameters are thermal properties such as heat conduction and storage capacity of the subsurface, surface roughness and further on a multiplicity of different terrain attributes.

In the past a lot of climatologic investigations have been carried out relating to the influence of land use, but to a much lesser extent to the influence of topography and topology on local climate resp. thermal circulations in cities and rural areas. Against this background, nocturnal autochthonous near ground air temperatures have been investigated in a diploma thesis (DIETRICH 2006). The research aimed to improve and extend semi-empirical methods for climate spatial prediction and regional climate modelling, required in case studies and climate impact assessment e.g. in urban planning or agro-meteorological applications. Using land cover data and terrain parameters, it is possible to localise human bioclimatological relevant compensation spaces, connecting structures and ventilation paths etc. Additionally, it is a support in order to quantify the cold air potential of a cold air contributing area as well as to assess agro-meteorological risks due to frost danger in spring.

In 2005 and 2006 the measurement campaign started in the surrounding areas of the Hessian climatic health resort Melsungen. Using a mobile measuring unit, data of near ground air temperature have been collected. The acquisition of data focused on the nocturnal spatiotemporal variability of temperature anomalies, cooling rates and vertical gradients between 50 and 200 cm above ground. The campaign aimed to localise areas, which produce high volumes of cold and fresh air, to detect where katabatic winds preferably occur and where cold air accumulates in large quantities.

To obtain reasonable spatial high-resolution climate information about local observations of site-specific cold air distribution, it was necessary to estimate spatially continuous climate data from point source temperatures, using geostatistical interpolation and terrain parameterisation methods. For that purpose, controlling terrain parameters had to be identified by using a statistical model (stepwise multiple regression). In this context 10 terrain parameters have been analysed and evaluated. The derivation of terrain parameters from DTM as well as the subsequent spatialisation of local climate has been accomplished with SAGA 2.0.

2 STUDY AREA AND MATERIAL

The study area Melsungen in the Fulda Valley (city centre Melsungen 9°33'E, 51°08'N) is located 20 km southeast of Kassel, Hessia (Germany). The model domain containing the empirical data observation net covers an area of approximately 20 km² (Fig. 1). However, the DTM extent of about 42,7 km² exceeds the model domain in order to cover the cold air contributing areas (cf. Chapter 4). The area includes different topographic elements and can be divided into three principle components comprising (1) the central Fulda Valley, (2) the bordering mountain ranges, which gently and sometimes steeply rise up, and (3) two bigger tributary valleys (Kehrenbach, Kesselbach), which lead out of the forested surroundings of Melsungen.

Sub routes 1 and 2 (Fig. 1) are primarily east exposed and present predominantly agricultural areas, whereas sub routes 3 and 4 have a higher percentage of west exposed settlement structures and show less homogeneity in exposition than the other sub routes.

The landscape has a height between 165 m at the bottom of the Fulda Valley and 430 m reached at the tops of the adjacent hills. Geologically being part of the German Trias, the region near Melsungen is situated in the Kurhessisches Bergland, which is dominated by New Red Sandstone with its typical steep terrain (GEOLOGISCHE KARTE VON HESSEN 1975). Besides agricultural activities in the Fulda Valley and the valley flanks, forestry can be found in the upper regions, whereas dense settlement structure resp. suburban and urban areas are restricted to the Fulda Valley. Particularly the valley flanks of the Fulda are structured by a bigger number of middle sized, rill-like slots, which have its seats in the upper parts of the valley flanks and which open out to the main and the two tributary valleys.

The climate of the region is characterised by a mean annual temperature of 8.6°C and a mean annual precipitation of 705 mm for the period of 1950-2003 (DWD 2005). The relative probability of calm and clear weather conditions being required to induce effective cold air production is almost 30% per year (after GERSTENGARBE *et al.* 1993). Actually the probability that thermal circulation patterns may occur is even higher in Central Europe with almost 40 % (HELDT & HÖSCHELE 1989), corresponding to own calculations. Containment criteria after HELDT & HÖSCHELE (1989) have been (1) wind velocity 10 m above ground <= 3 Bft. and (2) daily temperature range > 10°C. An accumulation of these weather types can be observed between April and October (GERSTENGARBE *et al.* 1993). Furthermore, the Fulda Valley shows below average annual wind velocities due to protecting effects of the surrounding terrain, so that the conditions for the development of thermal circulation within the research area can be categorised as well.

Empirical data has been collected at night by mobile temperature measurements. The measurement unit on a caravan has been equipped with three ventilated psychrometers, which have been installed at the forefront of the bus in three different heights above ground (50, 150, 200 cm). Temperature sensors have been technically inspected and calibrated Pt100 thermistors.

Basic substructure of terrain analysis and climate regionalisation has been a DTM with a grid cell size of 2500 m² (50 x 50 m). All terrain parameters, considered in chapter 4, have been derived from this source.

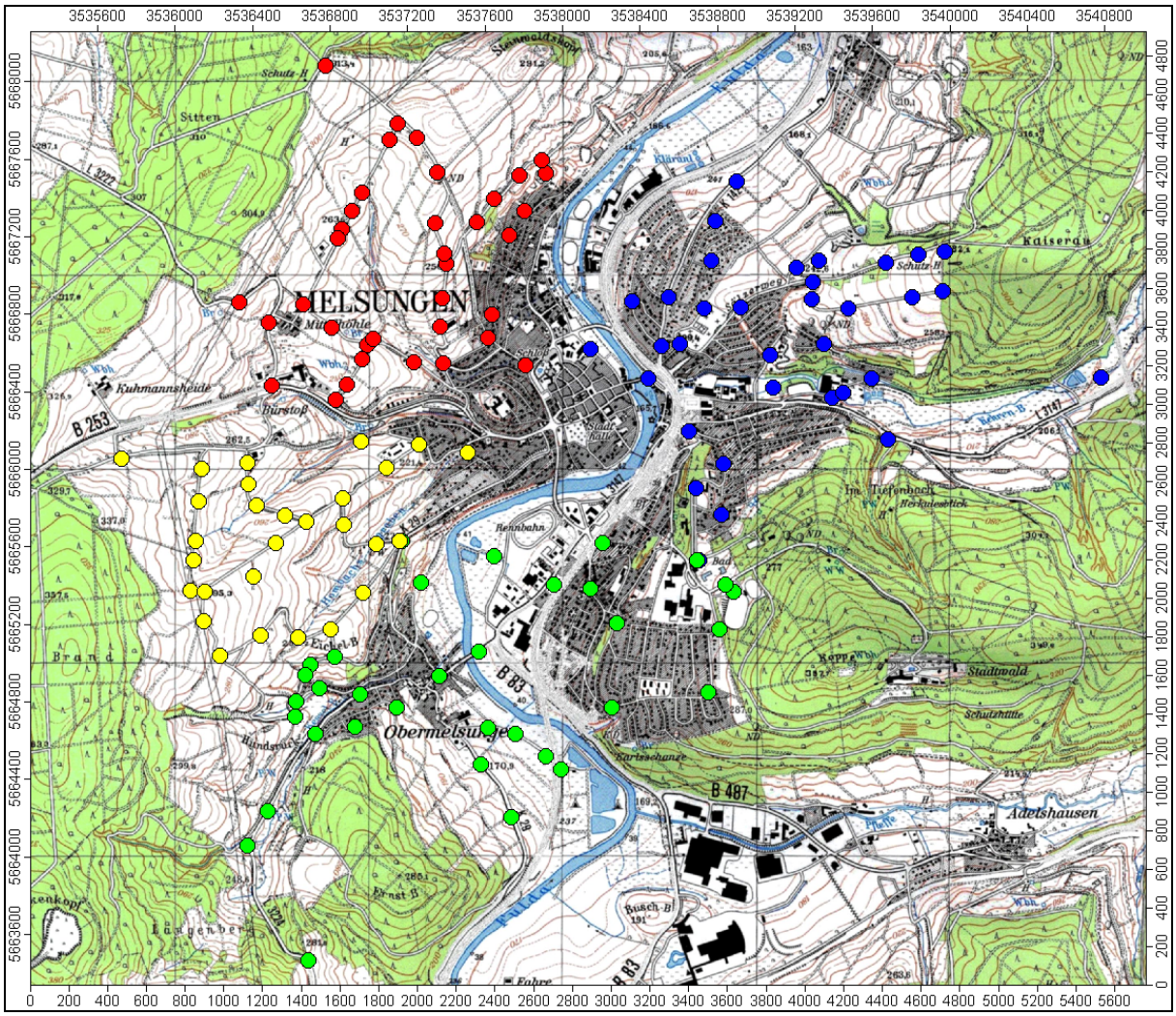


Fig. 1: Test site (observation net) of the climate health-resort Melsungen – route 1 (red), route 2 (yellow), route 3 (green) and route 4 (blue).

3 EMPIRICAL DATABASE

Mobile measurement campaigns deliver spatial high-resolution point source observations of climate parameters (KUTTLER & BARLAG 2002, KUTTLER & DÜTEMEYER 2003). This is an established method to localise areas of cold air production, flow and accumulation (KUNKA & OESTREICHER 2006). Against the background of mobile data collection, the research area has been divided into four individual sub routes including 130 measurement points. The local observation net was conceived under the dictum of a frequent recording of all topographically relevant terrain units like domes, depressions, crest lines, depth lines, valley flanks etc. in order to achieve a detailed topoclimatic representation.

All sub routes have been connected by a common initial and final point. This was necessary firstly to merge all sub routes and secondly to eliminate diurnal temperature variations which automatically would have influenced the results due to the duration of sampling (max. 1.5h for each sub route). In order to compensate this methodological error, temperature data of each sampling run had to be synchronised on the first moment of the sampling run (DANZEISEN 1983, STÜLPNAGEL VON 1987). Basis of the corrections have been the natural cooling rates of the initial point, which have been linearly allocated to the empirical temperature data (GROSS *et al.* 1996, PÜTTMANN 2002). Only after correcting, all data are comparable with each other (trend adjustment).

The empirical study about the influence of terrain attributes on cold air distribution was subject to a seasonal assignment. The analysis aimed to obtain information about seasonal variations in the behaviour of local thermal circulations caused by phenological variations and the resulting changes in the influences of the terrain. In contrast to a multitude of other climatologic investigations dealing with local climate during summer, this study tried to explore nocturnal autochthonous near-ground air temperature patterns in spring, autumn and winter. Hence, on the one hand the warmer part of the meteorological transitional period [TP] between winter and summer was inspected (April, May, September), and on the other hand the colder season [CS] without active vegetation cover except for withered grass and without any snow cover (November, December) which would have caused difficulties in carrying out mobile measurements. The specification of sample periods has been geared to the methodical differentiation of thermal seasons after RAPP & SCHÖNWIESE (1994). However, data volume of the colder season is not comparable to the amount of data of the transitional period owing to a rather infrequent occurrence of adequate synoptic weather conditions. By that reason, sub routes 3 and 4 could not be investigated in winter.

In order to comprehend the spatiotemporal process related differentiation of cold air production, flow and accumulation during the whole night, every measurement event was split up into four sections. The first sample section started directly after sunset and the second one immediately subsequent to the first one. The third section started approximately four hours before sunrise followed by the fourth section which finished directly before sunrise. The mobile measurement unit was stopped at every observation point for at least 90-120 seconds, so as to account for the intrinsic time constant of the thermistor and thus to minimise systematic errors as well as to intercept outliers. A total of more than 50 single sample runs have been carried out in 16 nights.

Three **temperature parameters** have been derived from the original data. First of all **anomalies** in two different heights above ground (50, 200 cm) have been calculated in order to match sampling runs featuring different temperature levels and to display positive and negative deviations of single local observation points resp. sub spaces of the research area (ROSNER 1992). The arithmetic mean of the corrected absolute temperatures of all observation points of a sub route acted as a reference isotherm (ROSNER 1992). The anomalies resulted from the deviation of a single value from the reference isotherm (temperature standardisation). The range of anomalies indicates the intensity of the spatial temperature pattern, the higher the range the more heterogeneous the pattern. Anomalies in 200 cm above ground have been considered in the subsequent regionalisation (Fig. 3).

Vertical **temperature gradients** as an absolute climate variable result from the differences between absolute temperatures in 50 cm and 200 cm above ground. Gradients do not have to be synchronised like anomalies because this parameter only serves as a pure snap-shot of the atmospheric condition in relation to specific locations. Basically, gradients inform about the near-ground atmospheric layering. Negative temperature gradients show a lower temperature level near the ground than in 2 m observations. Negative gradients indicate a stable layering with initial cold air production or flow, whereas positive gradients suggest heat emission of the surface, preferential in build-up areas. On this account temperature gradients have been generally used solely to detect sub spaces showing initial cold air production.

Finally, nocturnal **cooling rates** have been observed, but in opposite to the parameters mentioned above, only at a few selected measurement points. A specific cooling rate results from the temperature difference between two observations within one hour. The approach aimed to prove unequal cooling trends of different sub spaces. So only a few observation points had to be investigated. Moreover, cooling rates of the common initial and final point are of great importance referring to the elimination of diurnal temperature variations being inherent in original data.

4 TERRAIN ANALYSIS

There are many scientific publications dealing with digital terrain analysis and modelling of geo-systematic processes. Application fields of terrain analysis are for instance aeolian erosion processes (BÖHNER *et al.* 2004), pedological processes (BÖHNER & KÖTHE 2003, BÖHNER & SELIGE 2006), hydrological denudation (KREIKEMEIER *et al.* 2004) and regionalisation of soil-relevant climatic parameters (BÖHNER 2004). Terrain analysis can be seen as an efficient method for solving topo-climatic problems and analysing and understanding its controlling factors (cf. DIETRICH 2006).

Topo-climatology studies the impact of different land surface parameters on local climate and its spatial variations of near-ground atmospheric processes (BÖHNER & ANTONIC 2007). Spatial climate variability during autochthonous conditions mostly depends on land-cover and topographic variability. Besides terrain parameters like absolute height or slope inclination also the relation between a terrain segment and its environment further away plays a key role for topo-climate. This applies to thermal circulation, too. Already a few decades ago at an early stage of topo-climatic research GEIGER (1961) and HARRISON (1971) recognised the importance of topology for the development of small-scale spatial variability of climatic elements. It is essential to look at different spatial scales of cold air flow (THAMM 2000) because air flow is among the scale-invariant atmospheric phenomena with varying large reference areas resp. horizontal extensions (BENDIX 2004). In this context, the spatial relation to the surrounding area is absolutely indispensable for a sufficient characterisation of a specific terrain segment (BÖHNER & KÖTHE 2003). But considering topology, it is advisable to choose a adequate complex parameterisation of the terrain. Only by doing so, it is possible to accommodate the meaning of topology for small-scale climatic processes.

A total of 10 terrain parameters have been derived from the DTM (Tab. 1). The analysis has been carried out with regard to their impact on the development of slope winds, preferred flow paths, position and size of cold air pools. All below-mentioned terrain parameters are attributed to terrain analytical methods after WILSON & GALLANT (2000), BÖHNER & KÖTHE (2003), BÖHNER & SELIGE (2006) and BÖHNER & ANTONIC (2007).

Tab.1: Investigated terrain parameters.

Parameter	Unit	Description
hsl	[m]	height above sea level
inc	[°]	slope inclination
cao	[m²]	catchment area, original, run off area of above-situated grid cells
cam	[nondimensional]	catchment area, modified (lognormal), run off area of above-situated grid cells
hac	[m]	height above culmination reference to the height difference of immediate adjacent channel lines
hbc	[m]	height below crest line reference to the height difference of immediate adjacent crest lines
hno	[nondimensional]	height, normalised normalised height of the local environment
hst	[m]	height, standardised standardised height of the local environment
hma	[nondimensional]	height, middle slope
eno	[nondimensional]	aspect, normalised (eno) normalised distance of aspect to southwest

- **Catchment area**: The catchment area (cao) specifies the hydrologically defined contributing upslope area (WILSON & GALLANT 2000), frequently considered as an approximation for the cold air potential. The hydrologically defined catchment area, however, is closely related to the channel network and thus fails to model cold air flow particularly in broad valleys. Against this background, a modified catchment area (cam) has been implemented in SAGA (Fig. 2), which yields a better approximation of cold air distribution and accumulation (BÖHNER 2004).
- **Relative heights**: Height above culmination (hac) describes the vertical offset of a grid cell to its according channel line considering the local relief energy in relation to the local erosion base (BÖHNER 2004). Referring to linear depressions like valleys or rills this parameter records processes of cold air convergence and linear run off. Thus, the linear correlation between height above culmination and nocturnal minimum temperatures, which has already been described in HARRISON (1971), is sufficiently considered.
- **Height below crest line:** The height below crest line (hbc) refers to the relative height of a terrain position to a nearby ridge. Crest lines are always zones of cold air divergence with its typical positive temperature anomalies (VOGT 2001). Both channel lines and crest lines determine transport processes like cold air flow or hydrological run off and therefore display morphometrically structuring factors of the terrain concerning the elementary development of nocturnal temperature patterns.
- **Normalised and standardised height:** Besides the vertical offset of a grid cell, normalised height (hno) and standardised height (hst) also consider the extension of a catchment area of a specific terrain point (cf. BÖHNER & SELIGE 2006). Normalised height allots value 1 to the highest and value 0 to the lowest position within a respective reference area. Standardised height is the product of normalised height multiplied with absolute height. Hence, both parameters do not only relate to adjacent channel and crest lines but comprise the widespread topology. This is necessary to illustrate a procedural differentiation of local air flow, because divergent, convergent and linear cold air flow depend on the direct environment, the catchment area, and the relative height differences at the same time (THAMM 2000).
- **Mid-slope position:** The mid-slope position (hma) is commonly considered in topoclimatic analysis, to cover the warmer zones of slopes (BENDIX 2004). This parameter assigns mid-slope positions with 0, whereas maximum vertical distances to the mid-slope in both valley or crest directions are assigned with 1 in order to represent the temperature drop towards upper and lower parts of a slope.
- **Aspect**: Normalised distance of aspect (eno) aims to represent daily air and surface temperature maxima in the later afternoon or early evening caused by anisotrophic heating of slopes (BENDIX 2004). Besides the angle distance to southwest, this parameter also considers mean slope inclination of a terrain unit.

The analysis of relationships between independent predictor variables (terrain) and the dependent predictant variable (temperature) has been carried out using stepwise multiple linear regression ($Y \rightarrow X_1 + ...+ X_n$) with a minimum significance of $p \leq 0{,}05$. Both seasonal observation periods have been studied separately to detect seasonal variations of the influence of the terrain on nocturnal air temperature distributions. In some cases the multiple regression analysis identifies only one predictor parameter, although the correlation matrix indicates more than one parameter being significantly correlated. This effect is called a multi-collinearity problem (BAHRENBERG *et al.* 2003) and advertises that several parameters are redundant, that is to say, some parameters have a similar climatologic functionality showing a high linear correlation.

In the following sections, only **anomalies** in **200 cm** above ground will be discussed in detail, primarily for the reason that these anomalies have the highest correlation of all observed variables. Furthermore, anomalies in 200 cm above ground are elementary of greater relevance to human-climatologic aspects. The high level of correlation in 200 cm above ground is caused by the fact that an increasing ground level strictly heightens the thermal impact of the land cover and simultaneously weakens the influence of the terrain (GEIGER 1961).

Table 2 shows a regression matrix of all sub routes and the cumulative measurement net. It is structured in two observation periods and four sections of a sampling event (one whole night). To the right all investigated terrain parameters are listed including the succession of identification by the stepwise multiple regression analysis as well as the corresponding regression coefficients. The coefficient of determination R^2 and the regression constant are given to the left.

Tab. 2: Correlation between temperature anomalies in 200 cm above-ground and specific terrain indices.

ANOMALIES 200 cm a. gr.					Terrain indices									
Route	Sample Section & "Season"	Regression-constant	R^2 (corrected)		cam	cao	hac	hbc	hno	hsl	hst	inc	eno	hma
1	**Trans. Period [TP]**													
	after sun set	-0.3999	0.44	succesion		1							3	2
				coefficient of gradient		-5,81E-06							16.0200	- 4,6751
	second section	4.1590	0.33	"	1									
				"	-0.3717									
	third section	3.8377	0.34	"	1									
				"	-0.3458									
	before sun rise	-0.6365	0.39	"		1		2						3
		(mean)	**0.37**	"		-4,56E-06		0.0147						- 1,821
	Colder Season [CS]													
	after sun set	2.7183	0.59	succesion	1					2				
				coefficient of gradient	-0.3570					0.0051				
	second section	***no regression (p ≤ 0,05)***		"										
				"										
	third section	-2.6906	0.56	"		2				1	3			
				"		-2,21E-06				0.0141	-0.0048			
	before sun rise	-	-	"										
		(mean)	**0.58**	"										
2	**Trans. Period [TP]**													
	after sun set	0.7185	0.49	succesion		1								
				coefficient of gradient		-8,02E-06								
	second section	0.3350	0.51	"		1								
				"		-4,44E-06								
	third section	-0.0409	0.56	"		1	2							
				"		-4,42E-06	0.0128							
	before sun rise	-2.0841	0.62	"		1				2				
		(mean)	**0.55**	"		-3,35E-06				0.0095				
	Colder Season [CS]													
	after sun set	3.5660	0.75	succesion	3	1				2		4		
				coefficient of gradient	-0.1162	-2,08E-06				-0.0070		-3.6205		
	second section	6.7328	0.57	"	1				2					
				"	-0.5241				-1.8151					
	third section	-	-	"										
				"										
	before sun rise	-2.0672	0.72	"						1				2
		(mean)	**0.68**	"						0.0100				1,9415
3	**Trans. Period [TP]**													
	after sun set	10.4226	0.51	succesion	4					2			3	1
				coefficient of gradient	-0.4403					-0.0199			6.8879	3,4884
	second section	1.9824	0.40	"					1	2				
				"					4.2240	-0.0148				
	third section	5.9781	0.46	"	1									
				"	-0.5161									
	before sun rise	0.8508	0.39	"										1
		(mean)	**0.44**	"										2,8838
4	**Trans. Period [TP]**													
	after sun set	***no regression (p ≤ 0,05)***		succesion										
				coefficient of gradient										
	second section	-1.1099	0.43	"							1			
				"							0.0120			
	third section	-0.7110	0.25	"					1					
				"					1.7231					
	before sun rise	-0.6958	0.22	"					1					
		(mean)	**0.30**	"					1.6735					
ENTIRE	**Trans. Period [TP]**													
ROUTE	after sun set	3.6892	0.36	succesion	2	1						3		
				coefficient of gradient	-0.2775	-3,40E-06						-3.2745		
	second section	-1.3496	0.40	"		2		3	1					
				"		-2,92E-06		0.0129	2.5184					
	third section	2.5461	0.43	"	1	2								
				"	-0.2366	-1,90E-06								
	before sun rise	2.2969	0.29	"	1	2								
		(mean)	**0.37**	"	-0.1941	-1,68E-06								
	Colder Season [CS]													
	after sun set	3.0018	0.39	succesion	1									
				coefficient of gradient	-0.2696									
	second section	0.1509	0.16	"		1								
				"		1,35E-06								
	third section	-	-	"										
				"										
	before sun rise	-	-	"										
		(mean)	**0.27**	"										

Terrain analysis of the transitional period [TP]: The aggregation of all sub routes implicates that the entire measurement net shows a lower correlation (R^2 [TP]: 0,29-0,43, mean [TP]: 0,37) than the separated sub routes (R^2 [TP]: 0,22-0,62, mean [TP]: 0,30-0,55). Presumably the weakening of the correlation of the entire observation net has to be ascribed to the combination of specific thermal land cover and terrain dependent characteristics of all sub routes, by which the detectable influence of the terrain could have been diminished.

During the warmer season the highest R^2 can be found at sub route 2 (R^2 [TP]: 0,49-0,62, mean [TP]: 0,55). By reason of predominant homogenous land cover, the heat balance consistently develops, so that the land cover signal only shows a weak impact on the temperature distribution near ground. Moreover, sub route 2 reveals only minor regional differences in aspect (cf. Chapter 2). Thus, the radiation balance is rather uniform, too. On that account, sub route 2 most likely validates the simplifying assumptions of the terrain parameters, followed by sub route 3 (R^2 [TP]: 0,39-0,51, mean [TP]: 0,44). But in consideration of stronger developed differences in aspect and a higher percentage of urban areas the correlation of sub route 3 is already lower.

Sub route 2 also exemplifies a gradual nocturnal increase of correlation (R^2 [TP]: from 0.49 up to 0,62). This can be explained by a constant influence of topographic factors and topology on temperature patterns, whereas other geographical factors have a decreasing effect on air temperature near ground later in the night (UPMANIS & CHEN 1999). It is to assume that the approximate ideally typical character of sub route 2 is responsible for a clear development of the chronological increase of terrain impact. All other sub routes predominantly show an inverse behaviour because of their higher percentages of urban land cover, which modifies the intrinsic thermal regime of the surface. Related to sub routes 1, 3 and 4, the urban land cover signal seems to overlay the influence of the terrain by intensified long-lasting heat emissions until early in the morning. Besides, higher surface roughness, which is a characteristic property of residential and urban areas, substantially influences orographically induced thermal circulation patterns by partially obstructing cold air flow alongside anthropogenic obstacles (cf. GROSS *et al.* 1996, SCHWAB 2000). This presumably leads to a chronologically diminishing impact of the terrain on temperature over the course of the night.

In the next step different terrain parameters have been analysed with regard to their percentage of explained variance of temperature distribution. In two cases no parameters could be identified by stepwise multiple regression ($p \leq 0{,}05$) due to the weather conditions. Referring to the entire route, catchment areas (cam, cao) are the most explaining terrain parameters during the whole night. The catchment area is applicable to the volume of cold air flow being expected at a specific position (WILSON & GALLANT 2000). Additionally, in the first half of the night also relative heights (hbc, hno) and slope inclination (inc) have been identified, but only to a moderate extent, whereas actually no definite temperature trend could be observed depending on absolute height (hsl). Most likely, this appears to be related to the fact that different levels of altitudes of the sub routes do not mirror an observable height determined trend of temperature distribution.

Discussing sub routes in the warmer transitional period [TP], it becomes apparent that there are large differences between sub routes 1, 2 and sub routes 3, 4. Similar to the entire observation net, temperature distribution of sub routes 1, 2 is essentially controlled by catchment area, too, whereas cold air production and flow within sub routes 3, 4 are predominantly determined by relative heights (hno, hma). The observed positive temperature trend with increasing normalised height (sub routes 3, 4) has different reasons. On the one hand, many local observation points, which show negative anomalies, are located in local depressions, such as the bottom of the tributary Kehrenbach- and Kesselbach Valleys, featuring low housing density with high potential of autochthonous cold air production. On the other hand, the higher parts of sub routes 3, 4 are dominated by west-exposed urban areas with its typical advanced nocturnal heat emission bringing out positive anomalies. For that reason, higher locations face fundamental positive anomalies associated with pronounced cool temperatures in lower regions at the same time.

Another result during the warmer season [TP] is the identification of the warmer mid-slope zones (hma) for sub routes 1 and 3. However, an oppositional trend could be observed. Focusing on sub route 3, west-exposed, residential areas located at mid-slopes and showing positive temperature anomalies might be responsible for an overdrawn indication of the warmer mid-slope zones and therefore also for a significant and strong temperature decrease towards upper and lower parts of the test site. In contrast to that, regarding sub route 1, a positive temperature trend towards upper and lower regions has been diagnosed, tantamount to obvious negative anomalies in medium elevation positions. This can only be explained by the position of numerous observation points near channel lines of rill-like slots alongside the

wide-stretched flanks of the Fulda Valley (cf. Chapter 2). These linear depressions represent zones of cold air production resp. convergence and therefore effect negative temperature anomalies at an altitude, where - considering the terrain of the whole investigated region - usually positive anomalies should have been expected. So the typical warmer mid-slope zone has only been observed in parts due to the overbalance of effective local cold air production within the east-exposed and agriculturally used area. This apparent paradox unfolds the dependence of the expressiveness of empirical data on the sampling design of the observation net.

With regard to the first sampling section after sunset [TP], normalised distance of aspect (eno) has been identified twice by the regression analysis, but only ranking third in the order of detection (sub routes 1, 3). The almost total (statistical) disregard of normalised aspect (eno) and height below crest line (hbc) resp. height above culmination (hac) confirms the assumption that it is necessary to consider the extended topology, which is a substantial controlling factor on a large-scale cold air production and flow (THAMM 2000, VOGT 2001). The above-mentioned parameters only quantify the relative position and therefore are not able to reproduce and to explain the principle of cause and effect of micro-scale temperature distribution.

Terrain analysis of the colder season [CS]: The colder season is characterised by an explicit higher correlation. As expected, once again sub route 2 exhibits the strongest correlation (R^2 [CS]: 0.57-0.75, mean [CS]: 0.68), followed by sub route 1 (R^2 [CS]: 0.56-0.59, mean [CS]: 0.58). The nearly absence of vegetation cover effects an intense reduction of thermal differences specific to land cover with a simultaneous change in surface roughness (SPRONKEN-SMITH 2002, STREIFENEDER *et al.* 2002). By this means, on the one hand differences in thermal behaviour between several kinds of land cover are reduced, and on the other hand, the impact of the terrain on air flow induced by gravitation increases due to the decline of aboveground flow resistance. Both effects obviously lead to a stronger influence of the terrain on local thermal circulation patterns. Nevertheless, the same trend of a lower correlation with respect to the entire route can be observed during the colder season (R^2 [CS]: 0,16-0,39, mean [CS]: 0,27), at least for the available data of sub routes 1 and 2 together (R^2 [CS]: 0,56-0,75, mean [CS]: 0,58-0,68).

A further seasonal difference in thermal behaviour of local climate becomes evident, looking at the controlling terrain parameters. Although catchment areas play cross-seasonally the most important role, absolute height (hsl) has cumulatively been established by regression analysis, what is in opposition to the warmer period. In this regard, a predominant positive temperature variation with increasing absolute height has been discovered, what is a consequence of persistent and intense radiation inversions above the large cold air pool at the bottom of the Fulda Valley.

Further findings: In order to clarify the decreasing influence of terrain with increasing ground level (GEIGER 1961), an overview can be seen in Table 3 showing the lower correlation between other investigated temperature parameters and terrain attributes.

Tab. 3: Correlation between temperature anomalies, gradients and digital terrain information.

All temperature parameters				
Route	Sample Section & "Season"	Ano. 200 cm a.gr. R^2 (corrected)	Ano. 50 cm a.gr. R^2 (corrected)	Gradients (50-200 cm) R^2 (corrected)
ENTIRE ROUTE	Trans. Period [TP]			
	after sun set	0.36	0.19	0.28
	second section	0.40	0.25	0.19
	third section	0.43	0.28	0.16
	before sun rise	0.29	0.20	0.19
	(mean)	**0.37**	**0.23**	**0.21**
	Colder Season [CS]			
	after sun set	0.39	0.25	0.07
	second section	0.16	0.16	0.05
	third section	-	-	-
	before sun rise	-	-	-
	(mean)	**0.27**	**0.21**	**0.06**

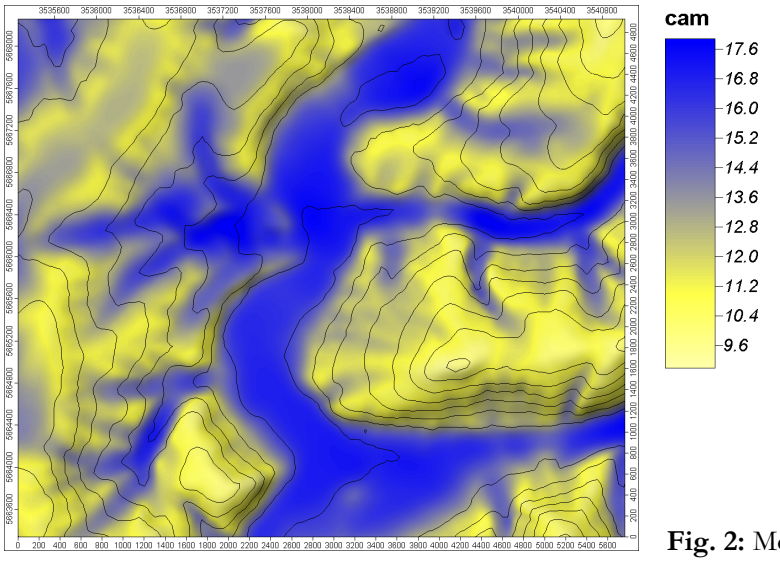


Fig. 2: Modified catchment area (cam)

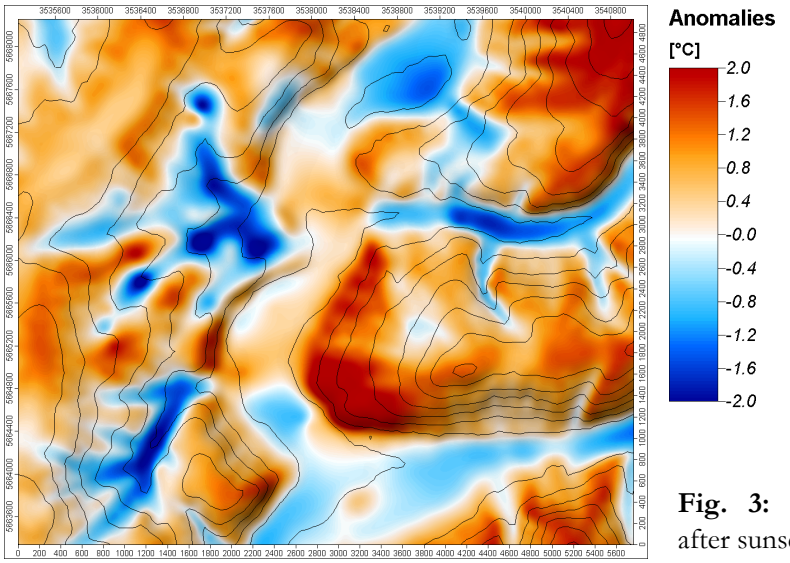


Fig. 3: Estimated Temperature distribution after sunset [TP]

5 CONCLUSIONS

Current terrain analysis techniques offer advanced opportunities for investigations of autochthonous local climate. Especially in a quite hilly topography, where cold air production and flow are related to the extended topology, morphometric terrain parameters are stringently required for comprising and reproducing the procedural complexity of thermal circulation patterns near ground. Complex terrain parameters like catchment areas or relative heights allow to sufficiently explain the topographically determined temperature distribution. In this context, as to be expected, particularly catchment areas play the most important role (cf. THAMM 2000, VOGT 2001). This indicates the imperative demand of considering a model domain, far in excess of the extension of the observation net itself.

Referring to several sub routes, stepwise multiple linear regression analysis on complex terrain parameters yields the highest coefficient of determination at an average of 68%, whereas the aggregation of all sub routes (entire observation net) only shows a lesser correlation at an average of 27-37 %. This is the result of the combination of sub-route-specific thermal land cover and terrain dependent characteristics. Moreover, it could be proved that the percentage of explained variance of temperature distribution basically increases within the colder season. This can be explained by decreased aboveground flow resistance as well as by alleviated thermal differences in land cover resulting from a decline of active

vegetation cover in November and December. Preferentially during the warmer season, unexplained temperature variance has to be ascribed, in the first instance, to the factor of land cover and its consequences for thermal behaviour of the surface in terms of varying heat conduction, storage and emissivity (SPRONKEN-SMITH 2002, STREIFENEDER *et al.* 2002).

Above all, empirical observations of this investigation coincide with findings of OTAVIO & FITZJARRALD (2001), whereupon there is only a quantifiable but not a qualitative seasonal change in nocturnal air temperature patterns. Thus, the location of those sub spaces showing cold air production, flow or accumulation do not alter cross-seasonally but only become less pronounced in winter.

In general, the analysis results reveal that the role of the terrain as an important control for nocturnal temperature distribution, whilst the land cover is most likely responsible for a high percentage of unexplained variance. Additionally, in certain circumstances it becomes apparent that the explanatory power of observed empirical data is strongly dependent on the sampling design, restricted to trafficable roads and path networks.

The presented semi-empirical approach is obviously able to reproduce nocturnal temperature distributions sufficiently neglecting seasonal and phenological variations. Through coupling DTM and land cover information resulting e.g. from satellite imagery, it should be generally possible to precisely predict agricultural risks or to estimate negative human-bioclimatologic influences in dense built-up areas. With regard to high-resolution numerical simulation, the synthesis of terrain parameters and land cover data will be a main target of applied research in future, particularly within the scope of microscale topo- and urban-climatology. The primary aim must be to consistently enhance the accuracy of forecast and of analysis of local climatologic processes in heterogeneous rural and complex anthropogenic environments. But in addition to terrain and land cover, also potential flow obstacles have to be concurrently considered having equal priority in modelling thermal circulation systems. In this context it is essential to adopt statistical downscaling methods in order to integrate small obstacles and hence to be able to develop a precise and computational efficient forecasting tool.

REFERENCES

BAHRENBERG, G., GIESE, E. & J. NIPPER (2003²): Statistische Methoden in der Geographie. Band 2.: Multivariate Statistik. – Borntraeger. Berlin, Stuttgart.

BENDIX, J. (2004): Geländeklimatologie. – Gebrüder Borntraeger. Berlin, Stuttgart.

BÖHNER, J. (2004): Regionalisierung bodenrelevanter Klimaparameter für das Niedersächsische Landesamt für Bodenforschung (NLfB) und die Bundesanstalt für Geowissenschaften und Rohstoffe (BGR). – Arbeitshefte Boden 4: 17-66.

BÖHNER, J. & R. KÖTHE (2003): Bodenregionalisierung und Prozessmodellierung: Instrumente für den Bodenschutz. – Petermanns Geographische Mitteilungen 147: 72-82.

BÖHNER, J. & T. SELIGE (2006): Spatial prediction of soil attributes using terrain analysis and climate regionalisation. – – In: BÖHNER, J. MCCLOY K.R. & J. STROBL (Eds.): Göttinger Geographische Abhandlungen 115: 13-28, 118-120.

BÖHNER, J. & O. ANTONIC (2007): Land-surface parameters specific to topo-climatology. – In: HENGL, T & H.I. REUTER (Eds.): Geomorphometry: concepts, software, applications. – Office for Official Publications of the European Communities, Luxembourg.

BÖHNER, J., GROSS, J. & M.J.P.M. RIKSEN (2004): Impact of land use and climate change on wind erosion: Prediction of wind erosion activity for various land use and climate scenarios using the Weels Wind Erosion Model. – In: GOOSENS, D. & M.J.P.M. RIKSEN (Eds.): Wind Erosion and Dust Dynamics: Observations, Simulations, Modelling. – ESW Publications. Wageningen (Netherlands): 169–192.

DANZEISEN, H. (1983): Experimentelle Untersuchung bodennaher Lufttemperatur- und Feuchteverteilungen in Stadtgebieten mit Hilfe eines Messwagens. – In: Beitr.Landespfl.Rheinland-Pfalz 9: 7-34.

DIETRICH, H. (2006): Kalt- und Frischluftpotentiale in Melsungen (Fulda). – Diplomarbeit am Geogr. Inst. d. Georg-August-Universität Göttingen, Abt. Phys. Geogr. Selbstverlag [unveröffentlicht].

DWD (2005): Klimadaten der DWD-Messstation Nr. 0157919 (Melsungen), 1950-2003.

GEIGER, R. (1961⁴): Das Klima der bodennahen Luftschicht - Ein Lehrbuch der Mikroklimatologie. – Vieweg. Braunschweig.

GEOLOGISCHE KARTE VON HESSEN 1:25.000 (1975²). Kartenblatt 4823 Melsungen. – HESSISCHES LANDESAMT FÜR BODENFORSCHUNG [Hrsg.].

GERSTENGARBE, F.W., WERNER, P.C., BUSOLD, W., RÜGE, U. & K.O. WEGENER (1993): Katalog der Großwetterlagen Europas nach Paul Hess und Helmuth Brezowski, 1881-1992. – In: DWD [HRSG.]. Selbstverlag des DWD. Offenbach.

GROSS, G., FREY, T., MOSIMANN, T., TRUTE, P. & R. LESSING (1996): Die Untersuchung kleinräumiger Kaltluftabflüsse mittels empirischer Abschätzung und numerischer Simulation. – Meteorologische Zeitschrift 5: 76-89.

HARRISON, A.A. (1971): A discussion of the temperatures of inland kent with particular reference to night minima in the lowlands. – Meteorological Magazine 100: 97–111.

HELDT, K. & K. HÖSCHELE (1989): Hang- und Bergwinde am Rheintalrand bei Karlsruhe. – Meteorologische Rundschau 41: 104-110.

KREIKEMEIER, A., DAMM, B., BÖHNER, J. & J. HAGEDORN (2004): Wildbäche im Fulda- und Oberwesereinzugsgebiet (Nordhessen und Südniedersachsen) - Fallbeispiele und Ansätze zur Abfluss- und Abtragsmodellierung. – Zeitschrift für Geomorphologie 135: 69-94.

KUNKA, R. & R. OESTREICHER (22.11.2006): Flächenhafte Regionalisierung von Klimagrößen aus Messfahrtspuren durch Gis-gestützte mathematische Analyseverfahren. – URL: http://www.tlug-jena.de/contentfrs/fach_02/umweltdaten/links/klimafahrten.pdf (aufgerufen am 29.02.2008).

KUTTLER, W. & A.-B. BARLAG (2002): Mehr als städtische Wärmeinseln. Essener Unikate – Berichte aus Forschung und Lehre 19, Vitalität einer Region: 84-97.

KUTTLER, W. & D. DÜTEMEYER (2003): Umweltmeteorologische Untersuchungsmethoden. – Promet 30: 15-27.

OTAVIO, C.A. & D.R. FITZJARRALD (2001): The early evening surface-layer transition, temporal and spatial variability. – Journal of the Atmospheric Sciences 58: 2650-2667.

PÜTTMANN, H. (2002): Bioklimatische Bewertung von Bebauungsstrukturen - am Beispiel zweier Wohngebiete in Halle / S. – Dissertation. Selbstverlag der Universität Halle-Wittenberg.

RAPP, J. & C.D. SCHÖNWIESE (1994): 'Thermische Jahreszeiten' als anschauliche Charakteristik klimatischer Trends. – Meteorologische Zeitschrift 3: 91-94.

ROSNER, H.-J. (1992): Geländeklimatologische Untersuchungen im Raum Feldberg-Altglashütten (Hochschwarzwald) und die Regionalisierung der Ergebnisse mit Methoden der digitalen Bildverarbeitung. Dissertation. – Freiburger Geographische Hefte 39.

SCHWAB, A. (2000): Reliefanalytische Verfahren zur Abschätzung nächtlicher Kaltluftbewegungen. Dissertation. – Freiburger Geographische Hefte 61.

SPRONKEN-SMITH, R.A. (2002): Comparison of summer- and winter-time suburban energy fluxes in Christchurch, Nea Zealand. – International Journal of Climatology 22: 979-992.

STREIFENEDER, M., STEINICKE, W. & J. TIMMER (2002): Klimauntersuchung – Nachbarschaftsverband Heidelberg-Mannheim. Abschlussbericht. – In: STEINICKE & STREIFENEDER UMWELTUNTERSUCHUNGEN GBR (Hrsg.). Selbstverlag. Freiburg.

STÜLPNAGEL, A. VON (1987): Klimatische Veränderungen in Ballungsgebieten unter besonderer Berücksichtigung der Ausgleichswirkung von Grünflächen, dargestellt am Beispiel von West-Berlin. – Dissertation. Selbstverlag der TU Berlin.

THAMM, H.P. (2000): Modellierung von Flächendatensätzen der Oberflächentemperatur, Strahlungsbilanz und Kaltluftproduktion in Strahlungsnächten. – Dissertation. Selbstverlag der Universität Freiburg.

UPMANIS, H. & D. CHEN (1999): Influence of geographical factors and meteorological variables on nocturnal urban-park temperature differences – a case study of summer 1995 Göteborg, Sweden. – Climate Research 13: 125-139.

VOGT, J. (2001): Lokale Kaltluftabflüsse. – In: MEURER, M., BURGER, D. & VOGT, J. [Hrsg]: Karlsruher Schriften zur Geographie und Geoökologie 14. Institut für Geographie und Geoökologie der Universität Karlsruhe (TH). Karlsruhe.

WILSON, J.P. & J.C. GALLANT (2000): Terrain analysis: principles and applications. – Wiley. New York.

NUTRIENT IN- AND OUTPUT FLUXES OF AN AGROFORESTED SUB-CATCHMENT OF THE HANA AT THE BORDER OF TAÏ NATIONAL PARK, CÔTE D'IVOIRE

Elke Fischer [1], *Michael Bock* [1] & *Gerhard Gerold* [2]

[1]Institute of Geography – Section Physical Geography – University of Hamburg, Bundesstr. 55, D-20146 Hamburg (Germany)

[2]Institute of Geography – Section Landscape Ecology – University of Göttingen, Goldschmidtstr. 5, D-37077 Göttingen (Germany)

Abstract: Within a sub-catchment mainly cultivated by cocoa plantations the monthly fluxes of PO_4, NO_3, K, Ca, Mg, Na, Si, and Fe were determined concerning inputs through rainfall and throughfall precipitation and linear surface runoff exports via the Hana stream. Calculations are based on volume / discharge weighted mean concentrations, rainfall and runoff volumes over a total number of 13 months and the total size of the sub-catchment that was delimited und spatially defined with SAGA.

Major component analyses including a variety of parameters such as pH, conductivity, major anions and cations and organic carbon compounds led to the identification of the most relevant parameters. Substance specific analyses of periods with elevated nutrient inputs via precipitation and/or release through linear surface runoff were identified and parameters were grouped according to their behaviour within the system. "Input parameters" meaning parameters that above all are imported to the system via rainfall are phosphate, potassium and sodium. "Ouput parameters" mostly are those derived from mineral weathering and lateral soil transports such as silica and iron. Ca, and Mg, and the nitrogen compounds as well show varying tendencies of high input and output rates depending on seasonal water fluxes.

1 INTRODUCTION

The largest remaining continuous primary forest tract throughout western Africa lies within the Taï National Park (PNT) in South-western Côte d'Ivoire. The Park is surrounded by areas that are mainly cultivated by plantations forming the centre of cocoa production in the country and thus, the world. Besides the direct influence on the forest of the National Park such as poaching and illegal timber exploitation, an indirect influence through hydrological and atmospheric transport originating from the surrounding agricultural areas is probable, which would require an additional 'buffer' zone around the park for complete protection.

The investigation of nutrient fluxes is an essential component for the estimation of nutrient budgets and for the identification of possible sources, sinks and storage capacities of an ecosystem. Furthermore, the interaction of nutrient supply and demand clearly reflects whether the system is closed or balanced and indicates the speed of nutrient turnover. Within this paper, two aspects are addressed: First, monthly nutrient fluxes of rainfall, throughfall and stream water are presented. Secondly, the investigated parameters are classified via cluster analysis and patterns of nutrient input/ouput tendencies are characterized.

2 STUDY SITE

The study site is located within a sub-catchment of the stream Hana covering 37.3 km² at the border of Taï National Park, Côte d'Ivoire (at about 5°50' N / 6°50' W, figure 1). The Hana stream originates in this area and crosses the National Park. The total catchment covers more than 90 % of the the Park area (total Hana catchment area covers 4300 km²).

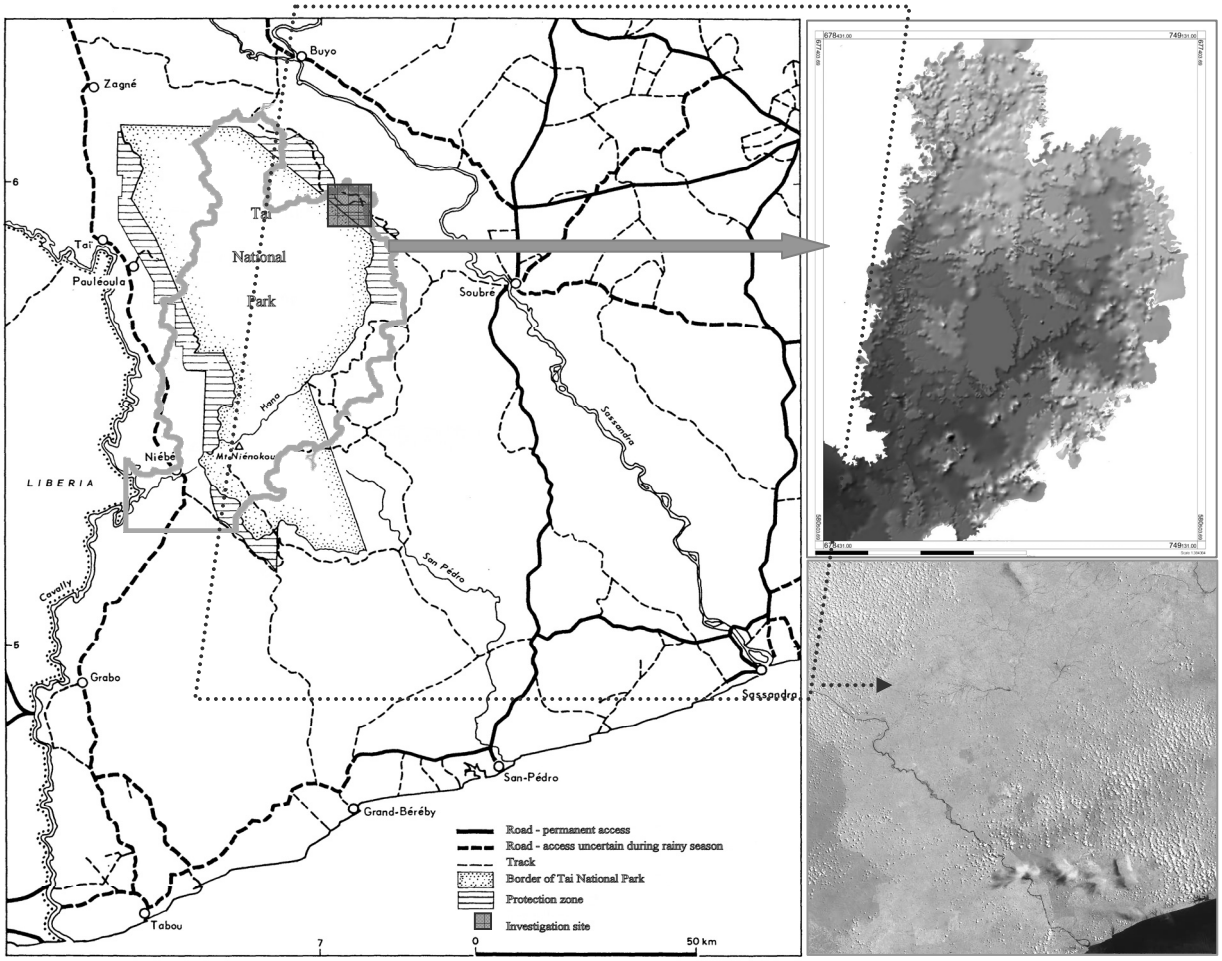


Fig. 1: Taï-National Park and location of the Hana catchment (modified according to RIEZEBOS *et al.* 1994); Digital Elevation Model (50 m distances) and satellite view (2001).

The climate is subequatorial with a mean annual rainfall ranging from 2000 mm in the southwest and 1600 mm at the northeast margin of the Park (VAN ROMPAEY 1993). In general, the rainy season lasts from February to November and the dry season from December to January. During dry season the region is influenced by the Harmattan, a cold and dry wind that is characterized by high dust loads and occurs for about 3 weeks, in general (STOORVOGEL 1993, HERRMANN 1997. MOUNKAILA *et al.* 2003).

The geological setting throughout the catchment is dominated by sericite and green schists of the Sasca domain (PAPON *et al.* 1973; WOZAZEK & KRAWINKEL 2002). The area is a peneplain with inselbergs declining from 225 m asl in the northeast part of the catchment to 50 m asl in the south-west (BLOKHUIS 1994). Soils are predominately ferralitic with deeply weathered Acrisols, Ferralsols and/or Cambisols.

For the investigated sub-catchment (37.3 km²), a ground cover classification was performed by Landsat-7 ETM+ 2001 (overall accuracy 97.6 % compared to overall ground truth). The distribution of cacao or coffee plantations, forest, *hevea* (rubber tree) plantations and annual crops or fallow were 46 %, 22 %, 17 % and 15 %, respectively (SÜLTMANN 2004).

The farmers of 161 surveyed local households stated, that harvest is done twice between September and December and average yields are 0.8 t/ha/y cocoa beans. Fertilizers containing N, P and K (4:1:1) in general are applied March to April (1.2 to 1.8 t/ha/y depending on plantation age). Pesticides and herbicides (most commonly Endosulfan Endosulfan ($C_9H_6Cl_6O_3S$), Basssudine ($C_{12}H_{21}N_2O_3PS$) and Gamma 20 ($C_6H_6Cl_6$), are applied up to two times per year (July/August and December/January).

3 METHODS

The investigation was conducted within a sub-catchment of the Hana stream on 3 plots of 0.1 km² each, representing forest and two cacao plantations cultivated for 7 and 25 years, respectively. The sub-catchment provides representative conditions concerning land use with dominating cocoa plantations and comprises the total area of the catchment before the Hana enters the mainly forested parts of the total catchment. The plots themselves were chosen according to comparable characteristics of topography, slope, soils, distances, climatic conditions and accessibility.

Each investigation plot was equipped with seven bulk throughfall samplers containing of plastic funnels covered with a mesh with 2 mm width to exclude large plant parts or insects (1 m above ground surface, diameter 12 cm, volume 1l). Additionally, two plots were instrumented with rainfall collectors and bulk deposition samplers (1 m above ground surface) to gain open field precipitation. The Hana gauging station was supplied with an automatic sampler (ISCO 6700) and a multiparametric sensor (Sebaterm), continuously recording water level, temperature, pH, electric conductivity and concentration of dissolved oxygen.

The first set of samples was discarded in order to condition the sampling equipment. For the results presented here, samples taken from January 2002 to October 2002 and again between May 2004 and September 2004 were considered since these months had no missing samples or data (besides single missing values due to maintenance/malfunctions and contamination or complete lack of data over two periods of time with a total number of 6 months in 2003 occurred that is due to the beginning of the civil war and the retarded decision of the ongoing financial support due to the same fact). All samples were collected at weekly intervals, although rainfall volume was quantified daily. Samples intended for laboratory analyses were cooled and transported to the laboratory unit in Soubré. All samples were filtered by means of syringe filters (0.45 µm, Millex-HA cellulose mixed ester, Millipore). Sample preservation after collection combined freezing conditioning and the addition of HNO_3 (2 %, Suprapur®), respectively.

Rainfall amounts were recorded between December 2001 and November 2004 on a daily basis at two plots with 15 samplers each. All samples of precipitation, and throughfall were analysed for temperature, electric conductivity, and pH. Data on temperature, pH, electric conductivity, concentration of dissolved oxygen, oxygen saturation and the water level of the stream were measured at 15 minute intervals and retrieved weekly. For discharge calculation the stream velocity was determined weekly with a current meter in a regularly surveyed cross-section of the stream. The data on the cross-section area, water level and velocity were used for discharge calculation (regression water level – discharge, $R^2 = 0.95$).

The photometric analyses of the samples for the concentrations of phosphate, ammonia, nitrate, sulphate, and chlorine were done with a Nova 60 (Merck) at the laboratory unit in Soubré. After airfreight transport to Germany, the conserved samples were analysed for aluminium, calcium, copper, iron, potassium, magnesium, manganese, sodium, phosphorus, sulphur, silicon, and zinc with Inductively Coupled Plasma-Optical Emission Spectroscopy (ICP-OES, Perkin Elmer). Additionally, a sample aliquot was analysed for the concentration of Organic Carbon (total and following 0.45 µm filtration dissolved – TOC, DOC), and Total Nitrogen bound (TNb) with a Dimatoc 100/Dima-N (Dimatec).

Special attendance was drawn on the delimitation of the subcatchment in order to provide reliable nutrient flux estimations of the ecosystem. To produce manually delineated catchments of streams based on Topographic Maps is time consuming and susceptible to uncertainties and errors. In times of GIS technology it is even outdated. On the other hand the uncritical usage of GIS tools to delineate catchments automatically on the base of Digital Terrain models (DTM) and digital stream networks may lead to wrong delimitations, as well since almost no DTM has information about the real drainage paths of streams - no matter what quality it possesses or which data source it is based on. In reality, a stream is a local depression of earth surface and on its way the height decreases more and more. In the model of a DTM, these assumptions are invalid. A DTM has no consistent pathway and the auxiliary way to produce one artificially is to use certain algorithm (there are e.g. different modules in SAGA to handle this problem in the module library "Terrain analysis->preprocessing"). However, these procedures lead to channel

networks and delineated catchments that do not mean "real" streams. If a digital stream network corresponding with the real streams is given, the solution is to adapt the DTM to the digital stream network. This includes not only the manipulation of the grid cells including stream transfers but to adapt the cells in a wider spatial context.

For this study the SRTM DTM (NASA 2008) was processed in the way described above to delineate the catchment area. The SAGA module 'DTM to stream network' (scilands GmbH 2005) that was created to deal on that problem manipulated the neighbourhood around the stream network. The preprocessing procedure was carried out in the following. In a final step the catchment was delineated from that manipulated DTM with the SAGA module 'drainage basins' (scilands GmbH 2002) and the total area was calculated to 37.3 km². The mentioned SAGA modules are not in the open source libraries but special tools of scilands GmbH, Göttingen.

All results were tested for outliers and adjusted. Values below detection limits were set to zero. Normal distribution (Kolmogorov-Smirnov), significances of mean value differences (Mann-Whitney-U) and correlations (Spearman) were tested with the software package SPSS 14.0. 15-minute values and daily concentrations determined for the stream Hana were flow weighted and combined to weekly values. Stream discharge was calculated based on an algorithm derived from stream velocity, water level and the area of the cross-section.

4 RESULTS & DISCUSSION

Table 1 shows the monthly nutrient fluxes for 2002 of rainfall, throughfall of all three sites and nutrient export fluxes through the stream water. Net throughfall fluxes indicate that all substances (except zinc where the differences were not significant) are enriched while passing the canopy layer. This process can either be passive since the canopy combs out dusts, salts and other particles that are washed off the leaves and branches which are represented by substances such as sodium and silica. On the other hand, substances are actively leached out of leave tissue which is especially known for potassium, nitrate and phosphate.

Taking into account the vertical / lateral course of nutrient fluxes and nutrient export via the stream several substances show nutrient retention within the system (on a significant level α=0.01 PO_4, NO_3 and Na) and others are continuously released such as iron and silica. Ca, Mg and K are retained within the biotic circle throughfall – percolation – infiltration – runoff. Concerning potassium, phosphorous and nitrogen compounds it has to be considered that more than 78 % of the sub-catchment are used for agriculture (above all cocoa with 48 % of the total area) that requires a high input of fertilizers adding N, P, and K to the nutrient balance of the ecosystem.

Tab. 1: Mean nutrient fluxes (kg/ha*month) of rainfall and throughfall (tropical rain forest, cocoa plantations, 2002) at the border of Tai National Park, Côte d'Ivoire.

	PO_4	NO_3	Ca	Mg	K	Na	Fe	Si	Zn
Rainfall	0.8	4.9	0.2	0.1	0.3	1.6	0.01	0.1	0.02
Throughfall									
forest	0.8	7.3	1.1	0.5	3.2	2.3	0.02	0.3	0.02
cocoa 7y	1.5	6.9	1.4	0.6	4.9	2.3	0.01	0.4	0.02
cocoa 25 y	1.1	6.7	0.7	0.5	4.3	2.1	0.01	0.3	0.01
Net Throughfall									
forest	0.0	2.4	0.9	0.4	2.9	0.7	0.01	0.2	0.00
cocoa 7y	0.7	2.0	1.2	0.5	4.6	0.7	0.00	0.3	0.00
cocoa 25 y	0.3	1.8	0.5	0.4	4.0	0.5	0.00	0.2	-0.01
Stream	0.1	0.8	0.4	0.2	0.2	0.5	0.20	0.6	0.00

To regroup the parameters according to their tendencies to behave within the vertical and lateral course hierarchical cluster analysis of all components (rainfall, throughfall and stream water) were carried out.

The cluster analysis of the results on open field precipitation (Fig. 2) shows a clear distinction between parameter groups. The marked distance gap indicates that in open field precipitation the grouping of 4 clusters is relevant: Fe as a single and outstanding parameter, phosphorous compounds (total P and PO_4), all nitrogen compounds (TNb, NH_4, NO_3) plus K and as a fourth cluster Mg, Ca, Si, Na plus DOC.

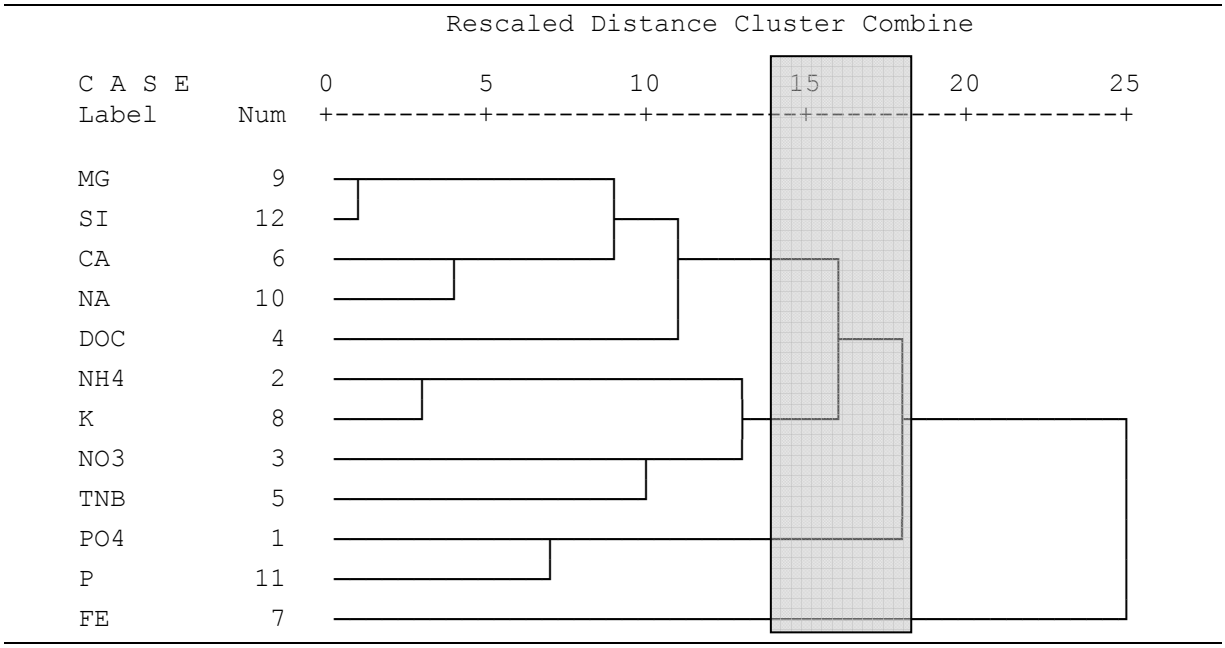


Fig. 2: Hierarchical cluster analysis: volume weighted means of selected parameters of open field precipitation using average linkage between groups (selected distance gap between parameter groups is marked).

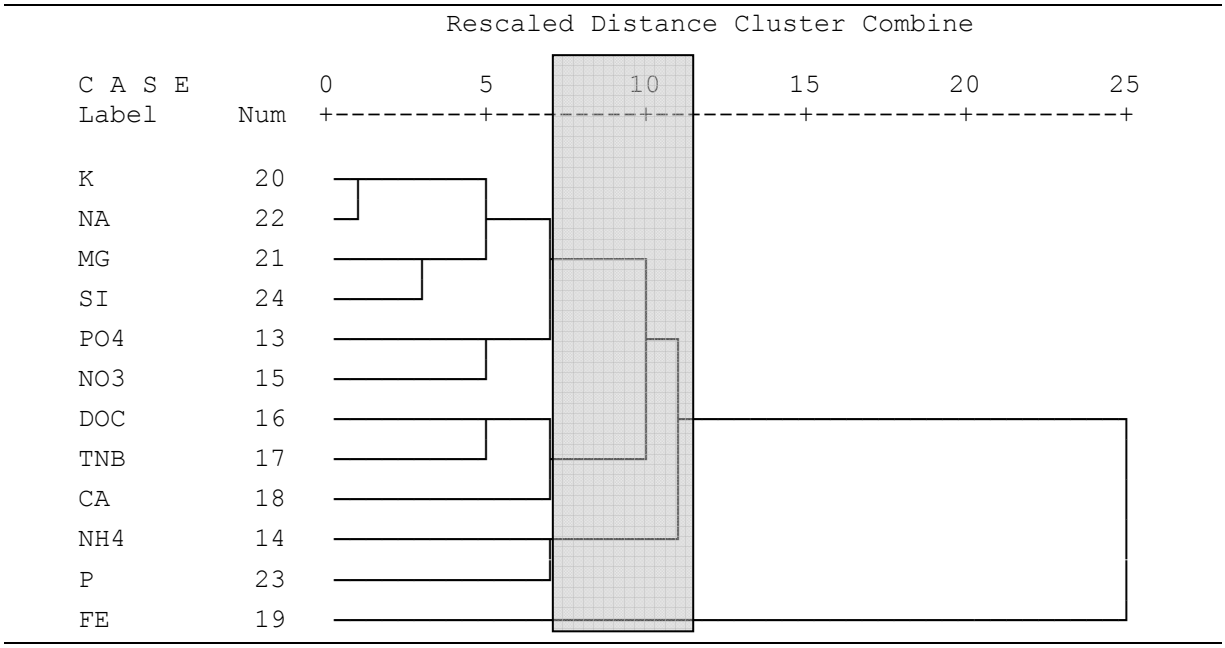


Fig. 3: Hierarchical cluster analysis: volume weighted means of selected parameters of throughfall precipitation using average linkage between groups (selected distance gap between parameter groups is marked).

Cluster classification of the results on throughfall samples again suggest a separation of 4 clusters and again, Fe displays as a single parameter. Compared to open field precipitation parameters the attribution shifts identifying P and NH_4 as a second cluster, DOC, TNb and Ca as the third and NO_3, PO_4, Si, Mg, Na and K as the fourth cluster. Though, it has to be stated that the differences for throughfall precipitation results are not as distinct as for open field precipitation.

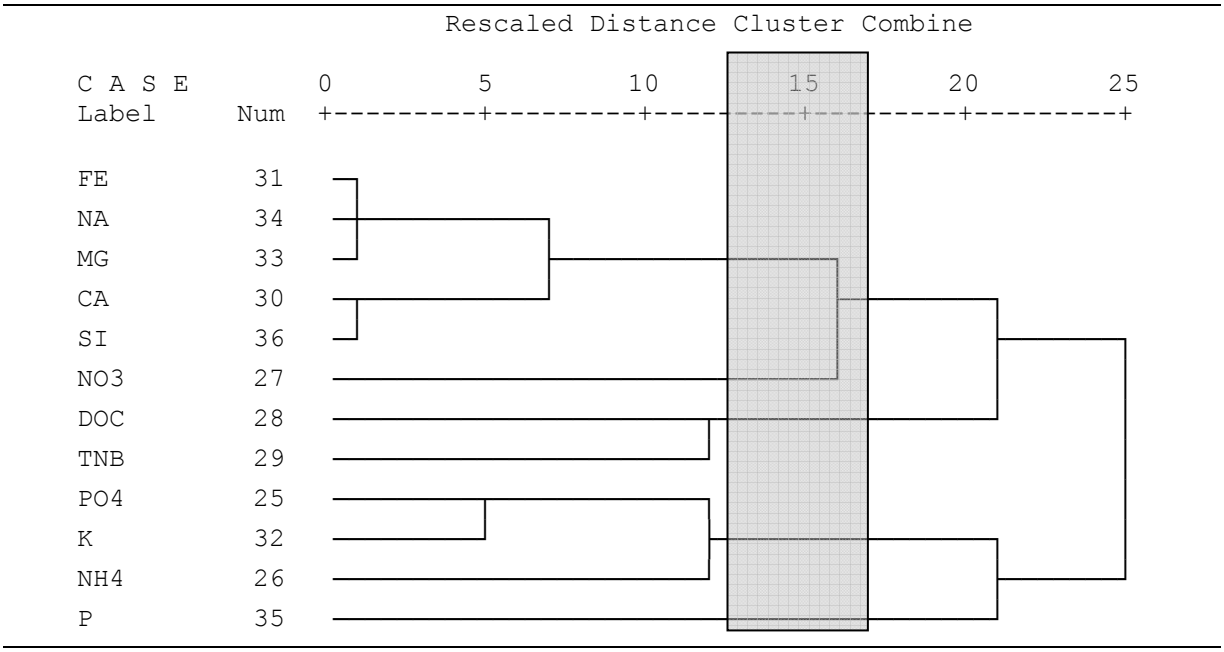


Fig. 4: Hierarchical cluster analysis: discharge weighted means of selected parameters of stream water using average linkage between groups (selected distance gap between parameter groups is marked).

Concerning the results on stream water, a number of 4 clusters is chosen, as well. In contrast to the precipitation components where Fe appeared as a single parameter, in stream water total P is separated from all others. A second cluster is built by PO_4, K and NH_4 while the third cluster comprises DOC and TNb. This set is closely related to NO_3 that belongs to the fourth cluster but would be singled out within a five-cluster solution. The other parameters of cluster four are Si, Ca, Mg, Na and Fe.

By means of cluster analysis the shifting of parameter groupings can be identified. Concerning the open field precipitation a distinct differentiation of phosphorous compounds, nitrogen compounds and the variables mostly provided by atmospheric transport of salts and dusts such as Na, Si and Mg is possible. This classification shifts within the throughfall precipitation samples towards a more biogenic influence e.g. through leaf litter leaching which is expressed by the regrouping of K, PO_4 and NO_3 in one cluster. Nevertheless, the difference gaps between the clusters are not as distinct as for the result of open field precipitation and thus, also might be a result of overlapping factors as nutrient inputs of deposited salts and dusts on leaves that are even enforced in times of the Harmattan and simultaneous active leaching of substances derived from plant tissue itself.

The regrouping of parameters according to the cluster analysis of stream water samples reveals the different influencing components. Cluster 1 and two with the phosphorous compounds and K mostly represent biogenic influences, cluster 3 with TNb and DOC and the adjacency to NO_3 mostly can be attributed to pedogenic influences via e.g. lateral flows. The fourth cluster (Fe, Na, Mg, Ca, Si) comprises the pedogenic / geogenic components resulting from sediment transport within the stream.

On a temporal scale, different periods of rainfall inputs and stream runoff can be distinguished. Figure 5 illustrates the water fluxes of the selected investigation periods.

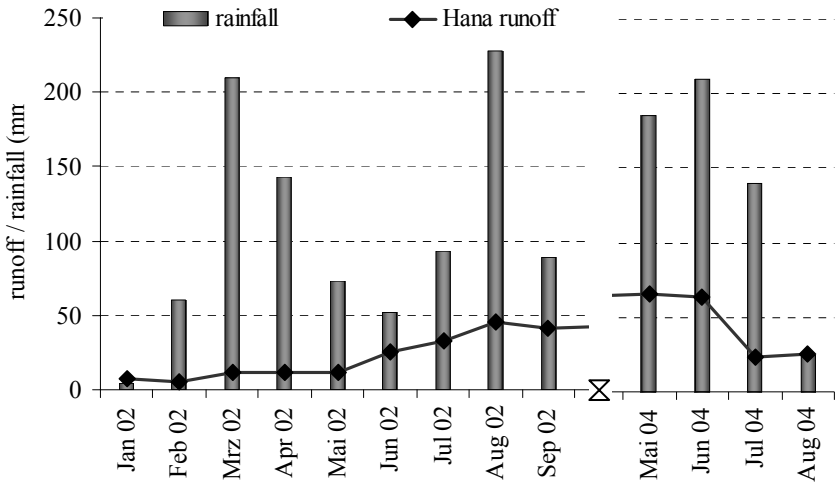


Fig 5: Monthly rainfall/runoff quantities (mm) for selected investigation periods (Hana sub-catchment, Tai National Park, Côte d'Ivoire).

The dry seasons generally lasts from December to January. In 2002 the dry season was prolonged until February which is both reflected in low rainfall and runoff amounts. In the following, rainfall amounts declined from May to July, so the release of water from the sub-catchment through linear surface runoff also fell short. "Normal" higher runoff therefore occurred in august 2002. In 2004 the dry season lasted even longer until March but no reduction of rainfall amount showed up in the following rainy season. Thus, the selected period represents the main rainy season 2004 with high rainfall and runoff quantities.

The monthly nutrient fluxes of rainfall, throughfall and stream water that correspond with rainfall input and stream runoff are shown in Figure 5 and 6. The selected parameters represent those parameters being extracted by factor analysis (major components). The initial set included a variety of parameters such as pH, conductivity, major anions and cations, further elements and organic carbon compounds. The selection of component specific most informative parameters reduced the initial set of 22 to a number of 4 to 6, respectively.

The selected parameters showed different patterns of in- and output tendencies within monthly nutrient fluxes. Substance specific analyses of periods with elevated nutrient inputs through precipitation and/or release via linear surface runoff were identified and parameters were grouped according to their behaviour within the overall system.

"Input parameters" meaning parameters that above all are imported to the system through rainfall are phosphate, potassium and sodium. Except for dry periods when no rainfall and ongoing export of substances by surface runoff occurs input fluxes are much larger than exports. Thus, the system retains these substances even by consuming or fixing. Especially K is provided through leaching within the canopy (and especially within cocoa canopy) whereas sodium above all is an airborne marine substance. The fluxes of sodium in throughfall precipitation dominates the rainfall fluxes only in periods with low rainfall quantities and longer periods of combing out and accumulating salts, dusts, etc. on plant surfaces. In the same time, salts and dusts deposing during rainless periods are not collected in the open field precipitation samplers and thus, do not show up in the results.

"Ouput parameters" mostly are those derived from mineral weathering and lateral soil transports such as silica and iron. Compared to these export rates the aeolian inputs through rainfall and dry deposition are negligible. Ca, and Mg, as well show tendencies of high output rates during wet seasons and high runoff quantities but also are imported by rainfall and throughfall to substantial proportions. The same shows up for the nitrogen compounds (i.e. NO_3). These parameters have periods of inputs exceeding outputs and vice versa. Whether any of these results can help to assess the balance of the system is above all a site specific or even plantation specific decision integrating factors such as physical properties, management practices and fertilizer applications.

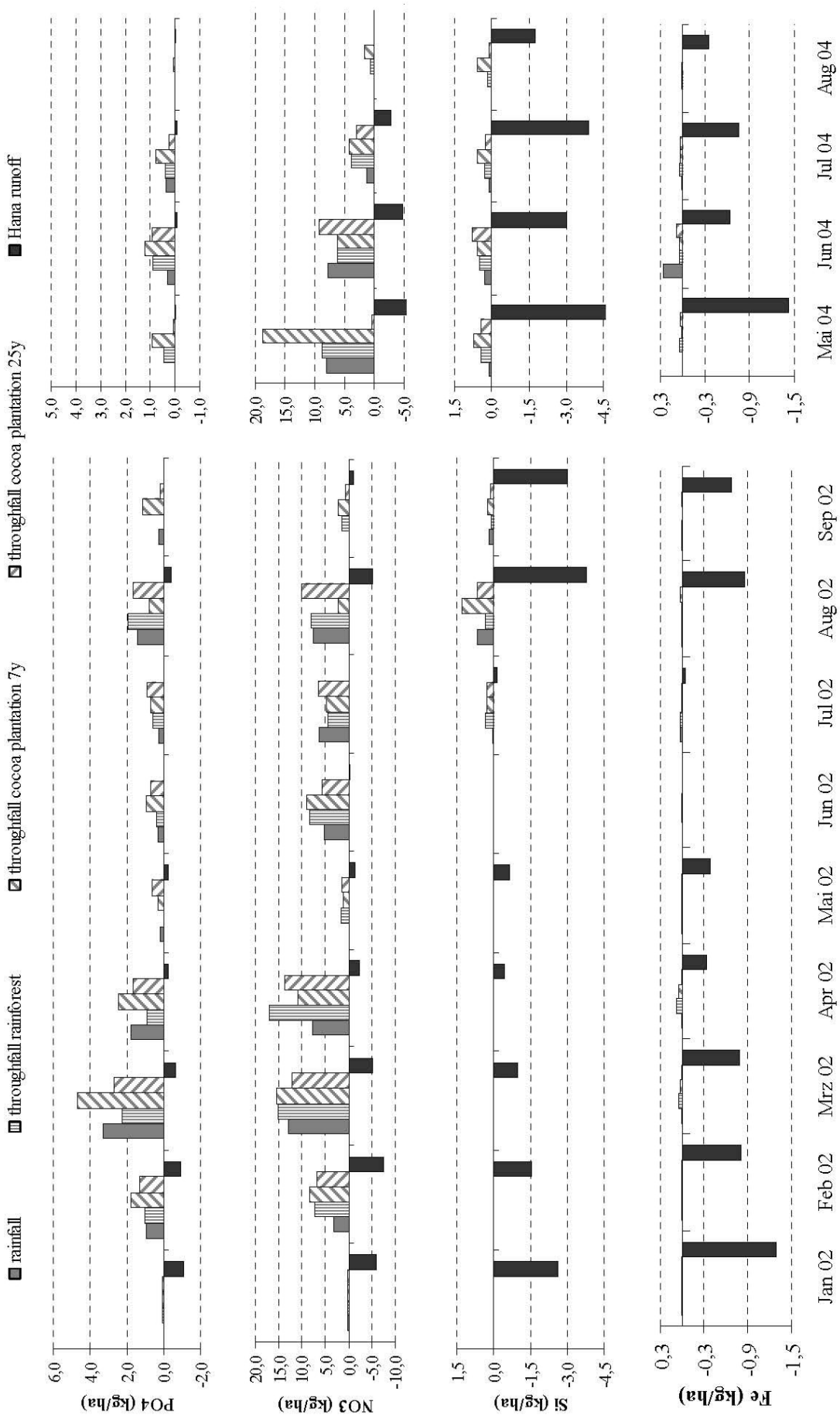


Fig. 6: Monthly nutrient fluxes of PO_4, NO_3, Si and Fe (kg/ha) in open field precipitation, throughfall precipitation (forest, cocoa 7y, cocoa 25y) and stream water for selected investigation periods (Hana sub-catchment, Tai National Park, Côte d'Ivoire).

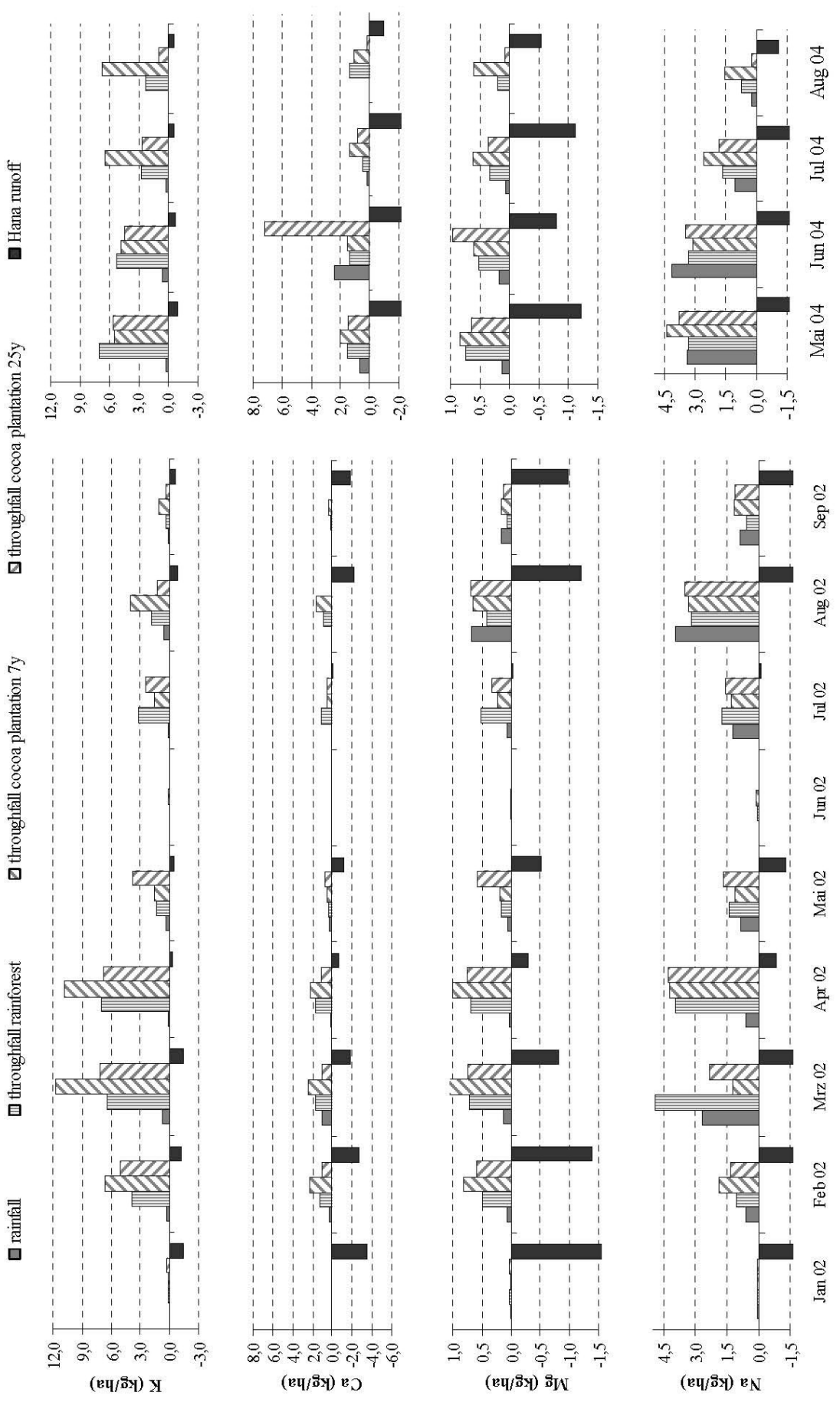


Fig 7: Monthly nutrient fluxes of K, Ca, Mg and Na (kg/ha) in open field precipitation, throughfall precipitation (forest, cocoa 7y, cocoa 25y) and stream water for selected investigation periods (Hana sub-catchment, Tai National Park, Côte d'Ivoire).

Acknowledgments: All presented data and results derive from the Tai region project on hydrology which is an active cooperation between the Centre des Recherches et d'Application en Télédétection (CURAT) at the University of Cocody-Abidjan, Côte d'Ivoire, and the Landscape Ecology Unit (Dept. of Geography) at the University of Göttingen, Germany. The project is financed by the DFG/BMZ programme (German Research Foundation/Federal Ministry for Economic Cooperation and Development) and was initiated in June (2001).

This research was conducted under a research grant and financial support for that we like to thank the German Research Foundation and the German Federal Ministry of Economic Cooperation and Development. Furthermore, we thank P. Voigt, A. Södje, K. Honoré and J. Grotheer for laboratory access and support as well as the people of Azoumanakro and Soubré 3 for their warm hospitality and support in the field.

REFERENCES

BLOKHUIS, W.A. (1994): Le milieu physique. Géologie, Géomorphologie & Pédologie. In: RIEZEBOS, E.P., VOOREN, A.P., GUILLAUMET, J.L. (Eds.): Le Parc National de Tai, Côte d'Ivoire. Tropenbos Series 8 : 28-41, 61-65.

HERRMANN, L. (1997): Staubdepositionen auf Böden Westafrikas – Eine Zusammenfassung. Mitteilungen der Deutschen Bodenkundlichen Gesellschaft 85: 227-253.

MOUNKAILA, M., HERRMANN, L., MAURER, TH., GAISER, TH. & K. STAHR (2003): Spectral and mineralogical properties of potential dust sources on a transect from Sahara to Sahel in Chad. – Mitteilungen der Deutschen Bodenkundlichen Gesellschaft 102: 743-744.

NASA (2008): Jet Propulsion Laboratory. Shuttle Radar Topography Mission. URL: ftp://e0srp01u.ecs.nasa.gov/srtm, (30.05.2008)

PAPON, A. (1973): Géologie et minéralisation du Sud-Ouest de la Côte d'Ivoire; synthése ds travaux de l'opération SASCA 1962-1986. – SODEMI, Abidjan. 285 p.

PNT (2004): Parc National de Tai. URL: http://www.parc-national-de-tai.org/, (22.11.2007).

RIEZEBOS, E.P., VOOREN, A.P., & J.L. GUILLAUMET (Eds.): Le Parc National de Tai, Côte d'Ivoire. – Tropenbos Series 8, Wageningen. 196 p.

STOORVOGEL, J.J. (1993): Gross Inputs and Outputs of Nutrients in Undisturbed Forest, Tai Area, Côte d'Ivoire. – The Tropenbos Foundation III (Tropenbos series 5), Wageningen. 148 p.

SÜLTMANN, C. (2004): GIS- und Satellitenbildgestützte Landnutzungsklassifikation mit Change detection im Westen der Côte d'Ivoire. – In: KAPPAS, M. (Ed.): Erdsicht – Einblicke in Geographische und Geoinformationstechnische Arbeitsweisen 1, Stuttgart. 118 p.

VAN ROMPAEY, R.S.A.R. (1993): Forest gradients in West Africa. A spatial gradient analysis. – PhD thesis, Agricultural University, Wageningen 142 p.

WOZAZEK, S & H. KRAWINKEL (2002): Development of the Côte d'Ivoire Basin: reading provenance, sediment dispersal, and geodynamic implications from heavy minerals. – International Journal of Earth Sciences 91: 906-92.

DIFFUSION, FLOW AND CONCENTRATION GRADIENT SIMULATION WITH SAGA GIS USING CELLULAR AUTOMATA METHODS

Ralph Heinrich [1] *& Olaf Conrad* [2]

[1]Umweltvorhaben Berlin – Brandenburg, Knesebeckstr. 18, D-10623 Berlin (Germany)

[2]Institute of Geography – Section Physical Geography – University of Hamburg, Bundesstr. 55, D-20146 Hamburg (Germany)

Abstract: Cellular automata are simple computational operators, but despite their simplicity, they allow the simulation of highly complex processes. This paper shows how to apply the concept of cellular automata to simulate diffusion and flow processes in shallow water bodies. Similar to real world processes, in the model the diffusion and flow process as well is dependent on the shape of the water body. The proposed model has been implemented as SAGA module, allowing far reaching automation and high performance gain when working with real world data. Two examples for the simulation of water quality are given to point out limits and potentials of the model.

1 INTRODUCTION

This paper presents a method how to simulate diffusion and flow processes within complex 2D shapes and how concentration gradients can be derived from it. The method has been developed to visualize concentration gradients in shallow lakes with in- and outflow, where monitoring data show concentration growth or decrease between the inflow and the outflow points. Parameters are for example nutrients like nitrate, which is reduced by denitrification process inside the water body.

The simulation is based on cellular automata methods. A cellular automaton consists of a regular lattice L, which is in theory endless. The cells z of the lattice L are defined by their position i, k (row, column):

$$L = \begin{pmatrix} z_{1,1} & \cdots & z_{1,k} \\ \vdots & \ddots & \vdots \\ z_{i,1} & \cdots & z_{i,k} \end{pmatrix} \quad [1]$$

Every cell has one state s out of the set of all possible states. Furthermore every cell has a neighbourhood N. For example in a rectangular lattice all cells having a common edge are called von Neumann neighbourhood (the cell itself plus four neighbour cells), cells with common edges or corners are called Moore neighbourhood (eight neighbour cells plus the cell itself, in total nine cells, Fig. 1).

A cellular automaton updates regularly in discrete time steps. Beginning from a start configuration C_0 at $t = 0$ all cell states will be updated simultaneously according a transition function f, which is the same for all cells and all time steps. It refers only to the cell states $s_{i,k}$ of the cell $z_{i,k}$.itself and the neighbourhood $N_{i,k}$ of the respective cell (see Fig. 1).

$$s_{i,k}(t+1) = f(N_{i,k}(t)) \qquad \text{with } N_{i,k} = (z_{i-1,k-1}, z_{i-1,k}, z_{i-1,k+1}, z_{i,k-1}, z_{i,k}, z_{i,k+1}, z_{i+1,k-1}, z_{i+1,k}, z_{i+1,k+1}) \quad [2]$$

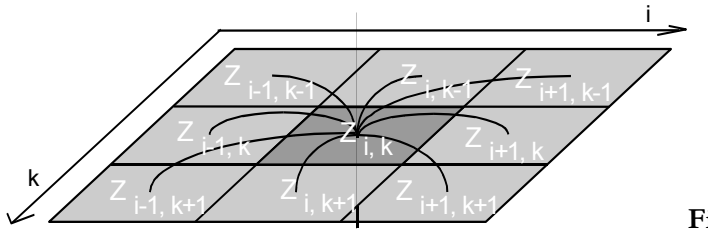


Fig. 1: Moore Neighbourhood.

This procedure can be iterated endlessly until a break occurs (a certain number of iterations or other break conditions). For definitions and application examples see for instance WEIMAR (2003), WOLFRAM (1996), WOLFRAM (2002) or GERHARDT & SCHUSTER (1995). Adopted for GIS, the cellular automata (1) interact with different layers and (2) get spatial limitations with borders, expressed by appropriate border rules.

With these modifications of the cellular automata concept it is possible to simulate spatial phenomena within closed areas. Especially the formulation of boundary conditions, which allows to reshape real world objects, is the innovation in the cellular automata method, that makes it applicable for simulation, where spatial limitations are essential.

2 DIFFUSION SIMULATION

For the simulation of 2d-diffusion processes a quite simple transition rule can be applied. In words this rule says: Create the average value of all cell states of the neighbourhood. Expressed as formula:

[3] $$s_{i,k}(t+1) = \sum_{(x,y)\in N_{i,k}} s_{x,y}(t) / N_{i,k}$$

Applied on a 9 by 9 cells automaton with all cell values s = 0, except the central cell, for which $s_{5,5} = 100$, the transition from $t = 0$ to $t = 1$ is calculated as follows:

$$\begin{aligned} s_{5,5}\,(1) &= [s_{4,4}\,(0) + s_{4,5}\,(0) + s_{4,6}\,(0) + s_{5,4}\,(0) + s_{5,5}\,(0) + s_{5,6}\,(0) + s_{6,4}\,(0) + s_{6,5}\,(0) + s_{6,6}\,(0)\,] \,/\, 9 \\ &= [0 + 0 + 0 + 0 + 100 + 0 + 0 + 0 + 0] \,/\, 9 \\ &= 11{,}11 \end{aligned}$$

After several iterations (time steps) a circular distribution of state values is created (Fig. 2, left). Looking at the value distribution along horizontal and diagonal cuts after 4 iterations, it can be seen that a Gaussian bell shape curve is produced (Fig. 3).

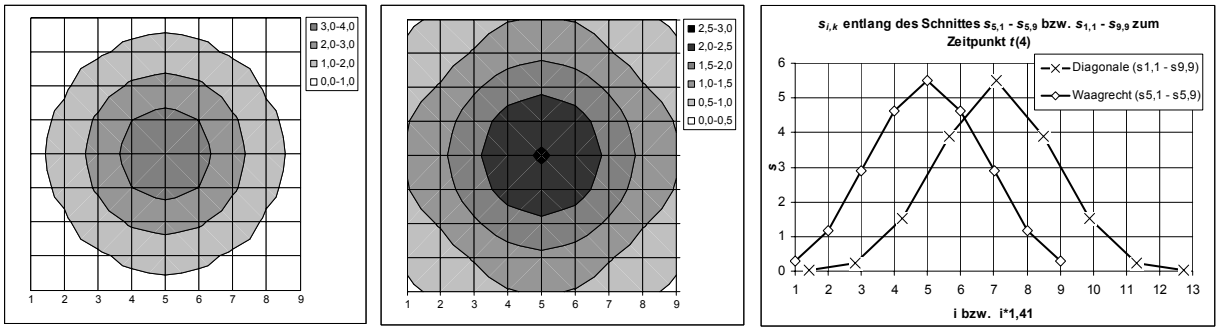


Fig. 2: State value distribution after several iterations.

Fig. 3: Values along horizontal and diagonal cuts.

The total value (sum of all cell states) is maintained, if suitable boundary conditions are chosen (see below). In that case the cellular automaton creates a diffusion function, where the cell value is step by step distributed over the whole area inside the boundaries. With increasing numbers of iterations all cell states converge to an equal distribution.

Boundary Conditions: Different possibilities to formulate boundary conditions have been investigated (HEINRICH 2005, 2006). Suitable for maintaining the total cell values is a boundary condition, where the neighbourhood $N_{i,k}$ of cells at the border of the investigated area inside the cellular automaton is modified. The neighbourhood consists only of those cells, which are inside the investigated area (the white cells in Fig. 4). The grid L itself (as a rectangular grid area) is larger (e.g. the whole grid in Fig. 4). Therefore N does not have the same size for all cells. Only cells, which are not directly at the border, have

a neighbourhood of 9 cells, those at the border have smaller neighbourhoods. Fig. 4 gives values for N for an example area. The smallest possible value for N is 2, as $N = 1$ would mean that the cell is isolated.

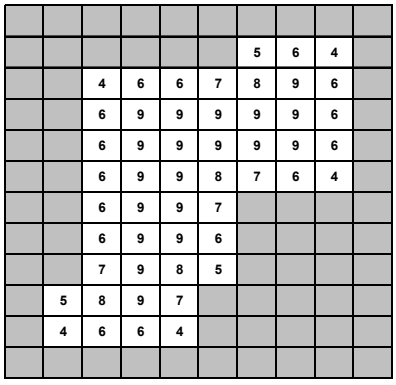


Fig. 4: Neighbourhood $N_{i,k}$ for closed border conditions.

The update after a time step for a cell in a corner (e.g. the lowest cell in the left corner, Fig. 4) will be done based on 4 cells only:

$$s_{2,11}\,(t+1) = [s_{2,10}\,(t) + s_{2,11}\,(t) + s_{3,10}\,(t) + s_{3,11}\,(t)\,] \,/\, 4$$

Depending on the position of the cells along the border the update rule has to be adopted in a similar way. As the cell values (state of the cells) is distributed only amongst the cells inside the investigated area no state value in total is lost. The sum of all cell states is constant:

$$\sum_{i,k=0}^{n,m} s_{i,k}, k(t_0) = \sum_{i,k=0}^{n,m} s_{i,k}(t_\infty) = const. \qquad [4]$$

n, m = maximum for row or column numbers

3 FLOW SIMULATION

Diffusion as described in chapter 2 levels (for example) different concentrations, between areas with high concentrations and those with lower ones. Regarding a cross section square to the diffusion direction a compensating stream is existing there. The stream strength depends on the gradient between the both concentrations and will be zero, when the concentrations are equal.

Using this effect with some modifications a constant flow can be achieved. For that two cells (or group of cells) get fixed state values. If one cell has permanently a higher value than the other, a gradient will be built up between both cells, as the difference between both cells cannot be balanced. With other words a constant flow of value will be generated, the cell with the fixed high value constantly adds value, while the cell with fixed low value takes value off.

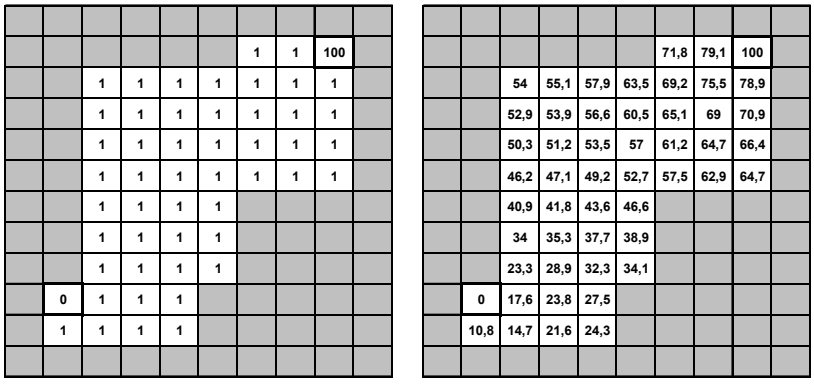


Fig. 5 : Cell states between high and low fixed values, left: at $t = 0$, right: at $t = 600$.

The gradient as the difference between maximum and minimum state of a cell neighbourhood is not constant over all cells but depends on the width of the area, through which the flow is going. Applied to a 2D water body the cell states after numerous iterations represent an elevation surface, on which water flow is directed from high to low elevations. Therefore the fixed cells can be seen as inflow and outflow, respectively. The flow velocity is corresponding to the gradient of each cell's neighbourhood.

						15,6	30,8	49	
		2,17	5,04	9,6	15,2	18,6	34,9	31	
		4,78	7,65	12,3	15,7	18,5	17,7	14,2	
		7,68	10,3	13,4	15,9	16,3	13,4	8	
		10,3	12,6	15,3	17,5	18,1	8,93	3,54	
		13,1	15,2	17,4	19,8				
		18,5	20,4	17,7	14,3				
		35,3	20,1	15,1	11,4				
	46,5	28,9	17,6	12,5					
	17,6	23,8	12,8	5,92					

Fig. 6: Gradient of cell neighbourhoods.

4 CONCENTRATION GRADIENTS

Often concentration differences exist between inflow and outflow of a larger water body. Such as nutrients like nitrate, which are degraded by biological processes (denitrification). This process can be roughly described by a time dependent first order reaction. This means that the reaction velocity depends only on the concentration of the substance itself. Such a process can be described by an exponential decay function:

[5] $$c\,(t) = c_0\, e^{-\lambda t}$$

$c(t)$: concentration at t, c_0: initial concentration at $t = 0$, λ: decay constant

Concentration growth would be described respectively by an exponential growth function. As it is necessary to calculate in discrete steps (from grid cell to grid cell) in fact a geometric progression describes the process of decay (eq. 6). Furthermore the time t has to be replaced by the position x (eq. 7):

[6] $$c\,(t+1) = c\,(t) \cdot (1 + k) \quad \text{or} \quad c\,(t+1) = c\,(t) \,/\, (1 + k)$$

[7] $$c\,(x+1) = c\,(x) \cdot (1 + k) \quad \text{or} \quad c\,(x+1) = c\,(x) \,/\, (1 + k)$$

left: growth, right: decay, k: growth or decay factor analogue to the exponent λt in [5]

Within the factor k the flow velocity v has to be integrated. Base for the estimation of concentration gradients is the consideration that with constant flow velocity along the flow direction the time dependent decay is also location dependent. This relation is reciprocally proportional. On higher flow velocities the spent time t of a water particle needs to flow along a certain distance is shorter than on low velocities. Therefore the time factor t is replaced by the factor $1/v$, as this represents the passed time for a certain distance. In total the growth and decay formula is:

[8] $$c\,(x+1) = c\,(x) \cdot (1 + f/v) \quad \text{or} \quad c\,(x+1) = c\,(x) \,/\, (1 + f/v)$$

left: growth, right: decay, f: correction factor

A correction factor f is necessary to adjust the function to different cell sizes and flow velocity, as the gradient representing the flow velocity (see chapter 3) is only a relative value.

5 IMPLEMENTATION IN GIS

This chapter gives an overview on how the different grids have to be calculated and how they are connected. The result grid needs several intermediate steps from a base grid representing a water body. The following abbreviations are used for this schematically description:

G1 / g1: grid 1 / cell from grid 1
- or { }: empty cell
b / h: width / height of a grid
N : neighbourhood (Moore-neighbourhood of 9 cells)
N (m4): Moore neighbourhood of all cells of grid 4
min / max: minimum / maximum
CI / CO: concentration inflow / concentration outflow

First step is the preparation of base grids, which work as input for the subsequent application of cellular automata, i.e. the water body, in- and outlet definition (Fig. 7). With this information the first cellular automaton is used to interpolate a continuous surface between inlet and outlet (Fig. 8). The gradient can be derived directly from the surface interpolation result (Fig. 9). The gradients for in- and outflow cells must be doubled, because these cells do not have neighbourhood cells with higher values. The inflow (and outflow) cells do in fact have higher (lower) neighbours, which are, according the idea, those cells where they get their inflow from (or give their outflow to). Such cells lie outside of the considered area.

The second cellular automaton is used for decay and growth calculation. The decay calculation tries iteratively to find a concentration surface that reflects given inflow and outflow concentrations (Fig. 10). If the inflow concentration (CI) is larger than the outflow concentration (CO) just the procedure according decay has to be taken. In case CI > CO in addition the procedure for growth calculation given in Fig. 11 has to be considered.

	1	2				b
1	-	-	-	-	-	-
2	-	-	1	1	1	-
	-	1	1	1	1	-
	-	1	1	1	1	-
	-	1	1	1	-	-
h	-	-	-	-	-	-

G1: Cells marked with 1 belong to the water body. G1 can be created easily e.g. using the SAGA module *Shapes to Grid* .

	1	2				b
1	-	-	-	-	-	-
2	-	-	5	6	4	-
	-	5	8	9	6	-
	-	6	9	8	5	-
	-	4	6	5	-	-
h	-	-	-	-	-	-

G2: The cell neighbourhood is derived from G1. The values 2 to 9 indicate the number of water cells inside the Moore neighbourhood N.

	1	2				b
1	-	-	-	-	-	-
2	-	-	1	1	100	-
	-	1	1	1	1	-
	-	1	1	1	1	-
	-	0	1	1	-	-
h	-	-	-	-	-	-

G3: The water body, in- and outlet, where 100 marks the inlet, 0 the outlet, and 1 all other cells belonging to the water body.

Fig. 7: Preparation of Base Grids

	1	2				b
1	-	-	-	-	-	-
2	-	-	xx	xx	100	-
	-	xx	xx	xx	xx	-
	-	xx	xx	xx	xx	-
	-	0	xx	xx	-	-
i	-	-	-	-	-	-

G4: Identical to G3 at t = 0 (initial condition)

	1	2				b
1	-	-	-	-	-	-
2	-	-	xx	xx	xx	-
	-	xx	xx	xx	xx	-
	-	xx	xx	xx	xx	-
	-	xx	xx	xx	-	-
h	-	-	-	-	-	-

G5: calculated from G4. Transition rule:

If g4 = { } then g5 = { }

else g5 = Sum N (g4) / g2

	1	2				b
1	-	-	-	-	-	-
2	-	-	55	66	100	-
	-	45	50	60	66	-
	-	34	40	50	55	-
	-	0	34	45	-	-
h	-	-	-	-	-	-

G6: Output of G4 at t = default value

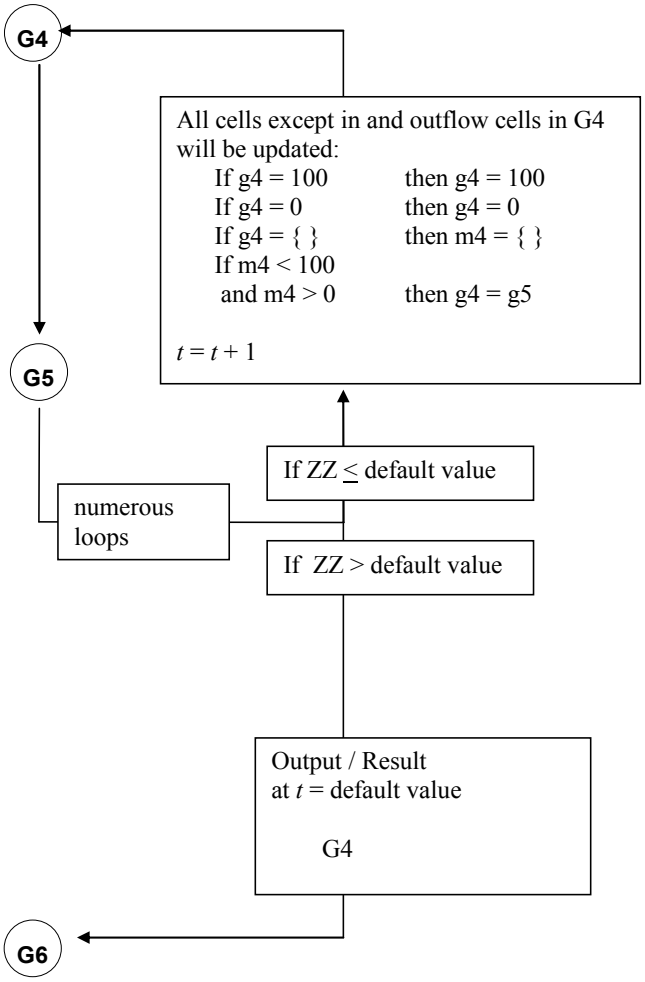


Fig. 8: Cellular Automaton for the Interpolation of a Continuous Surface

	1	2				b
1	-	-	-	-	-	-
2	-	-	22	50	79	-
	-	22	33	60	50	-
		50	60	33	22	
	-	79	50	22	-	-
H	-	-	-	-	-	-

G7: Gradient from the differences to the neighbourhood cells. Transition rule (from G3, G6) is:

If g3 = { } then g7 = { }

g3 = 1 then g7 = max (Ng6) - min (Ng6)

g3 = 100 then g7 = (max (Ng6) - min (Ng6)) * 2

g3 = 0 then g7 = (max (Ng6) - min (Ng6)) * 2

Fig. 9: Gradient calculation

	1	2				b
1	-	-	-	-	-	-
2	-	-	0	0	5	-
	-	0	0	0	0	-
		0	0	0	0	
	-	0	0	0	-	-
h	-	-	-	-	-	-

G8: Initial conditions for decay using G3 and known inflow concentration (e.g. CI: 5 mg/l).

If g3 = { } then g8 = { }
g3 = 100 then g8 = ZK
g3 = 1 then g8 = 0
g3 = 0 then g8 = 0

	1	2				b
1	-	-	-	-	-	-
2	-	-	3,6	4,2	5	-
	-	3	3,6	4,2	4,2	-
		3	3,6	3,6	3,6	
	-	3	3	3	-	-
h	-	-	-	-	-	-

G9: Propagation of concentration using CA method based on G7, G8 using transition rule

If g8 = { } then g9 = { }
g8 = CO then g9 = CO
Else g9 = max (Ng8) / (1 + (f / g7))

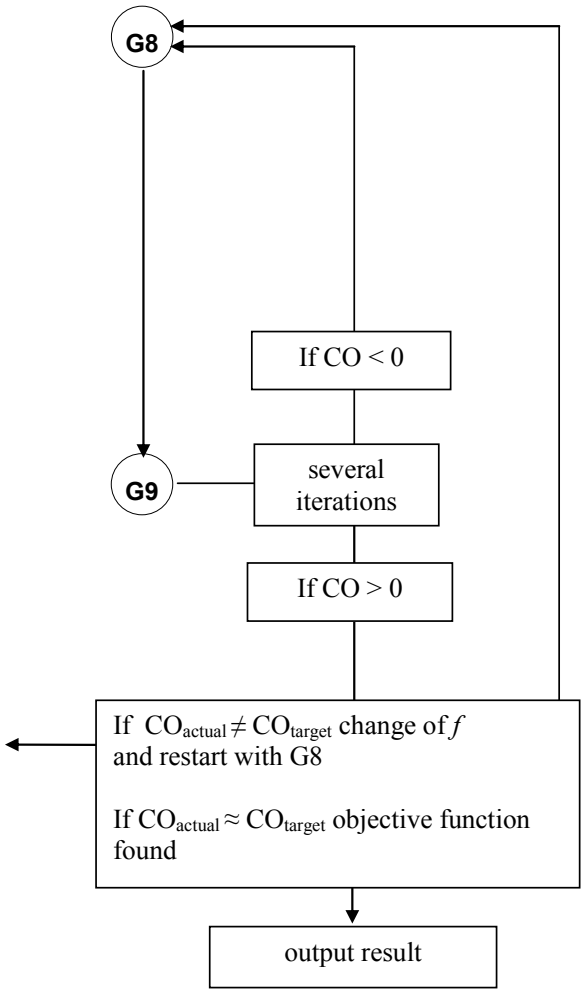


Fig. 10: Decay

If CI < CO

1) CO will be recalculated, so that a decay function can be used (instead of growth),
2) the propagation of the concentration will be calculated as described in Fig. 10, and
3) the results will be calculated back (as if a growth function has been used).

Example values: CI: 5 mg/l, CO: 7 mg/l.

to 1) Target value for CO = 7
identify the relation CO / CI; here: CO / CI = V_{OI} = 7 / 5 = 1,4
next: CO_{new} = CI / V_{OI}; here: CO_{new} = 5 / 1,4 = 3,57...
(calculated target value for an intermediate grid G9n)

to 2) the propagation of the target value CO_{target} will be calculated as described in Fig. 10

	1	2				b
1	-	-	-	-	-	-
2	-	-	0	0	5	-
	-	0	0	0	0	-
		0	0	0	0	
	-	0	0	0	-	-
h	-	-	-	-	-	-

G8

	1	2				b
1						
2			3,99	4,47	5	
		3,57	3,99	4,47	4,47	
		3,57	3,99	4,00	3,99	
		3,57	3,57	3,58		
h						

G9n

to 3) for all cell values $g9_{new}$ the relation g9n to CI is calculated: g9n / CI
next: calculating G9: g9 = CI / (g9n/ CI)

	1	2				b
1						
2			0,80	0,89	1	
		0,71	0,80	0,89	0,89	
		0,71	0,80	0,80	0,80	
		0,71	0,71	0,71		
h						

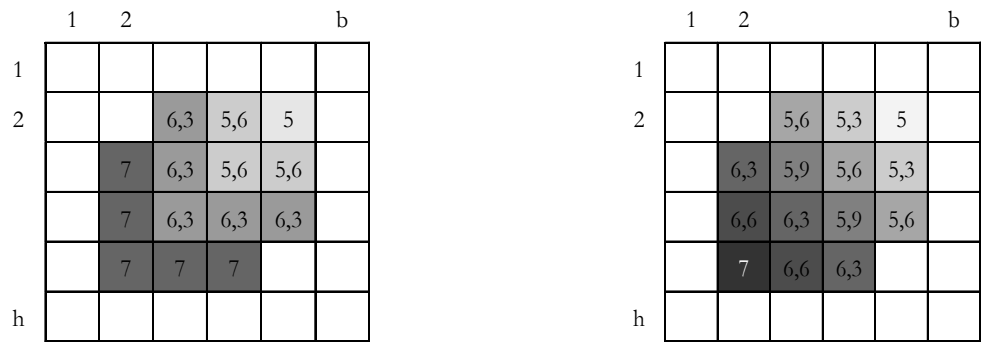
G9n / CI

	1	2				b
1						
2			6,3	5,6	5	
		7	6,3	5,6	5,6	
		7	6,3	6,3	6,3	
		7	7	7		
h						

G9 = CI / (G9 / CI)

Fig. 11: Growth

The resulting grid G9 shows a quite unrealistic angular shaped value distribution. The diagonal direction is accentuated, when using Moore neighbourhoods. Alternatively von Neumann neighbourhoods could be used, but in that case a rhomb like distribution is produced (Fig. 12). Combining both grids G9 (mo) and G9 (vn) a more realistic round gradient can be produced. To gain this G9 has to be calculated twice, one time using the Moore neighbourhood, one time with the von Neumann neighbourhood. Finally both grids have to be merged by obtaining the cell wise average values. The result gives a much more natural impression.



G9 (mo) – Calculated using Moore neighbourhood

G9 (vn) - Calculated using von Neumann neighbourhood

Fig. 12: Optimisation

6 APPLICATION WITH SAGA

The concept of cellular automata is not new to SAGA, since two famous ones, *Life* and *WaTor* (e.g. GERHARDT & SCHUSTER 1995), are included in the standard distribution since its first release in 2004. Both models are rather of academic interest than applicable to real world simulations. But, because of its simplicity, Life gives a very good introductory example for dynamic process simulation in a lecture concept for SAGA module programming (CONRAD 2006). Apart of these two you find the cellular automata concept in a number of other SAGA modules, though it is in most cases not explicitly mentioned as such. The *Vertical Distance to Channel Network* module for instance interpolates channel network base levels using a cellular automaton similar to the surface interpolation proposed here (BÖHNER & KÖTHE 2003, BOCK & KÖTHE 2008).

The original interpolation algorithm had been implemented using the spreadsheet software Excel in combination with the programming environment VisualBasic for Applications. Preparation of input data as well as the visualisation of results has been done with ArcView3.x (HEINRICH 2007b). The whole process chain was obviously time consuming, since spreadsheet software is not optimised for such calculations and many manual interactions had to be done. Therefore motivations for a SAGA module implementtation have been a far reaching automation in one software, which supports a geospatial context, and with an improved performance. This allows experimentations using different boundary conditions and parameter settings. But also weak points of the model become visible, which in turn influence the further model adjustment. Some recent results are given as examples.

Application for Water Quality Interpolation: The first example focuses on the Lake Templin near Berlin, Germany, and simulates the interpolation of total nitrogen concentration, which is an important water quality parameter, indicating eutrophication. The concentration along the lake is decreasing (see Fig. 13; the water flows from north to south). Concentration data is available from two monitoring stations on the inflow and outflow of the lake.

The interpolated surface (Fig. 13a) shows steep slopes at the inflow and the outflow, where the lake is narrow. At wider areas of the lake the slope is much more flat. With this the surface shows behaviour like a water table with flat backwater behind a narrowing and steeper gradient inside narrow segments. The gradient (Fig. 13b) makes this behaviour obviously again and shows in addition that on the one hand narrow sections and changes in flow directions cause steeper gradients (higher flow velocities), and on the other hand that bays with no flow (like the bay in north east) have very flat gradients.

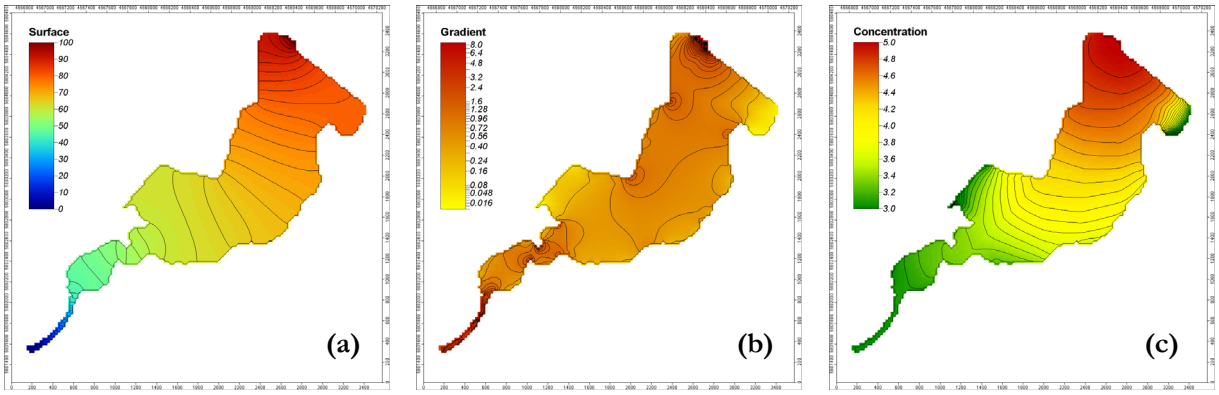


Fig. 13: Water quality interpolation, example southern Templiner See (Germany, southwest of Berlin)

The quality interpolation (Fig. 13c) shows a consistent trend and gives a good impression how the concentration changes under the conditions of a steady decay. In contrast to the surface the concentration changes in narrow segments are slow, whereas the changes in wider areas are faster. This reflects the mechanism that on low flow velocity a water package needs more time to pass a certain distance, so that there is more time for decay processes.

No natural analogy has the very fast concentration reduction in those areas with very slow water movement (like at the north-east bay). Therefore a first model adjustment has been made, which allows the definition of a minimum gradient that cannot be undershot. With this adjustment a diffusion process is simulated. Diffusion is the reason why such steep gradients do not exist in nature.

Application for Diffusion Simulation: A second example has been calculated for the inflow of a point source of polluted water into a shallow coastal bay. The bay has contact to the open sea in the east and a very small outlet to the sea in the south (Fig. 14), which is nearly closed. Two scenarios were calculated, one with an open outlet (Scenario I), the other with a closed outlet (Scenario II).

Assuming an concentration gradient between the inflow in the north west of the bay and the open sea in the east the concentration gradient does show only very little insignificant differences between the both scenarios (Fig. 14b, d).

The left maps show the calculated surface (similar to Fig. 13). In this case this can be seen as the diffusion of the pollutants in the bay. In both scenarios (Fig. 14a, c) it can be seen that the dam and the island inside the bay prevent the diffusion so that the bay is divided into two parts, one with higher concentrations, one with lower. You can also say that the clean sea water does not reach the bay behind the dam, as the exchange process is constrained. The difference between the both scenarios is that the open outlet in the south has an influence on the southern bay, but this influence is limited.

The maps in Fig. 14b and 14d show some significant breaks. The concentration gradient is quite uneven. Further adjustments to the cellular automata interpolation will avoid this kind of unexpected artefacts, which in this case trace back to the influence of the island.

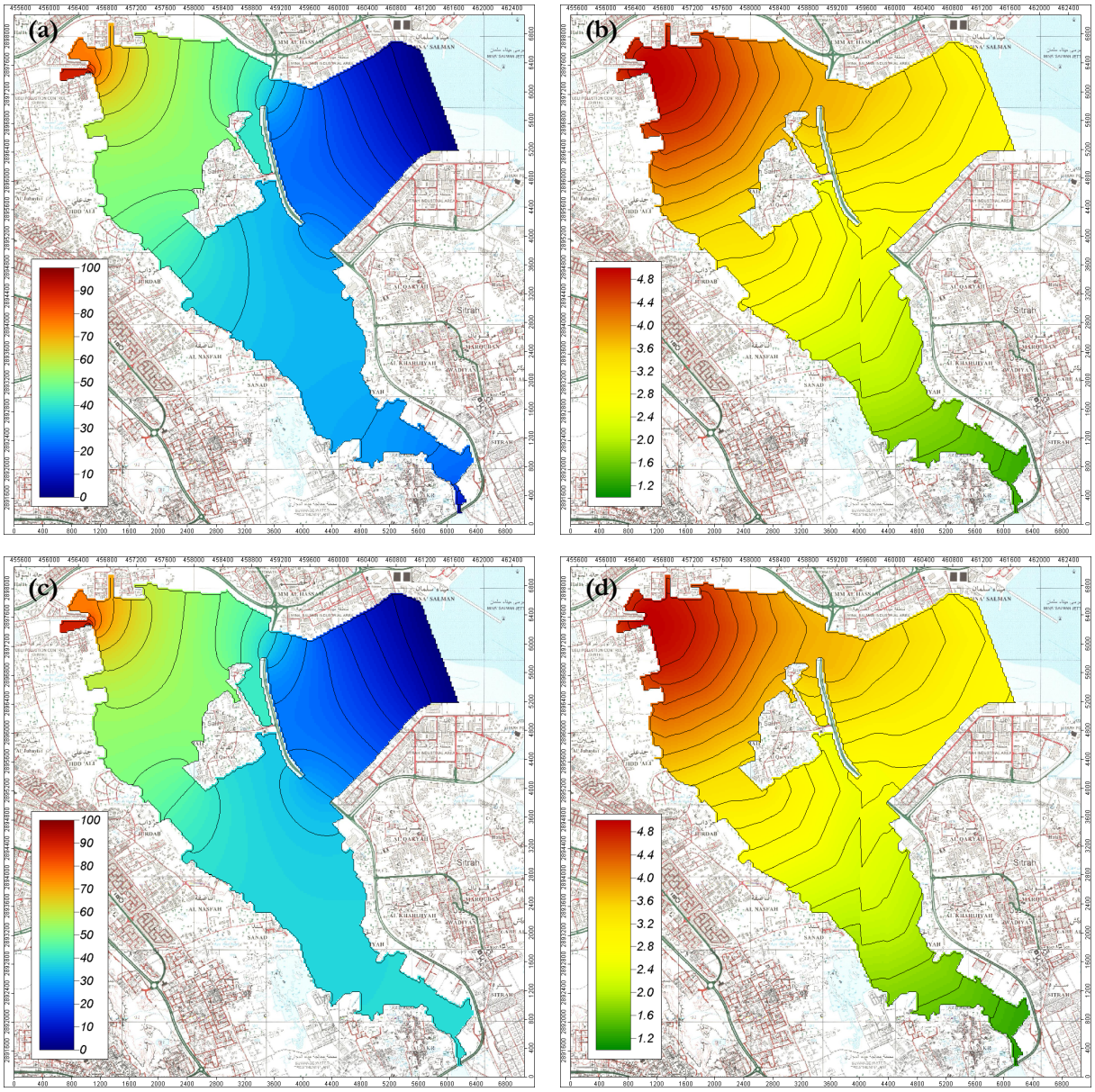


Fig. 14: Two different scenarios for interpolation of a point source inflow into shallow coastal bay, left: diffusion, right time dependant decay.

REFERENCES

BOCK, M. & R. KÖTHE (2008): Predicting the depth of hydromorphic soil characteristics influenced by ground water. – (this issue).

BÖHNER, J. & R. KÖTHE (2003): Bodenregionalisierung und Prozessmodellierung: Instrumente für den Bodenschutz. – Petermanns Geographische Mitteilungen 147 (2003/3): 72-82.

CONRAD, O. (2006): SAGA - Entwurf, Funktionsumfang und Anwendung eines Systems für Automatisierte Geowissenschaftliche Analysen. Dissertation. – [Online] URL: http://webdoc.sub.gwdg.de/diss/2007/conrad/conrad.pdf.

GERHARDT, M. & H. SCHUSTER (1995): Das digitale Universum, Zelluläre Automaten als Modelle der Natur. – 1-320, Vieweg, Braunschweig, Wiesbaden.

GNAUCK, A., HEINRICH, R. & B. LUTHER (2003): Simulation of Eutrophication of Shallow Lakes of the Lower Havel River. – pp. 963-966. In: GNAUCK, A. & R. HEINRICH (Eds.): The Information Society and Enlargement of the European Union. – 17th International Conference Informatics for Environmental Protection, Part 2: Applications, Workshops, Posters, Metropolis-Verlag, Marburg.

HEINRICH, R. (2006): Visualisierung der Gewässergüte in GIS mittels zellulärer Automaten. – In: GNAUCK, A. (Hrsg.): Theorie und Modellierung von Ökosystemen. – Workshop Kölpinsee 2005. Shaker Verlag, Aachen, pp. 55-80, ISBN 3-8322-5477-3.

HEINRICH, R. & A. GNAUCK (2006): Visualization of Surface Water Quality Data - An Interpolation Method using Cellular Automata in GIS. – In: KREMERS, H. & V. TIKUNOV (Eds.): InterCarto-InterGIS 12. International Conference on Geoinformation for Sustainable Development, Berlin, Germany, August 28-30, 2006, Proceedings, Berlin: Deutsche Gesellschaft für Kartographie, 2006, pp. 174-181, ISBN 3-00-019239-5.

HEINRICH, R. (2007a): Interpolation von Gewässergütedaten in GIS mit Zellulären Automaten. – In: GNAUCK, A. (Hrsg.): Theorie und Modellierung von Ökosystemen. – Workshop Kölpinsee 2006. Shaker Verlag, Aachen, pp. 137-158, ISBN 978-3-8322-6058-3.

HEINRICH, R. (2007b): Ein Interpolationsansatz auf Basis räumlich begrenzter Zellulärer Automaten - am Beispiel der Visualisierung von Gewässergütedaten in GIS. – In: STROBL, J., BLASCHKE, T. & GRIESEBNER, G. (Hrsg.): Angewandte Geoinformatik 2007. – Beiträge zum 19. AGIT-Symposium Salzburg. Herbert Wichmann Verlag, Heidelberg, pp. 218 – 227, ISBN 978-3-87907-451-8.

WEIMAR, J.R. (2003): Simulation with Cellular Automata. – 2.ed., 1-200, Logos-Verlag, Berlin.

WOLFRAM, S. (1996): Cellular Automata and Complexity, Collected Papers. 1-596, Addison-Wesley.

WOLFRAM, S. (2002): A New Kind of Science. 1-1196, Wolfram Media, Champaign IL/USA.

POTENTIALS OF A COUPLED GIS-DATABASE APPROACH FOR ENVIRONMENTAL RESOURCES MANAGEMENT

Stephan Liersch [1] *& Raffaele Giordano* [2]

[1]Helmholtz Centre for Environmental Research UFZ, Permoserstr. 15, D-04318 Leipzig-Halle (Germany)
[2]National Research Council – Water Research Institute, Via F. de Blasio 5, 70123 Bari (Italy)

Abstract: Environmental resources management requires adequate tools to deal with various types of data. Due to the different characteristics of environmental data it is necessary to use appropriate tools to administer the data according to their nature. An Advanced Monitoring and Information System (AMIS) is proposed here that is able to integrate data provided by different sources in order to produce targeted information. The AMIS is an open source software tool consisting of three main components. A GIS, acting as user interface, provides functions to visualize and analyze spatial data. An object-relational database management system is used to administer data in various formats. To exchange data between GIS and database an interface was implemented in the GIS. The application of the AMIS is presented on the basis of two different case studies. In the first case study we demonstrate the capabilities of the AMIS in supporting soil salinity assessment and monitoring by integrating local knowledge as alternative source of information in a data-poor region. In the second case study the development of a rainfall-runoff database to support flood risk management is described. Therefore, a rainfall generator and a hydrological model were implemented in the system.

1 INTRODUCTION

Environmental monitoring can be considered as a producer of large amounts of data that need to be stored and organized adequately. Data on the environment are provided by different sources and thus occur in various formats, such as time series data provided by the monitoring network, remote sensing data, model results, reports, spatial information on geographic entities, etc. Appropriate tools are required that convert data into accessible, useful and tailor-made information according to the user's requirements. An important basis of decision making in the context of environmental resources management is information on the current state of the environment as well as information on trends and changes of different environmental variables in order to learn about the effects of implemented management actions. Both can be provided and supported by a well structured environmental monitoring and information system. In the following the system is referred to as AMIS (Advanced Monitoring and Information System) that can be considered as a software tool that provides different methods to deal with spatial and temporal data. Furthermore, the system was extended by specific tools tailored for two different case studies. The focus of the first case study is on supporting the current soil salinity monitoring by integrating local knowledge provided by farmers in Uzbekistan. A user interface was developed that assesses the soil salinity based on qualitative information as well as expert knowledge. In the second case study a rainfall-runoff database was developed to support flood risk management in a German catchment. Here it was necessary to incorporate a rainfall-runoff model in the system.

The AMIS is a combination of a Geographic Information System (GIS) and a relational database management system (RDBMS) that is able to integrate data provided by different sources. For this purpose SAGA GIS was equipped with an interface to exchange data with the PostgreSQL/PostGIS database. In order to be technically sustainable, all used software components are open source and freely available. The AMIS is currently under development in the frame of the European NeWater project (http://newater.info). Section two of this article focus on the technical description of the AMIS components and sections three and four on the application of the AMIS in two different case studies.

2 THE ADVANCED MONITORING AND INFORMATION SYSTEM (AMIS)

An important objective of the AMIS is to make alternative sources of information available, where quantitative data are missing, basically in data-poor regions. We distinguish here between "hard" data, such as measured data provided by a monitoring network, and "soft" data or qualitative information, respectively. Sources of alternative information can be diverse and data formats manifold, for instance information provided by surveys, a questionnaire, expert or local knowledge in form of interviews, reports, pictures, model results etc. (LIERSCH & VOLK 2007). Hence, the system must be able to deal with various data formats and types of information.

Generally, the AMIS is composed of three main components fulfilling different tasks. The core of the system is the GIS SAGA, acting as user interface and providing functions for GIS-specific operations. The second component is a relational database management system used to store and organize available data. The database can be installed locally on a computer or, which is more straightforward, on a web server. The GIS-Database interface can be considered as the third main component enabling the GIS to exchange data with the database (see Fig. 1). The interface is directly implemented in the GIS. Moreover, SAGA was equipped with additional components tailored for the two case studies, such as hydrologic models and an interface to input qualitative information based on local knowledge for soil salinity assessment. The AMIS is currently under development and a prototype version will be implemented to support the soil salinity monitoring in the Khorezm region in Uzbekistan in November 2008. A more detailed description of the three main components is given in the following paragraphs.

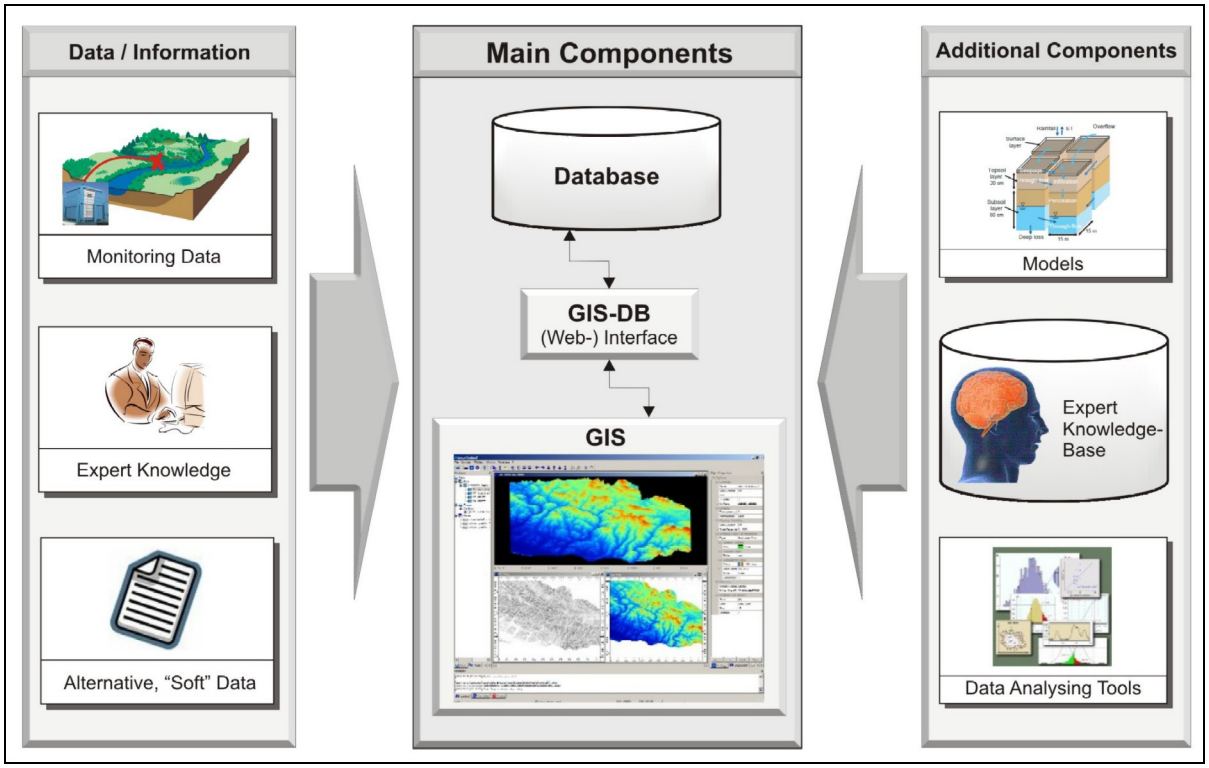


Fig. 1: Structure of the AMIS

GIS: Almost all data and information used in environmental resources management have a spatial reference, which requires the usage of GIS technology. SAGA has been extended and adapted to the specific requirements of the case studies, described in section three and four. Therefore, it was necessary to intervene in the SAGA API and SAGA GUI. The OGC-conform Simple Feature format (OGC 2005a) for geographic vector data was implemented as new data-object type in the SAGA API. Therefore, the new data-object type was derived from the existing type Shapes in order to inherit all implemented

SAGA-Shapefile functions for Simple Features. In order to exchange data with the database an interface was developed and integrated in SAGA. A new menu called Database was implemented in the SAGA GUI (see Fig. 3) providing import / export functions as well as specific AMIS functions that are related to the database.

SAGA has been selected as the basis of the AMIS because it is freely available, open source, platform independent and developed in an object oriented approach. Moreover it is rather user-friendly compared to other GIS with comparable functionalities and has a good performance in operating large datasets.

What is missing in SAGA: An interface, comparable to the plug-in approach, to extend SAGA in a convincing and developer-friendly way; enabling the developer to create new menus for instance. To extend SAGA for specific applications such as the AMIS it is necessary to deeply intervene in the SAGA GUI which is rather time consuming and leads to different SAGA branches which is not in the interest of the SAGA developer team.

Database: The object-relational database management system PostgreSQL (http://www.postgresql.org/) is predominantly used to store environmental data, such as time series from a monitoring network. The extension PostGIS (http://postgis.refractions.net/) enables PostgreSQL to deal with geographical data in vector format implementing the OpenGIS Simple Feature Specification for SQL (OGC 2005b). Relevant geographic data in the AMIS context are for instance, point data (climate stations), line data (channel network), and polygon data (catchment borders). Via look-up tables environmental time series data can spatially linked to their location in the real world and thus, easily accessed by the GIS interface. The database can be installed on a web-server allowing the user to access data via the Internet from any location, providing all users with the same datasets. Sharing data with colleagues is much easier this way, and if time series are updated centrally, everyone has automatically access to the same new datasets. Furthermore, the database provides functions to analyze time series data as well as spatial data.

GIS-DB interface: The interface between GIS and database is implemented in the GIS environment and its functions manifests in the menu Database, which is not available in the common version of SAGA GIS. It provides comprehensive functions to access data in a PostgreSQL database as well as to analyze and query the data using database functions. SAGA can import geodata in the common ESRI Shapefile format and the interface provides functions to export them to the database as OGC conform Simple Features. Any SQL command, depending on the user's permissions, can be executed in the database via the interface and the results can be visualized in the GIS. The strength of SAGA is originally on grid analysis. Using the spatial capabilities of PostGIS, the database interface extends the GIS towards advanced vector functionality.

3 CASE STUDY: SOIL SALINITY ASSESSMENT AND MONITORING

During the Soviet time Uzbekistan was the largest centre of cotton production with intensive agriculture in the desert and semi-desert. In the early 1990s, Uzbekistan still accounted for about 20 percent of world trade and thus was the third largest cotton producer in the world (ERS 2006). The inefficiency of the irrigation network, inadequate drainage systems, and intensive agricultural production were leading to severe soil degradation (salinization). 55% of the land in the Khorezm oblast (the study area) is medium to severe salinized (UNDP 2007). Every year large amounts of (not always available) water are used for irrigation and to leach the salt from the soils to the drainage network, causing water shortage and reducing water quality in downstream areas.

Currently, the regional soil salinity monitoring network is based on soil sampling stations where one station represents an area of approx. 50 ha. Additionally, the Hydromeliorative Expedition (HE), a branch of the Amelioration Expedition – a governmental agency – defines homogeneous areas according to plant growth characteristics at the end of the growing season of cotton plants. Homogeneous areas are neighboring agricultural fields with similar plant growth characteristics, representing similar degrees of soil

salinization. In each homogeneous area a soil sample is taken manually and analyzed in a laboratory. The obtained salinity value is assigned to the whole homogeneous area and the result is a refined map of soil salinity. According to the degree of salinization the amount of water required for leaching in the following season is estimated and plans to allocate the water among users are defined. The weak point in this approach is that the only parameter taken into account to define homogeneous areas – the plant growth – is influenced by a variety of factors, such as seed quality, agricultural management practices, climatic conditions etc., to name only a few. During our fieldwork, several people working in the management of the monitoring system, water managers, and chiefs of the Water User Associations were interviewed about the current monitoring system. According to their opinions, the soil salinity map is reliable only at the regional scale, while at the local scale the information provided are often not correct due to the wrong definition of homogeneous areas. This results in wrong assumptions for water allocation among farmers (GIORDANO *et al.* 2008).

Due to the scarcity of financial resources for soil salinity monitoring in the study area, the objective of our work is to develop an affordable monitoring strategy supported by local knowledge. The aim is not to substitute existing monitoring practices, but to integrate "new" and structured information provided by local farmers to better define homogeneous areas. In order to facilitate a long term engagement of the farmers it is important to integrate the monitoring activities in their normal course of life. In the case study area farmers are organized in Water User Associations (WUA) where each WUA has an agricultural expert who frequently interviews the farmers and suggests management practices. These circumstances facilitate the collection of the required information from farmers for soil salinity assessment and monitoring. During several semi-structured interviews and a workshop with experienced farmers we learned what factors farmers take into consideration when they assess the degree of salinity of their fields in a qualitative way. Based on this we identified the mental models of the farmers with the main factors: *Soil colour*, *Quality of water used for leaching and irrigation*, *Amount of water used for leaching and irrigation*, *Quality of the drainage system* and the *Surface characteristics of the field*. These factors are representing the local knowledge and will be collected by the agronomist using a structured questionnaire. Moreover, the parameters *Groundwater level*, *Groundwater salinity*, *Soil texture*, and *Plant growth* are taken into account to assess the soil salinity at the end of the growing season. These factors are considered as expert knowledge. Figure 2 shows the combination of the mental or cognitive models representing local and expert knowledge.

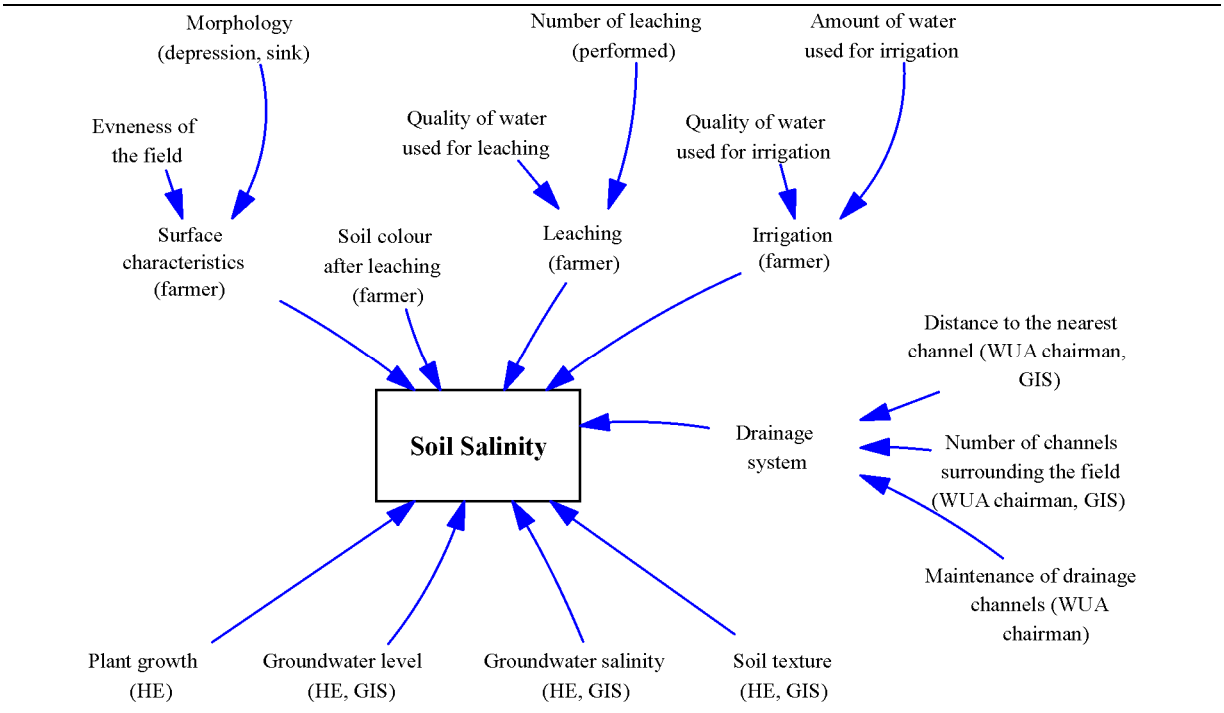


Fig. 2: Combined cognitive models for soil salinity assessment.

It should be emphasized here that soil salinity maps used and produced by the HE are still in analog format. Beside the proposed new approach in soil salinity monitoring and the definition of homogeneous areas we also try to establish the usage of GIS technology. All parameters shown in Figure 2 will be assigned to each cultivated field using GIS. The method is described in the section below. The parameters shown at the top of Figure 2 are collected by the agronomist from farmers during the leaching and vegetation period using a structured questionnaire. Concerning the quality of the drainage system, the parameter *Maintenance of drainage channels* is defined by the WUA chairman and parameters *Number of channels surrounding the field* and *Distance to the nearest channel* will be defined by the WUA chairman or using GIS operations. The parameters at the bottom of the figure are representing the expert knowledge and will be collected by the HE. Parameter *Plant growth* is obtained by visual observations of the fields as described above and will be assigned field by field to the GIS map. The parameters *Groundwater level*, *Groundwater salinity*, and *Soil texture* can be assigned manually field by field or by using overlay functions of the corresponding maps in the GIS.

Soil salinity assessment and data input into the AMIS: The technical basis to realize the soil salinity assessment and monitoring is a digitized map of agricultural fields of the area of investigation. Hence the first step is to load an existing POLYGON map (shapefile) to the SAGA Workspace. In the second step the map must be exported to the database in *Simple Feature* format. Therefore, function "*Create Soil Salinity Base Map*" was implemented in the AMIS (see Fig. 3).

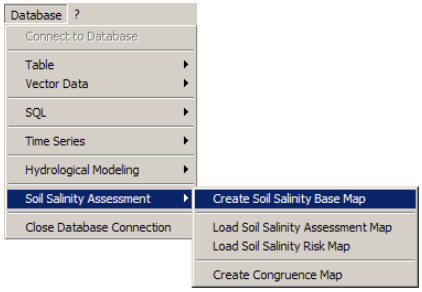


Fig. 3: Menu *Database*.

Specific attributes (listed in Tab. 1) are required by the soil salinity assessment method and will be automatically added to an attribute table in the database. This attribute table is linked to the simple feature by its name, where the prefix "_attr_" is added. To avoid redundancies it is important to split the soil salinity attributes from the geometric attributes, because the soil salinity data are comparable to time series data.

Tab. 1: Required soil salinity attributes.

Attribute name	Description
FID	Identification number of each feature (agricultural field). This will be an integer number and automatically act as the primary key in the database.
YEAR	The year in which the soil salinity assessment has been accomplished.
PERIOD	Two options are available. *Before Leaching* and *After Leaching*.
SALINITY	The salinity value as a word: *low, medium* or *high*.
SALINITY_V	The salinity value as a number between 0 and 1.
MS_LOW	The degree of membership to the fuzzy membership function low. A value between 0 and 1.
MS_MEDIUM	The degree of membership to the fuzzy membership function medium. A value between 0 and 1.
MS_HIGH	The degree of membership to the fuzzy membership function high. A value between 0 and 1.
PLANT	The plant that was growing in the period before the soil salinity assessment.
CONTROL SETTINGS	For each control in the soil salinity assessment dialog the selected value will be stored in a separate column.

Once, the soil salinity base map was exported to the database successfully, the system is ready for data input. Therefore, it is necessary to load the map (simple feature) from the database, because data are not assigned to a shapefile, but to the attribute table in the database. Here the user must use function "*Load Soil Salinity Assessment Map"* from the database menu shown in Figure 3. The map will be added to the SAGA workspace as data type *SimpleFeatures* and SAGA "knows" that this map is a specific map – the map of soil salinity assessment. Only in this case the functions: "*Load Soil Salinity Assessment*" and "*Open Soil Salinity Assessment Dialog*" will be provided in the context menu of the map (see Fig. 4).

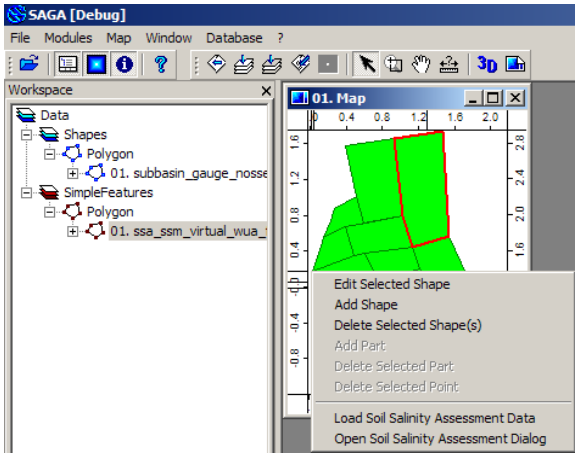


Fig. 4: Context menu of soil salinity assessment map.

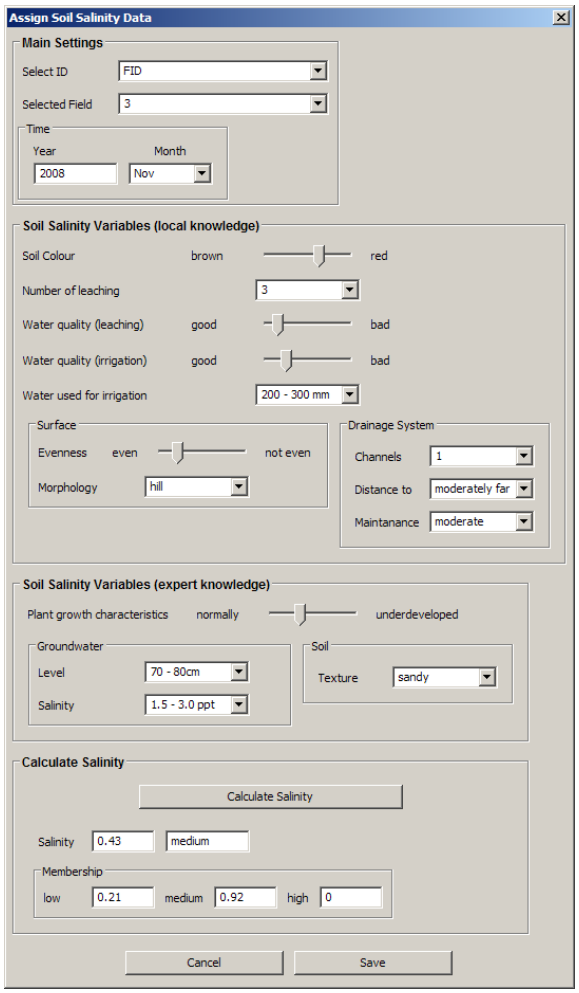


Fig. 5: Soil salinity assessment dialog.

Function "*Open Soil Salinity Assessment Dialog*" shows the data input dialog shown in Figure 5. Here the user can set the values of the controls field by field according to the information gained from the questionnaire and the parameters representing the expert knowledge. The dialog provides controls to input the information according to the parameters characteristics. In order to avoid the usage of numbers we implemented slide bars to input qualitative information, such as the *soil colour*, the *evenness of the field* etc. The terminology used to describe the minimum and maximum values of the slide bars are the terms used by the local community. To input discrete information, such as the *number of leaching performed*, the *morphology of the field*, etc. we use drop-down controls where the user can select the value from the list.

Finally, by pressing the button *Calculate Salinity* the settings of the controls are converted to numbers and the degree of salinity is calculated and displayed in the text field *Salinity*. The salinity value is in the range of 0 to 1 and is calculated using a simple approach based on weighting factors. The weight or importance of each parameter was defined by the local experts in a workshop. The sum of all weights was normalized to 1.

$$\text{Salinity} = \sum_{i=1}^{n} f_i w_i \quad \text{with: } f = \text{factor, } w = \text{weight}$$

Additionally, the degree of membership of the salinity value to the classifications *low*, *medium*, and *high* is shown in the membership fields. The calculations are based on a fuzzy logic approach using the functions shown in Table 2. In the example, illustrated in Figure 5, a soil salinity value of 0.43 is characterized as *medium*. The degree of membership to the fuzzy set *medium* is 0.92 and to the fuzzy set *low* is 0.21 giving an impression in which direction the classification is tending.

Tab. 2: Fuzzy membership functions

Fuzzy set	**Membership function**
low	$y = \cos(\text{Salinity} * \pi)$
medium	$y = \sin((\text{Salinity} + \pi / 4) * \pi^{1.60551})$
high	$y = \cos(\text{Salinity} * \pi + \pi)$

By pressing the *Save* button the date, the salinity values as well as the control values of the selected agricultural field are saved to the attribute table in the database. This procedure must be repeated for all agricultural fields in the map. To visualize soil salinity assessment data function "*Load Soil Salinity Assessment*" from the map's context menu must be used to open the load data dialog. Here the user can select the desired time step. The corresponding data are then loaded from the database attribute table and assigned to the attribute table of the soil salinity assessment map in the SAGA Workspace. Thus, every load command substitutes the existing data in the soil salinity assessment map.

In order to better understand the impact of different management strategies on soil salinization processes it is planned to extend the number of parameters to be monitored by farmers in the next step of the project and to incorporate remote sensing data. In this context it would be useful to investigate the influence of crop rotations and leaching as well as irrigation practices with respect to the salinization in different soil layers – not only the top layer. According to assumptions of several experts, too much water is currently wasted in agriculture for leaching and irrigation – unnecessarily increasing the groundwater level – which negatively feeds back to soil salinization by increased evaporation rates. Hence, our vision is to optimize agricultural water consumption by increasing the knowledge of farmers on causes and effects of different management practices on soil quality and agricultural productivity. Therefore, it is essential that the information gained from the monitoring system feeds back from managers to the practitioners, supporting the learning process.

4 CASE STUDY: RAINFALL-RUNOFF DATABASE FOR FLOOD RISK MANAGEMENT

In this section we present the development of a comprehensive rainfall-runoff database as a tool for flood risk management, representing an expert knowledge-base in the AMIS. For this purpose a rainfall scenario generator was developed generating flood-relevant rainfall events. These events serve as input for the simulation of runoff by using the metric conceptual rainfall-runoff model IHACRES (*Identification of unit Hydrographs and Component flows from Rainfalls, Evaporation and Streamflow data*). The IHACRES model was calibrated to the Mulde catchment in Central Germany that was strongly affected by the devastating flood event in August 2002. IHACRES was chosen because of its parsimonious approach to model parameterization and because of the simplicity to initialize the model. A detailed description of the model is given in JAKEMAN *et al.* (1990) and JAKEMAN & HORNBERGER (1993). The rainfall-runoff scenarios are directly exported to the database. The database can be used as an effective tool to easily assess possible streamflow situations assuming different rainfall volumes for the following days. The benefit of the database approach is that it can be easily used by persons who are not necessarily familiar with hydrologic modelling, because the modelling step has already been accomplished. A comprehensive description of this topic is given in LIERSCH & VOLK (2008).

The rainfall scenario generator is a C++ program that produces random rainfall scenarios according to user-defined rules. We produced rainfall scenarios with the duration of 20 days divided into three periods – a five-day initialization period, representing rainfall patterns of the previous five days, a two-day storm event, representing forecasted rainfall depth and a 13 days recession period characterized by low rainfall amounts. The generated rainfall scenarios capture a variety of artificial (not yet occurred) as well as real rainfall events with a large spectrum of different volumes and intensities. The scenarios are exported directly to the database and serve as input for the IHACRES model, implemented as a module in SAGA GIS. Hence the hydrologic model and the rainfall scenario generator can be considered as extensions of the AMIS. The IHACRES model was equipped with a calibration tool based on the Monte Carlo approach which is appropriate to calibrate six free parameters in a reasonable time.

Before the model can be used to simulate streamflow based on the rainfall scenarios, it must be calibrated adequately. Therefore, the model was calibrated at gauge Golzern with a catchment size of 5400 km². Arithmetic mean time series of all available precipitation and temperature gauges in the catchment served as model inputs. A comfortable tool was implemented in the AMIS allowing the user to load a simple feature of type POINT, representing rain gauges, and to calculate a mean time series based on the selected stations (see Fig. 6). Weighting factors can be assigned to the stations, in order to increase or decrease the influence of a certain gauge.

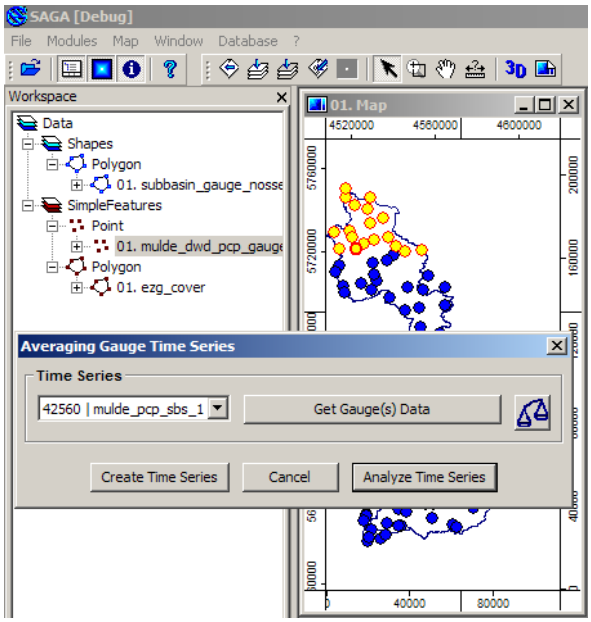


Fig. 6: Averaging time series data of selected gauges.

The performance of the model calibration was measured on a daily time step using three objective functions, the Nash-Sutcliffe efficiency (NSE, NASH & SUTCLIFFE 1970), a modified Nash-Sutcliffe efficiency (HOFFMANN *et al.* 2004) adapted to high flow conditions (ANSE) and the PBIAS in [%]. Reasonable model performance of mean NSE = 0.74 was achieved for the calibration period (1995 to 2002). Noticeable is that the model performs much better in the Mulde catchment in wet years than in dry years. With regard to the objective of our study it was important to analyze the model performance for flood events. During the period 1983 and 2002 three floods in 1983, 1995 and 2002 occurred in the summer periods with daily discharge values exceeding 500 m^3/s. Where the flood event in 1983 was overestimated by the model dramatically, good results were obtained for the events in 1995 and 2002.

With the development of the rainfall-runoff database our objective is to support applied flood risk management providing a large set of 20-day rainfall-runoff simulations. The calibrated model version was used to simulate 20-day streamflow events based on the rainfall scenarios stored in the database. Due to the fact that initial saturation conditions of the catchment significantly influence the streamflow behavior, each runoff scenario was calculated using different combinations of internal model states, namely the parameters: *Antecedent Precipitation Index* and *Streamflow* (t_0). These parameters were used to represent wetness conditions at time step 0. Finally, the number of rainfall-runoff simulations in the database is 2.2 Million, based on 10,000 rainfall scenarios and 220 different initial model states.

Figure 7 shows an example runoff result set, where only runoff simulations, based on rain-fall scenarios that are comparable to the 1995 flood, are displayed. Therefore, a database query in the form of:

SELECT *all simulations*
FROM *runoff simulation table*
WHERE *rainfall volume in the initialization period is* > 45 and < 50
AND *rainfall volume in the storm period is* > 90 and < 100
AND *initial streamflow is* < 30 m^3/s
AND *initial antecedent precipitation index is* > 0.1 and < 0.4

was performed.

The example shows that the runoff simulations capture the 1995 flood event, where some simulations are underestimating and some are overestimating the flood peak. The differences between the lowest simulated flood peak and the highest simulated flood peak can be considered as an uncertainty range.

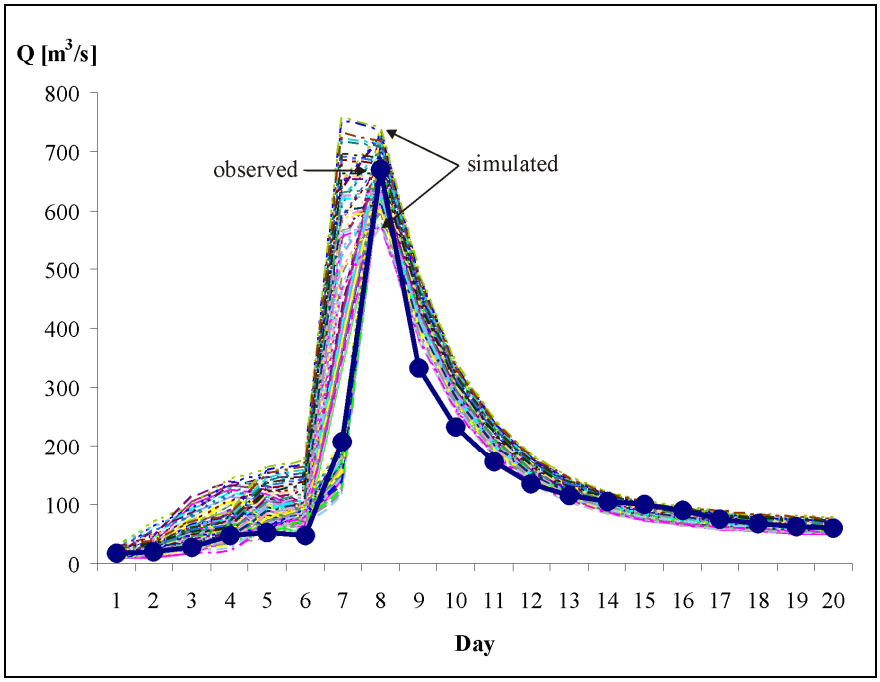


Fig. 7: The 1995 flood event captured by the database simulation result set.

5 CONCLUSIONS

With the development of the Advanced Monitoring and Information System our objective was to support environmental resources management by providing a comprehensive software tool able to integrate different types of data in order to produce tailor-made information. In the first case study we highlighted the necessity of a coupled approach of GIS and database in the context of the integration of local knowledge in soil salinity assessment and monitoring. The functions implemented in the AMIS support an improved definition of homogeneous areas and facilitate a long-term soil salinity monitoring in a technically convincing way. Furthermore, the AMIS can be used to identify trends and changes in soil degradation over the time. The benefits of the coupled approach for environmental modelling, described in the second case study, are manifold. Using a centralized database for the storage of environmental data as basis for modelling *avoids the usage of file-based data management* and if the data are updated centrally *all users have access to the same new data*. A pre-processing tool was integrated in the AMIS enabling the user to produce input data required by the hydrological model by selecting relevant climate stations in a map. Finally, the rainfall-runoff database to support flood risk management was developed using functions provided by the AMIS.

As shown in the two example case studies, a combination of different data management tools provide good potentials to support and facilitate environmental resources management. In order to be prepared for changing information needs demanding new requirements from the system, open source software was used in all AMIS components. This is an important prerequisite for extension and adaptation of the software tool in the future.

REFERENCES

ERS (2006): United States Department of Agriculture. Economic Research Service. – URL: http://www.ers.usda.gov/Briefing/Cotton/trade.htm.

GIORDANO, R., LIERSCH, S., VURRO, M. & V.F. URICCHIO (2008): The integration of Expert and Stakeholder Cognitive Models to support Environmental Monitoring. – In: SÀNCHEZ-MARRÈ *et al.* (Eds.): iEMSs 2008 International Congress on Environmental Modelling and Software. International Environmental Modelling and Software Society (iEMSs).

HOFFMANN, L., IDRISSI, A. E., PFISTER, L., HINGRAY, B., GUEX, F., MUSY, A., HUMBERT, J., DROGUE, G. & T. LEVIANDIER (2004): Development of regionalized hydrological models in an area with short hydrological observation series. – River Research and Applications 20: 243-254.

JAKEMAN, A.J., LITTLEWOOD, I.G. & P.G. WHITEHEAD (1990): Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. – Journal of Hydrology 117: 275.

JAKEMAN, A.J. & G.M. HORNBERGER (1993): How Much Complexity Is Warranted in a Rainfall-Runoff Model? – Water Resources Research 29: 2637-2649.

LIERSCH, S. & M. VOLK (2008): A rainfall-runoff database to support flood risk assessment. – In: SÀNCHEZ-MARRÈ *et al.* (Eds.): iEMSs 2008 International Congress on Environmental Modelling and Software. International Environmental Modelling and Software Society (iEMSs).

LIERSCH, S. & M. VOLK (2007): Towards empirical knowledge as additional information in data-based flood forecasting techniques. – In: OXLEY, L. & D. KULASIRI (Eds.): MODSIM 2007 International Congress on Modelling and Simulation. – Modelling and Simulation Society of Australia and New Zealand, December 2007. ISBN: 978-0-9758400-4-7.

NASH, J. & J. SUTCLIFFE (1970): River flow forecasting through conceptual models, Part 1 - a discussion of principles. – Journal of Hydrology 10: 282-290.

OGC (2005a): OpenGIS Consortium, Inc. OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 1: Common architecture Open Geospatial Consortium Inc.

OGC (2005b): OpenGIS Consortium, Inc. OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 2: SQL option Open Geospatial Consortium Inc.

UNDP (2007): Water, critical resource for Uzbekistan's future. – Publication in support of the Millenium Development Goals. UN Development Program Country Office for Uzbekistan. Tashkent.

MULTITEMPORAL FRAGMENTATION ANALYSIS OF PEAT SWAMP FOREST IN THE KLIAS PENINSULA, SABAH, MALAYSIA USING GIS AND REMOTE SENSING TECHNIQUES

Mui-How Phua [1], *Olaf Conrad* [2], *Kamlisa Uni Kamlun* [1], *Michael Fischer* [3] & *Jürgen Böhner* [2]

[1]School of International Tropical Forestry – Universiti Malaysia Sabah, Locked bag 2073, 88999 Kota Kinabalu, Sabah (Malaysia)

[2]Institute of Geography – Section Physical Geography – University of Hamburg, Bundesstr. 55, D-20146 Hamburg (Germany)

[3]Malaysian-German Forestry Education Project – School of International Tropical Forestry – Universiti Malaysia Sabah, Locked bag 2073, 88999 Kota Kinabalu, Sabah (Malaysia)

Abstract: Fragmentation is a contextual measurement of deforestation that receives increasing attentions. The peat swamp forest in the Klias Peninsula is an important ecological entity especially because it contains the largest pieces of peat swamp forest in the northern end of Borneo Island. We modified and implemented the fragmentation model proposed by RIITTERS *et al* (2000) in the SAGA-GIS. The module provides fragmentation classes of interior, perforated, edge, transitional and patch based on forest covers from supervised classification using maximum likelihood algorithm on multitemporal satellite images of 1985 (Landsat-MSS), 1999 (Landsat7-ETM+) and 2003 (SPOT4-HRVIR). The multitemporal peat swamp forest covers were analyzed for the fragmentation patterns. The peat swamp forest was fragmented in 1985 because of land use activities. Many forest patches were found in the Klias Peninsula but the interior areas were well conserved in two large pieces of peat swamp forest. Between 1985 and 1999, the interior areas have significantly decreased due to the fire in 1998. The main pattern of fragmentation between 1999 and 2003 was removal of small forest patches by land clearing for agriculture.

1 BACKGROUND

Deforestation and subsequently fragmentation are direct threats to biodiversity and sustainability of forest ecosystems (e.g. SAYERS & WHITEMORE 1991, LAURANCE *et al.* 1997). Peat swamp forests are being threatened or destroyed worldwide at an alarming rate (SETO & FRAGKIAS 2007). Deforestation, as the loss of forest, is obvious to see but the associating fragmentation is a different dimension of forest changes that is equally important to deal with. Fragmentation, resulted from deforestation is a contextual measurement that is not represented in the deforestation measure (RIITTERS *et al.* 2000). Fragmentation analysis necessitates analysis of forest pattern at certain spatial scale (window size) and thus giving the mandate to GIS as the suitable domain for characterization of such pattern. A forest fragmentation model proposed by RIITTERS *et al.* (2000) has been used to compare fragmentation pattern between the biomes. Based on the land use trend, tropical rain forest has comparatively higher risk of fragmentation in comparison to other forest types in the world.

Tropical peat swamp forest is found mainly in Southeast Asia (PHILLIPS 1998) and was estimated at about 20 million hectares: two-thirds of the total area of the world's tropical peat swamps (KYUMA 1992). The peat swamp forest ecosystem is well represented in Borneo (PHILLIPS 1998). The Klias Peninsula contained the largest pieces of peat swamp forest at the northern end of Borneo. However, the peat swamp forest has been rapidly disappearing due to land clearing and fires associated with El-Niño/ Southern Oscillation (PHUA *et al.* 2007). PHUA *et al.* (2007) examined the use of several vegetation indices in image differencing technique for detecting burned peat swamp forest. PHUA *et al.* (2008) has further investigated into a fast approach for detecting disturbances in multiple change events. While the loss of peat swamp forest area to fires was studied, fragmentation as a result of forest cover change has yet been

investigated. We aimed at examining the multitemporal patterns of fragmentation of peat swamp forest from 1985 to 2003 in the Klias Peninsula, Sabah, Malaysia. Specifically, we quantify the fragmentation of forest cover, generated from supervised land cover classification. We compare and discuss about the multitemporal patterns of fragmentation of peat swamp forest in the Klias Peninsula.

2 THE KLIAS PENINSULA

The Klias Peninsula is a coastal flat plain, located about 80 km southwest of Kota Kinabalu, the capital of the state of Sabah. The elevation is less than 15 m at the eastern part of the peninsula and gradually lower towards the western direction. The Klias Peninsula receives high annual rainfall, between 2,500 mm to 3,000 mm. Normally, rainfall is highest in October-December (monthly rainfall > 300 mm) and lowest in February-March (monthly rainfall < 200 mm) (IDRIS *et al.* 2005).

With forest debris accumulating over 10,000-12,000 years, peat can support complex peat swamp vegetation, such as six phase-communities in Kalimantan (ANDERSON 1964). Peat swamp forest in the Klias Peninsula was the largest remaining piece in the Northern end of Borneo Island. Peat layer as thick as 14 m was found in the Klias Peninsula. Mixed peat swamp forest is the most significant forest type. The peat swamp forest in the Klias Peninsula was selectively logged until, two protection forest reserves were established in 1984. The Binsuluk Forest Reserve (12,106 ha) and Klias Forest Reserve (3,620 ha) were gazetted for conserving a large portion of the remaining peat swamp forest in Sabah. A protection forest reserve is established for protecting watershed and maintaining the stability of essential climatic and other environmental factors. These areas cannot be logged (SABAH FORESTRY DEPARTMENT 2006a). Both reserves are similar in terms of species composition, i.e. the dominance of *Dryobalonops rappa*, *Dactylocladus stenostachys*, *Shorea platycarpa*, and *Gonystylus bancanus*. Kapur paya (*Dryobalonops rappa*) appears to be one of the most dominant species of the upper canopy. This is because past logging operations were selective, removing only certain commercial species (e.g. *Dactylocladus stenostachys* and *Gonystylus bancanus)*. The kapur paya was either not highly marketable or difficult to transport because the log sinks in water (SABAH FORESTRY DEPARTMENT 2006b).

The Klias Peninsula is typically a heterogeneous and rapidly changing landscape especially due to logging and agricultural activities. Drainage has been constructed to facilitate plantation development (GREER *et al.* 2005). Based on satellite data classification, there were still slightly less than 20,000 ha peat swamp forest in 1997 (PHUA *et al.* 2007). However, fires associated with recent El-Niño events especially in 1998 and 2003 have severely affected this important ecological entity at various extents.

3 SATELLITE DATA AND PREPROCESSING

The main objective of this study is to examine the deforestation and fragmentation of the peat swamp forest in the Klias Peninsula. The earliest satellite image of good quality is a Landsat-MSS image taken on June 29, 1985. Two more satellite images taken on December 07, 1999 (Landsat7-ETM+) and March 26, 2003 (SPOT4-HRVIR), which capture major disturbances were also acquired. The ETM+1999 and SPOT 2003 capture land cover after the extensive El-Niño fires in 1998 and 2003. About 15% cloud are found in the MSS 1985 and SPOT 2003 images The ETM+ 1999 contains about 10% clouds and haze, which are mainly outside the forest reserves. There was a small area within Klias Forest Reserve affected by haze.

The ETM+ 1999 image was rectified to UTM with 21 control points from a 1:50,000-topographic map with a second order polynomial. The georeference residuals in X and Y directions were 3.38 m and 2.37 m, respectively. The SPOT 2003 image was registered to the 1999 image with a second order polynomial. The RMS errors of the registration were 1.92 m to 2.87 m for X and Y directions, respectively. For the MSS 1985 image, rectification using the topographic map was carried out and the RMS errors were below 1 m in both X and Y directions. The resolution of the three sensor data ranges from 20 m for the SPOT 2003, 30 m for the ETM+ 1999 and 79m for the MSS 1985. For better comparability of land cover classification we decided to use 30m as common base resolution. To achieve this nearest neighbour resampling has been applied to all the images.

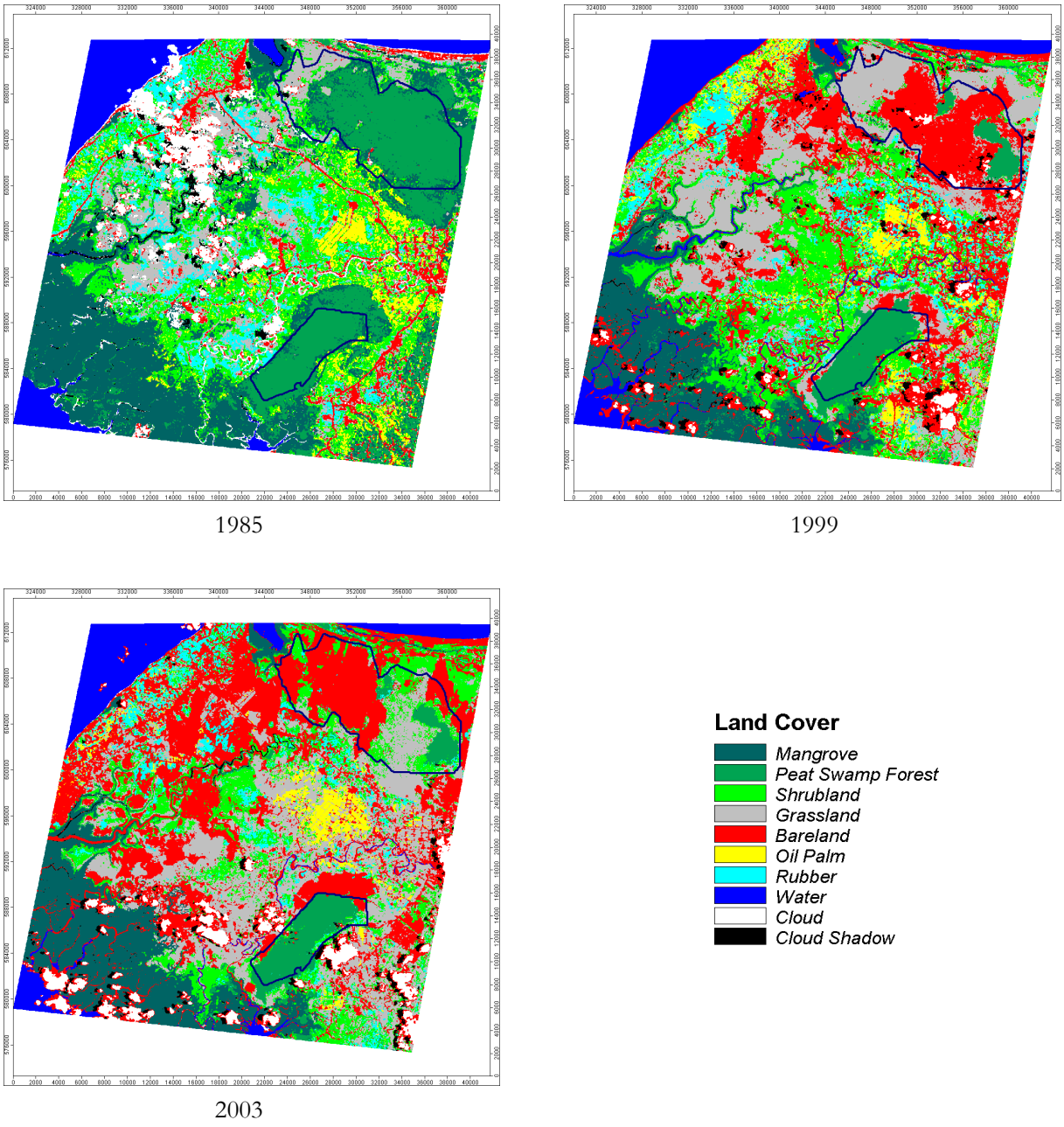


Fig. 1: Land cover classification of the Klias Peninsula. (blue outlines: protection forest reserves).

4 LAND COVER CLASSIFICATION

For examining the peat swamp forest change in terms of deforestation and fragmentation over time, supervised classification strategy was adopted in this study. We defined ten land cover classes, i.e. good peat swamp forest, mangrove, shrub, grassland, bare, oil palm, rubber, water, cloud and shadow. As all the satellite images were acquired in the past, the classification had to depend on ancillary data available. Training area collection was started on the ETM+ 1999 using the topographic maps, which also contain land cover/ land use information updated to 1993. The land cover map (1:100,000) updated in 1996 from Sabah Land and Survey Department was also used as a guide in the selection of training area. The training areas for the MSS 1985 and SPOT 2003 were collected based on the interpretation of the RGB colors of the unchanged classes.

The maximum likelihood algorithm was employed in the classification. Total accuracies of the training area sets for the three images were higher than 90%. The kappa accuracies for the MSS 1985, ETM+ 1999 and SPOT 2003 were 0.85, 0.94 and 0.97, respectively. The MSS 1985 has slightly lower accuracy compared to the ETM+ 1999 and SPOT 2003. This is probably due to the fact that the MSS has only four spectral bands with limited spectral regions (Green to Near Infrared). Two of them are relatively similar Near Infrared bands. With six spectral bands, the ETM+ is more superior in terms of spectral resolution. On the other hand, the SPOT has only 4 spectral bands but the spectral region was from Green to Shortwave Infrared. The focus of this study is the peat swamp forest where the producer's accuracies for the MSS 1985, ETM+ 1999 and SPOT 2003 were 90%, 91% and 96%, respectively. The user's accuracies for the MSS 1985, the ETM+ 1999 and SPOT 2003 were 95%, 87% and 100%, respectively. The accuracy of ETM+ 1999 was probably influenced of haze.

Some postprocessing operations were carried out to improve the classification results. This includes majority filtering (3x3) to reduce the 'salt-and-pepper' effect of the classification results. Misclassification of peat swamp forest as mangrove and vice-versa was relatively obvious in the land cover classification (LCC) of 1985. It was likely because of the relatively poor spectral resolution of the Landsat-MSS. The classification of peat swamp forest based on the ETM+ 1999 was relatively consistent with the SPOT 2003 except some areas in the Klias Forest Reserve affected by haze. We used the peat swamp forest and mangrove in LCC 2003 as the reference to correct the misclassification in the LCC 1999 and 1985. For example, if the pixel is peat swamp forest in 2003, it must be the same land cover in 1999 and 1985. Fig. 1a, b and c show the LCC of 1985, 1999 and 2003, respectively.

Fig. 2 shows the land cover types of the Klias Peninsula for the 18-year period since 1985. The Major changes noticeable between 1985 and 1999 are the significant decreases in peat swamp forest. In 1997, there were about 20,000 ha of good peat swamp forest in the Klias Penisula (PHUA *et al.* 2007). This means the deforestation rate between 1985 and 1997 was very low (about 200 ha/ year). It was the disastrous fires occurred in April 1998 that severely damaged the peat swamp forest in the Klias Peninsula (UNDP/GEF, undated). As a result, the bareland increased three times. It seemed that mangrove also experienced 'deforestation' but it was actually the influence of cloud and its shadow in the ETM+1999 and SPOT 2003.

On the other hand, grassland shows increasing trend in both periods. Because the ETM+ 1999 was acquired in end of 1999, grass and fern species that capitalized on the temporary nutrients released by the fires have colonized some burned areas. Field observation half year after the 2003 fire found that plant colonization dominated by grass and fern species occurred very fast. (PHUA *et al.* in press). This explains the increase of the grassland area in 1999 as well as in 2003. Half of the previously burned areas were burned again in 2003 (PHUA *et al.* 2008). Adding to land clearing, the bareland became the most dominant land cover in 2003. Fig. 2 also shows that the shrubland decreased steadily from 1998 to 2003.

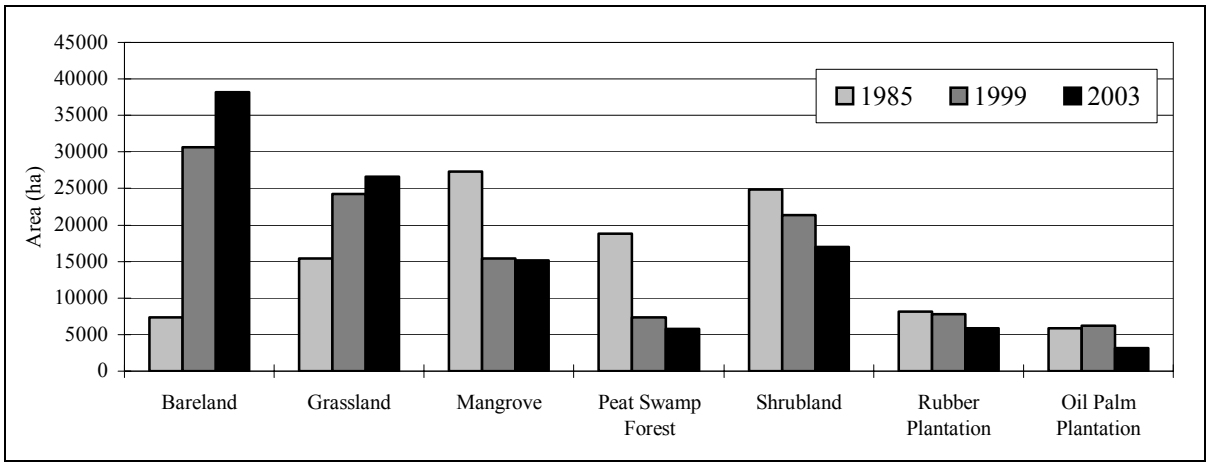


Fig. 2: Land Cover Classes of the Klias Peninsula from 1985-2003.

5 FOREST FRAGMENTATION ANALYSIS

As starting point for forest fragmentation estimations we chose the method introduced by RIITTERS *et al.* (2000). This fragmentation model works on raster based forest classifications and has originally been used to analyse global land cover maps derived from AVHRR satellite imagery with a spatial resolution of about 1km. In the following the model has been applied in several studies to 30m Landsat based land cover classifications to show the state or even the rate of change of forest fragmentation (RIITTERS *et al.* 2002, HURD *et al.* 2002). WADE *et al.* (2003) modified the method to distinguish between natural and anthropogenic causes for forest fragmentation, e.g. by urban growth. Based on the model results HURD *et al.* (2002) derived additional parameters to quantify fragmentation for specified areas like administrative units.

Input information needed by the fragmentation model is whether a pixel is forested or not. With this information two parameters are derived for each pixel, the forest density P_f and the forest connectivity P_{ff}. Both parameters refer to a given neighbourhood around each pixel, originally defined by a squared window with the analysed pixel in the centre. From the pixels within such a so-called moving window the density is calculated as

$$P_f = \frac{number\ of\ forest\ pixels}{number\ of\ all\ non\ water\ pixels}$$

Pixels classified as water are not counted, because the analysis shall usually be restricted to the land surface. The connectivity is likewise calculated as

$$P_{ff} = \frac{number\ of\ pixel\ pairs\ with\ both\ pixels\ forest}{number\ of\ pixel\ pairs\ with\ at\ least\ one\ pixel\ forest}$$

Pixel pairs are defined as those pixels, which are adjacent to each other in cardinal directions. Therefore a 3x3 moving window with a total number of 9 pixels has 12 pixel pairs (Fig.3).

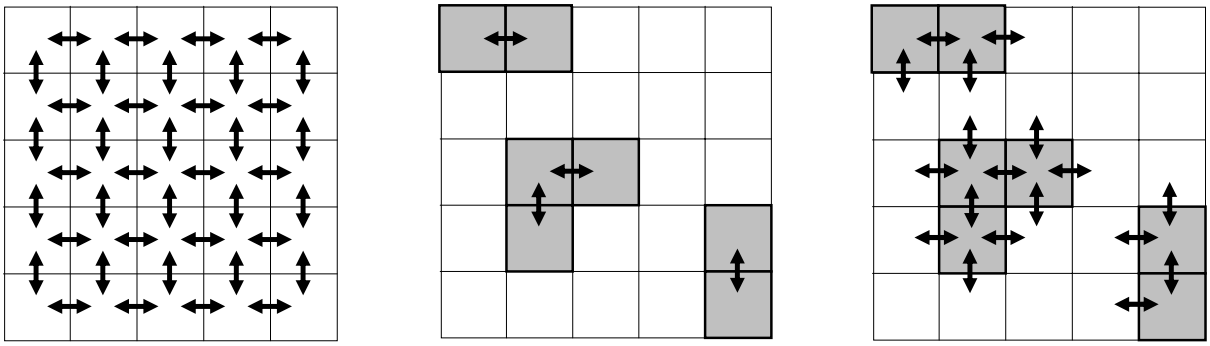


Fig. 3: Connectivity Pff calculation for a 5x5 moving window. a) total number of pixel pairs = 40. b) 4 forest-forest pixel pairs. c) 18 pixel pairs with at least one forest pixel. Pf = 4 / 18 = 0.22.

Finally using density and connectivity each pixel can be assigned to one of 6 forest fragmentation categories. If we express the density and connectivity ratios as percentages by multiplying them with the factor 100, then the categories are defined as followed (after RIITTERS et al. 2000, see Fig.4):

- interior, if density = 100%
- undetermined, if density > 60% and density = connectivity
- perforated, if density > 60% and density > connectivity
- edge, if density > 60% and density < connectivity
- transitional, if density is between 40% and 60%
- patch, if density < 40%

Categories might be defined differently using changed threshold values for determining to which category a pixel belongs. RIITTERS *et al.* (2002) also differentiated between interior forest with a density above 90% and core pixels with a density of 100%, whereas the rare category undetermined has been dropped (Fig.4). Fragmentation categories are first of all a classification of forest density. This is reasonable because the overall relation between density and connectivity is quite narrow (Fig. 5). Connectivity is only used to differentiate pixels with a density higher 60% but which have not yet reached the density of interior forest.

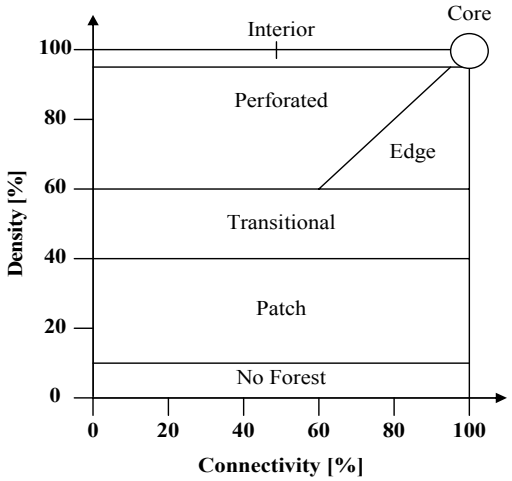


Fig. 4: Forest fragmentation categories (after RIITTERS *et al.* 2002).

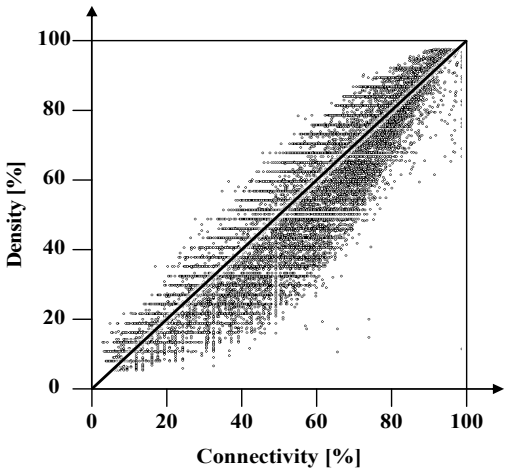


Fig. 5: Density plotted against connectivity

Fig. 6 shows a dataset created for test purposes, where a near circular forest area is surrounded by rings of gradually decreasing forest densities, as well as a few exemplary results for different window sizes. Obviously the moving window size is essential for the results. Generally increasing neighbourhood sizes correspond with more smoothed density and connectivity surfaces and will also result in decreasing core area sizes. For the test dataset the fragmentation categories show almost a circular shape for bigger neighbourhoods, while the effects of small clumps of forest pixels become less influential.

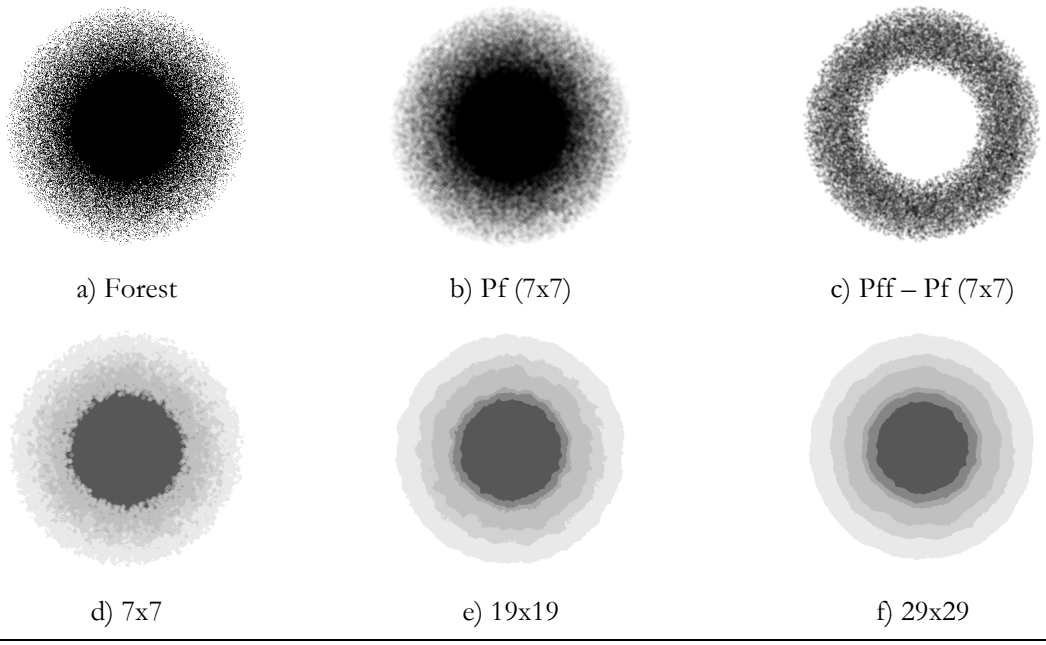


Fig. 6: Fragmentation model applied to a random forest distribution (400x400 pixels)

For this study the original fragmentation model has been implemented as SAGA module. A few optional enhancements have been made. Instead of using a squared window for the neighbourhood definition only pixels are taken into account, which lie within a given radius around the centre pixel (Fig. 5). Since distance relations are basically circular by nature, this gives more realistic results. For the connectivity calculation it is possible to include diagonal pixel pair relations additional to the ones in cardinal directions. But it turned out that this option has only minor influences on the results. Furthermore it is possible to modify the weighting of density and connectivity, when deciding, if a pixel belongs to category edge or category perforated.

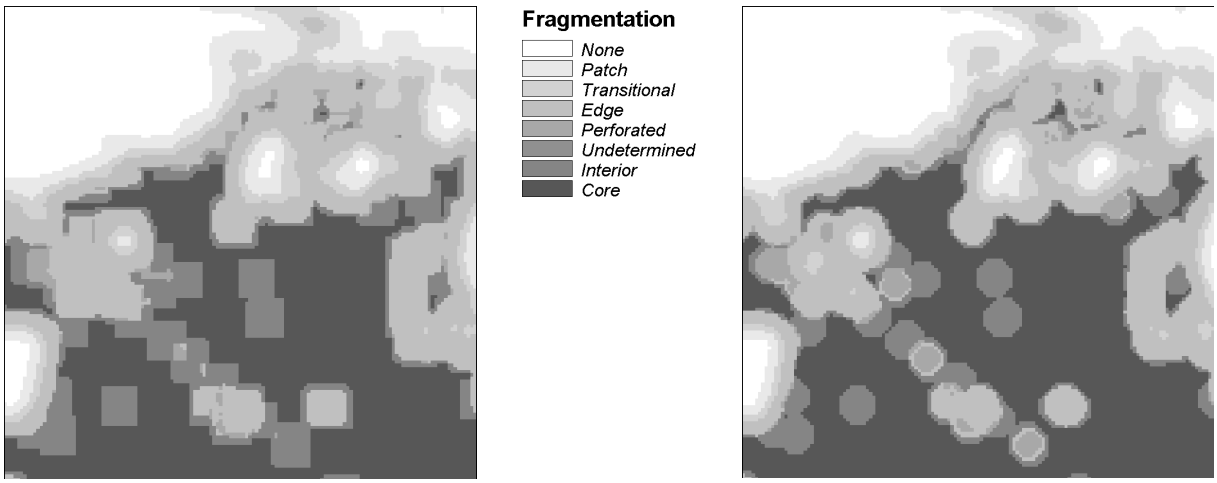


Fig. 7: Comparison of squared and circular neighbourhoods.

An alternative way for the calculation of density and connectivity has been implemented as further SAGA module. It is a two step resampling procedure, which is specially suited for investigating large neighbourhoods. In case of the density, each forest pixel gets assigned 100% density, all others a density of 0%. The first step is to resample this density raster to a raster with the desired neighbourhood as pixel size. Because the pixel size is increased, this process is called up-scaling. During the up-scaling process each resampled pixel is calculated as mean value of all original pixels it contains. In the second step the up-scaled raster is resampled back to the original pixel resolution using bicubic spline interpolation (Fig. 6). Both resampling techniques are directly supported by the SAGA application programming interface (CONRAD 2006). The connectivity calculation is similar except for the initialisation of the connectivity values, which are calculated with the standard method for the immediate neighbourhood, i.e. a 3x3 moving window.

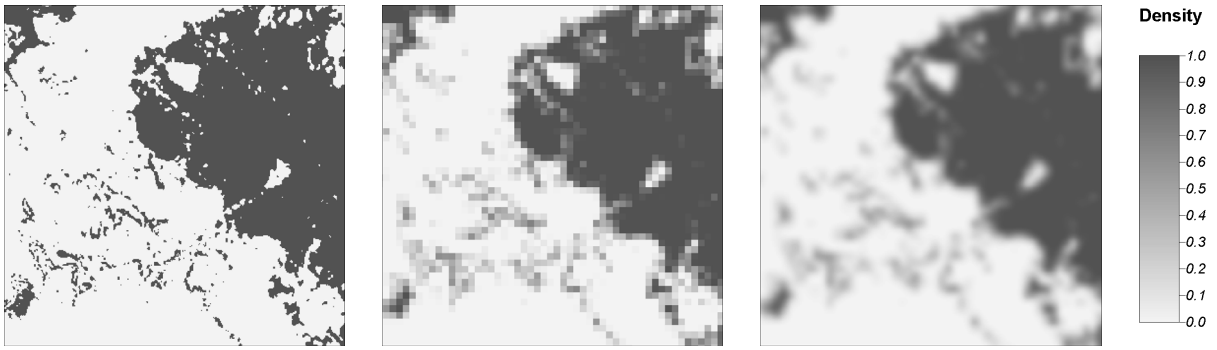


Fig. 8: Density calculation by coupled up and down scaling.

Bicubic spline interpolation fits a third order polynomial function to the 16 pixel values surrounding the interpolation point. Therefore the effective neighbourhood is higher than the pixel size chosen for the up-scaling, though the influence of nearer pixels remains dominant. Besides a big performance gain, the main advantage using the coupled resampling is that a smoother density and connectivity distribution is achieved, which results in less unexpected fragmentation category changes, induced by the fixed moving window size in the original model.

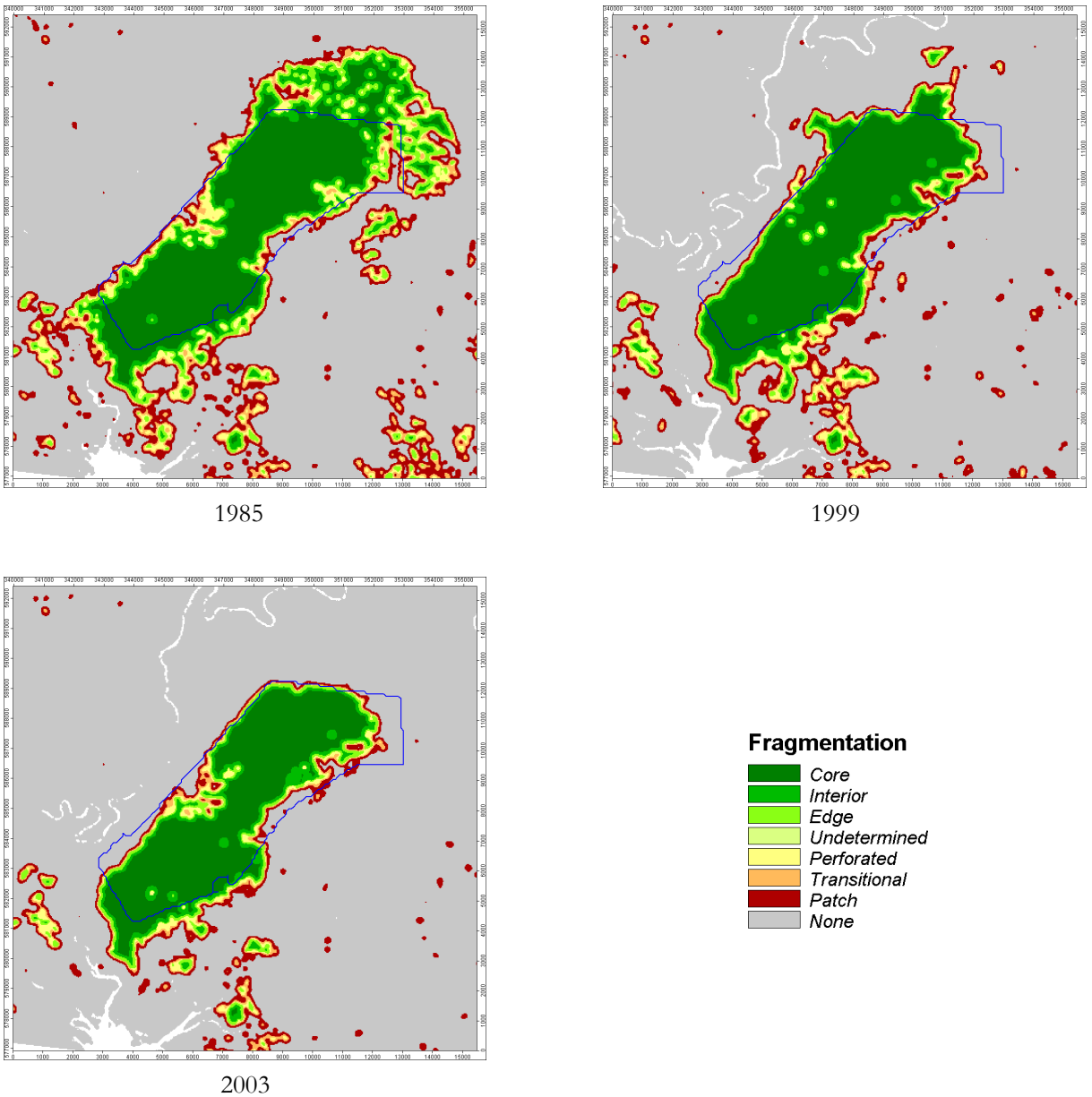


Fig. 9: Fragmentation categories calculated for peat swamp forest (effective neighbourhood 210m, blue outline: Klias Forest Reserve).

Forest Fragmentation in the Klias Peninsula: For analysing the fragmentation from 1985 to 2003, we selected the fragmentation results of the coupled resampling method with an effective neighbourhood of 210m, i.e. 7 times the pixel width of 30m, as the 'standard' for discussing the temporal pattern of fragmentation. Our discussion focuses on peat swamp forest (Fig. 9, Tab. 1).

Tab 1: Forest Fragmentation from 1985 to 2003 (ha)

Class	1985	1985-1999	1999	1999-2003	2003
Core	10414	-6344	4070	-636	3434
Interior	2291	-1647	644	-83	561
Perforated	2355	-1375	980	-370	610
Edge	2091	-1482	609	-136	473
Transitional	2199	-1033	1166	-485	680
Patch	5681	-1930	3751	-1677	2074
Undetermined	0	0	0	0	0
None	131827	13805	145632	3391	149023

Generally, the peat swamp forest in the Klias Peninsula has become fragmented in the past 18 years. In 1985, immediately after the establishment of the two protection forest reserves, more than half of the forest was still either core or interior. The core class comprised about 40% of the total peat swamp forest areas. The core and interior areas possess a high ecological integrity because the surrounding areas were forested. However, there was also significant number of patches. The patches summed to more than 5,500 ha and this is a quarter of the peat swamp forest areas. This indicates land clearing for development activities had led to fragmentation in the surrounding of the two large forest areas i.e. Binsuluk and Klias areas.

The deforested areas between 1985 and 1999 summed to almost 14,000 ha. It had led to a significant reduction of the core and interior classes where more than 7,500 ha were lost. The 1998 fires have caused large-scale deforestation in the Klias Peninsula (Phua et al., 2007). The high-integrity areas remained in 1999 was only less than 5,000 ha. It was further reduced partly by the recurrent fires in 2003 and partly by agricultural development. In 2003, the core and interior areas were less than 4,000 ha, which is only about one-third of the high-integrity areas in 1985.

There was a significant reduction of forest patches of about 2,000 ha to about 3,800 ha in 1999, but it was still relatively high. However, recent rapid agricultural development has accelerated the disappearance of the forest patches. It was further reduced to slightly more than 2,000 ha in 2003. This is attributable to small-scale agricultural development in the Klias Peninsula. There were recurrent fires in 2003 but occurred mostly at previously burned areas (Phua et al., 2008). This means the largest contributor to the deforestation of the small patches is likely the land clearing for agricultural development.

6 CONCLUSION

When changes on the land use leading to rapid disappearance of a precious ecosystem such as peat swamp forest, the need to know about the rate and cause of deforestation is obvious. It is also no lesser importance to monitor the contextual change in terms of fragmentation. The Klias and Binsuluk Forest Reserves had been converted from commercial forest to protection forest reserve in 1984 to conserve the peat swamp forest in the northern end of Borneo i.e. the Klias Peninsula. Due to very different satellite data, this study embarked on supervised classification of the satellite data, followed by multitemporal fragmentation analysis based on the fragmentation model of Riitters et al (2000). Some modifications of the model were also introduced.

Fire and agricultural activity have worked in different ways to cause the forest to disappear and became fragmented. Recurrent fires were the cause of the large-scale deforestation in the Klias Peninsula from 1985 to 2003. The two large pieces of peat swamp forest have shrunken. Many small forest patches were resulted but it was the agricultural activity that removed the patches and left the peat swamp forest to become two small islands in the Klias Peninsula. Some variations in the fragmentation classes calculated with different spatial scale were observable but the multitemporal patterns were consistent.

REFERENCES

ANDERSON, J.A.R. (1964): Tropical Lowland Peats in South-east Asia. – Royal Tropical Institute, Amsterdam.

CONRAD, O. (2006): SAGA - Entwurf, Funktionsumfang und Anwendung eines Systems für Automatisierte Geowissenschaftliche Analysen. – Dissertation. [Online] URL: http://webdoc.sub.gwdg.de/diss/2007/conrad/conrad.pdf

GREER, T., JESSEN, O., MTUNJAU, C. & S.F. YAP (2005): Hydrological assessment of the Klias Forest Reserve. In: KUGAN, F. & V.K. CHEY (Eds.): Conservation and Management of Other Wetlands in Sabah: Issues and Challenges. – Proceedings of the 9th SITE Research Seminar. Sandakan, Malaysia. pp. 39-49.

HURD, J., E. Wilson & D. Civco (2002): Development of a forest fragmentation index to quantify the rate of forest change. – Proc. 2002 ASPRS Annual Convention, Washington, D.C. 10p. [Online] URL: http://clear.uconn.edu/publications/research/tech_papers/Hurd_et_al_ASPRS2002.pdf

IDRIS, M. H., KURAJI, K. & M. SUZUKI (2005): Evaluating vegetation recovery following large-scale forest fires in Borneo and northeastern China using multi-temporal NOAA/AVHRR images. – Journal of Forest Research 10: 101–111.

KYUMA, K. (1992): Climate, geology and geomorphology of insular Southeast Asia in relation to formation and distribution of tropical peat. – In: KYUMA, K., VIJARINSORN, P. & Z. ZAKARIA (Eds.): Coastal Lowland Ecosystems in Southern Thailand and Malaysia. Showado Printing, Kyoto, pp. 31–40.

LAURANCE, W.F., LAURANCE, S.G., FERREIRA, L.V., RANKIN-DE MERONA, J.M., GASCON, C. & T.E. LOVEJOY (1997): Biomass collapse in Amazonian forest fragments. – Science 278: 1117-1118.

LINKE, J., BETTS, M.G., LAVIGNE, M.B. & S.E. FRANKLIN (2007): Structure, function, and change of forest landscapes. – In: WULDER, M.A. & S.E. FRANKLIN (Eds.): Understanding Forest Disturbance and Spatial Pattern: Remote Sensing and GIS Approaches. CRC-Taylor & Francis, Boca Raton, pp. 1-29.

PHILLIPS V.D. (1998): Peat swamp ecology and sustainable development in Borneo. – Biodiversity Conservation 7: 651-671.

PHUA, M.H. & S. TSUYUKI (2008): Evaluation of fire-induced deforestation of tropical peat swamp forest using remote sensing and GIS techniques: A case study of the Klias Peninsula, Sabah, Malaysia. – In: SANCHEZ, I.B. & C.L. ALONSO (Eds.): Deforestation Research Progress, New York, pp. 203-215.

PHUA, M.H., TSUYUKI, S., LEE, J.S. & H. SASAKAWA (2007): Detection of burned peat swamp forest in a heterogeneous tropical landscape: A case study of the Klias Peninsula. – Landscape and Urban Planning 82 (3): 103-116.

RIITTERS, K., WICKHAM, J., O'NEILL, R., JONES, B. & E. SMITH (2000): Global-scale patterns of forest fragmentation. – Conservation Ecology 4(2): 3. [online] URL: http://www.consecol.org /vol4/iss2/art3

RIITTERS, K., WICKHAM, J., O'NEILL, R., JONES, B., SMITH, E., COULSTON, J., WADE, T. & J. SMITH (2002): Fragmentation of Continental United States Forests. – Ecosystems 5: 815-822.

SABAH FORESTRY DEPARTMENT, 2006a. http://www.forest.sabah.gov.my/english/ForestResources/Forest-Reserves/tabid/72/Default.aspx (accessed March 2006)

SABAH FORESTRY DEPARTMENT, 2006b. http://www.forest.sabah.gov.my/caims/Level%202%20frame%20pgs/Class%201%20Frames/binsuluk_fr.htm (accessed March 2006)

SAYER, J.A. & T.C. WHITMORE (1991): Tropical moist forests: Destruction and species extinction. – Biological Conservation 55: 199-213.

SETO, K.C. & M. FRAGKIAS (2007): Mangrove conversion and aquaculture development in Vietnam: A remote sensing-based approach for evaluating the ramsar convention on wetlands. – Global Environmental Change, doi:10.1016/j.gloenvcha.2007.03.001.

UNDP/GEF (undated): Conservation and Sustainable Use of Tropical Peat Swamp Forests and Associated Wetland Ecosystems. – Project's Inception Report (Unpublished), Kota Kinabalu: Forest Research Institute Malaysia.

WADE, T.G., RIITTERS, K.H., WICKHAM, J.D. & K.B. JONES. (2003): Distribution and causes of global forest fragmentation.-Conservation Ecology 7(2): 7. [Online] URL: http://www.consecol.org/ vol7/ iss2/art7

WIEN, J.A. (1989): Spatial scaling in ecology. – Functional Ecology 3: 385-397.

CONTINENTAL-SCALE DIGITAL SOIL MAPPING USING EUROPEAN SOIL PROFILE DATA: SOIL PH

Hannes Isaak Reuter [1], *Luis Rodriguez Lado* [1], *Tomislav Hengl* [2] & *Luca Montanarella* [1]

[1]Institute for Environment and Sustainability – Land Management and Natural Hazards Unit – European Commission – Joint Research Centre (JRC), TP 280, Via Fermi 2749, I-21027 Ispra (Italy)

[2]University of Amsterdam – Institut for Biodiversity and Ecosystem Dynamic, Nieuwe Achtergracht 166, 1018 WV Amsterdam (Netherlands)

Abstract: We used a compilation of 12,333 soil pH measurements from 11 different sources to create a quantitative map of estimated soil pH values across Europe using a geostatistical framework based on regression-kriging. We made use of 54 auxiliary variables in the form of raster maps at 1km resolution to explain the differences in the distribution of soil pH_{CaCl2} and we added the kriged map of the residuals from the regression model. The goodness of fit of the regression model was satisfactory ($R^2_{adj} = 0.43$) and its residuals follow a Gaussian distribution. The map results agree with our previous knowledge about this soil property. The lowest values correspond to the soils developed on acid rock (granites, quartzite's, sandstones, etc), while the higher values are related to the presence of calcareous sediments and basic rocks. The validation of the model also shows that our models is quite accurate ($R^2_{adj} = 0.56$). This shows the validity of regression-kriging in the estimation of the distribution of soil properties when a large and adequately documented number of soil measurements are available.

1 INTRODUCTION

Soil acidity influences chemical, physical and biological soil properties as well as plant growth and the quality of fresh water in direct and indirect ways (SCHEFFER *et al.* 1992). The acidification of freshwater and soils constitutes an important environmental concern in Europe and North America since the first symptoms of soil water and degradation due to the deposition of atmospheric acidifying compounds were observed in the 1950's (ODEN 1968, ALCAMO *et al.* 1990) on lakes and forest ecosystems from northern areas in Europe. Although policies have already been adopted across Europe to reduce the emission level of these pollutants into the atmosphere, the acidification process is still one of the main threats for European soils as recently debated around the "Thematic Strategy for Soil Protection" in Europe. In addition to the atmospheric deposition of acidifying compounds, agricultural practices such as the use slurries and ammonium based fertilizers can contribute significantly to the acidification of soils (RASMUSSEN & ROHDE 1989). On the other hand, liming produces the opposite effect and corrects temporally the environmental negative effects of decreases in soil pH (INSAM & PALOJÄRVI 1995). It is clear that the effects of the acidification process do not only depend on intensity of the input of acidifying compounds but also on the sensitivity of soils to this degradation process. The same amount of acid inputs on a Distric Regosol in the north of Europe will produce more adverse consequences that the same acid amount on a calcic soil in the Mediterranean area. Thus, in order to determine the impact of an acidifying process we should previously determine the sensitivity of each type of soil to this threat.

According to ULRICH (1981, 1983) soils have a fixed sequence of buffering mechanisms in order to avoid pH reduction. As the acidification process progresses, these mechanisms enter into force sequentially, the buffering capacity of soils decreases progressively and only an abrupt pH descent is observed when each buffering mechanism is completely exhausted. In this sense, the value of the soil pH is an indicator of the buffer capacity of the soil (resistance of the soil to the acidification) and it can be used determine the risk in relation to this threat.

There is no consistent or harmonized soil profile database which would allow for the generation of an overview of pH values across Europe using digital soil mapping techniques. We therefore made an attempt to combine, harmonize and QA/QC check values from 11 available databases in order to create a quantitative map of estimated soil pH values across Europe. We are aware that these databases reflect biased sampling strategies, non precise coordinates, different sampling times (e.g. especially sensitive on agricultural soils) and many other factors which influence the predicted result. However, these databases represent a snap shoot of available data on soil pH currently available in ESDAC.

This paper describes the major distribution of soil pH with a reported quantitative uncertainty in order to determine the extent of the risk of acidification across Europe.

2 MATERIAL AND METHODS

Soil Database creation: Eleven soil profile databases which are available in the European Soil Data Center (ESDAC) have been used. ESDAC is the single focal point for soil information as one of the 10 data centers in Europe for Environmental Data (EEA, EUROSTATS and JRC). The main task is to work together with national and regional soil data centers to provide timely, quality controlled added value products for policy and research needs. The top layer of each soil profile stored in the following datasets has been obtained: ICP-Forest (FSS), soil profiles from the ecopedological map of Italy, FOREGS (SALMINEN *et al.* 2007), Spade (HIEDERER *et al.* 2006), Soveur (NACHTERGAELE *et al.* 2002), WISE (BATJES 2002), Galicia (RODRIGUEZ LADO 2008), Danube-SIS (Bavaria, Slovenia and Bulgaria) and Puglia. Data were projected to the INSPIRE complained Lambert Equal Area (LAEA) projection, if not already available in that projection. Secondly, the measurement of pH in the different databases were performed in different solutions (H_20, KCl, $CaCl_2$). For each database we adjusted if necessary to report the pH in $CaCl_2$ solution. The following linear pedotransfer functions were generated based on values reported for topsoil sampling locations from the WISE global soil database:

[1] $$pH_{CaCl2} = 0.9761 * pH_{H20} - 0.427 \quad (R^2 = 0.92, n=1997)$$

[2] $$pH_{CaCl2} = 1.0572 * pH_{KCl} + 0.123 \quad (R^2 = 0.90, n =377)$$

We tested the validity of the global dataset derived pedotransfer function (pH_{KCl}=0.87 pH_{H20}, R^2 =0.93) by comparing it against the fit obtained from pH data measured in KCl and H_20 from the Galicia DB (pH_{KCl}=0.84 pH_{H20}, R^2 =0.75, n = 414) and observed a good agreement.

Duplicate detection: We define a duplicate soil sampling location as a point which is within 1,000 m of another point. Some of the datasets used in the analysis have insufficient precision in their geo-location, allowing for false positives. We identified duplicates, and clustered these points with a minimum distance of 1,000 m to obtain single independent regions. From the total dataset of 12,333 sampling points, 2,093 were within 1,000 m of a neighboring point, resulting in 847 clusters.

Clustering of the duplicates occurred mainly in regions where denser sampling occurred (e.g. Italy had 571 duplicates out of 5,114 records) however even random sampling designs such as the 813 points extracted from the FOREGS DB contained two sampling locations within 1,000 m of neighboring records.

The average standard deviation of adjusted pH_{CaCl2} of all 847 clusters was 0.39 with a maximum standard deviation of 2.54, whereas the average difference between the minimum/maximum pH_{CaCl2} at each cluster location was 0.6 with a maximum difference of 4.1 reached at sampling location in Galicia. Minimum was in both cases 0.0. Therefore we can specify the measurement/ scale error for our analysis of ~0.5 unit pH_{CaCl2}.

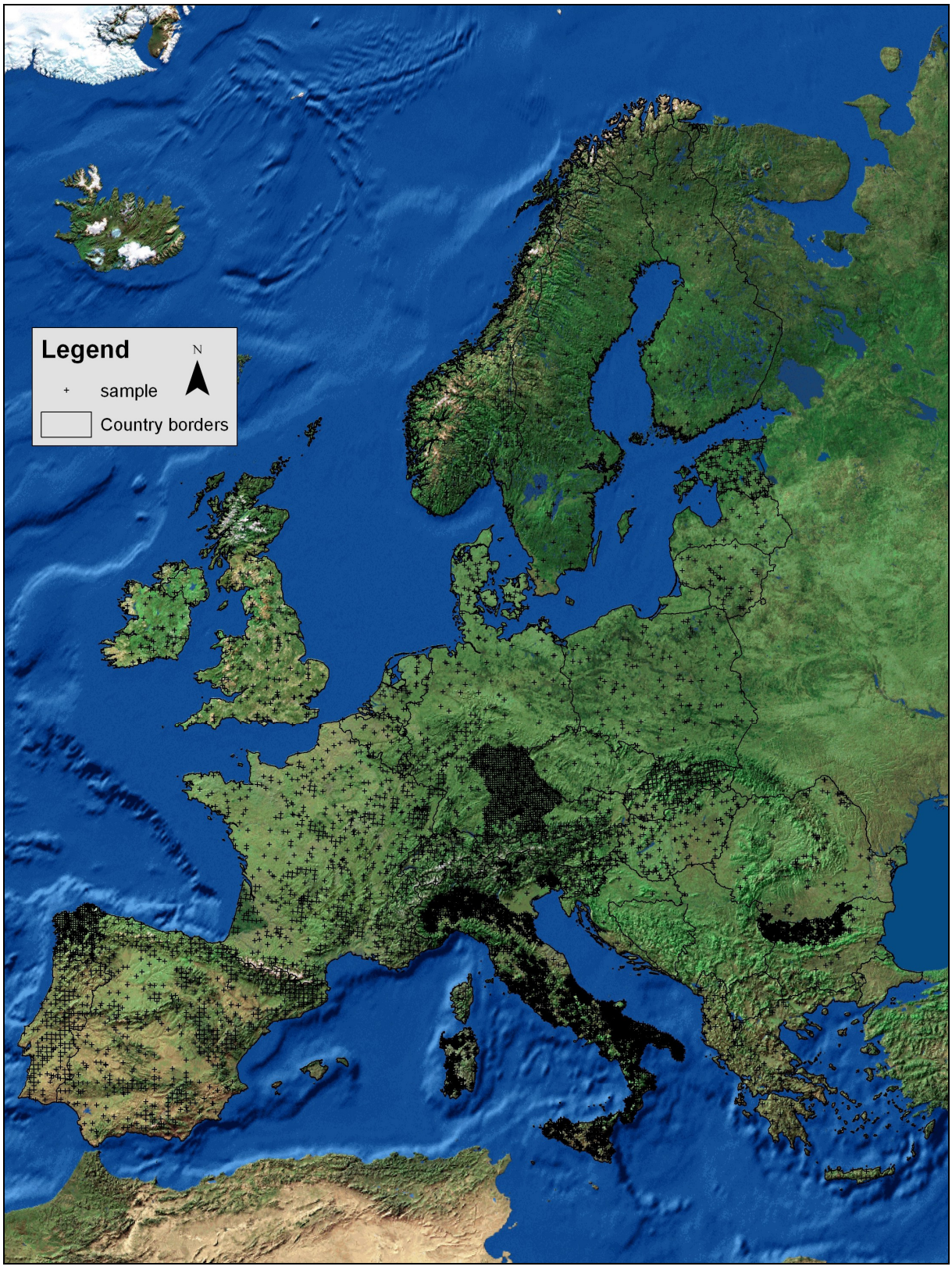


Fig. 1: Spatial distribution of soil sampling locations from the 11 different soil databases

We were interested to see if the chosen resolution was appropriate with respect to the pH_{CaCl2} point data. Coarsest, finest and recommended cell sizes were determined following HENGL (2006) and were ~200 m, ~1km and ~4.7 km respectively. The 0.5 probability to meet the next point was reached at ~3 km (0.95 at 30 km), for 2 points at ~5 km (0.95 at 43 km), for all points at 1,000 km (0.95 probability at 2,700 km).

Geostatistical mapping: We have used regression-kriging to estimate the pH_{CaCl2} values in soils of Europe. For a more in-depth description of the data preparation/treatment please refer to RODRIGUEZ *et al.* (2008), the general process of regression kriging is described in detail in HENGL (2007).

Firstly we create a linear regression model for the measured pH values against a number of auxiliary environmental variables and then we interpolate the residuals of this regression model by ordinary kriging. The final map is an additive combination of both models. This technique allows to take into account the boundaries of some environmental features that can highly influence the distribution of the studied soil property in the final map of estimates, and thus to obtain more realistic predictions.

The original dataset of observations (12,333 records) was divided using a random sampling function into a "*model*" dataset, that includes the 80% of the samples and a "*validation*" dataset, containing the remaining 20% of the samples, that was using for validation of the model. We used the "model" points dataset to build the multiple linear regression model. The derived regression equation has than be applied on the standardized 1 km resolution 56 auxiliary raster grids. The results have been aggregated to 5 km resolution. The auxiliary variables are either directly influencing soil pH or serve as a proxy for a factor:

- *Topography*: The 1 km DEM was derived from the SRTM30 V2 dataset obtained from the Jet Propulsion Laboratory. The SRTM DEM was used to derive a slope map, the Topographic Wetness Index (MOORE *et al.* 1991) and total incoming solar insolation (CONRAD 2001) using SAGA 2.0.
- *Geology*: We used the digital map of main geologic surface units in Europe (PAWLEWICZ *et al.* 2003). The original legend was reduced to 10 classes according to the genetic nature of each unit: (1) Granites, rhyolites and quartzites; (2) Paleozoic schists, phyllites, gneisses and andesites; (3) Shales and sandstones; (4) Mesozoic Ultramafic, basic phyllites, schists, limestones and evaporates; (5) Jurassic, Triassic and Cretaceous calcareous rocks; (6) Cenozoic serpentinites, gabros and sand deposits; (7) Tertiary basanites and andesites; (8) Neogene and Paleogene calcareous rocks; (9) Quaternary limestones and basaltic rocks; and (10) Other Ultramafic and undefined rocks.
- *EVI remote sensing images:* Monthly averaged MODIS images of the Enhanced Vegetation Index EVI at 1 km resolution for the period 01/01/2004 to 31/12/2006 were obtained from the MODIS Terra imagery at the Earth Observing System Data Gateway. We performed a Principal Component Analysis on the 19 complete mosaics and used the first five resulting components.
- *Image of Lights at Night:* The lights at night image for the year 2003 was obtained from the Defense Meteorological Satellite Program (http://www.ngdc.noaa.gov/dmsp/), which measures night-time light emanating from the earth's surface at 1 km resolution. This map is a proxy for urbanisation and is now increasingly used for quantitative estimation of global socioeconomic parameters as well as for human population mapping (SUTTON 1997, SUTTON *et al.* 1997, DOLL *et al.* 2000).
- *Distance to infrastructures*: The map of distances to roads, airports and utility lines was calculated using the distance operation in ILWIS and the GIS layers from the GISCO database of the European Commission (http://eusoils.jrc.it/gisco_dbm/dbm/home.htm).
- *Cumulative Earthquake's magnitude*: The cumulative earthquake's magnitude map was calculated by using the 90,000 registered earthquakes in period 1973-1994. These measurements were recorded in the Global Seismology point database (http://earthquake.usgs.gov/eqcenter/). We used the logarithmic measure of the "size" of an earthquake and then rasterized the point map to the 10 km grid by using the point density operation in ILWIS. This operation sums all earthquake magnitudes observed within a 10 km grid and gives a cumulative map of earthquake activity.
- *Land use:* We used the Corine Land Cover 2000 map of Europe generalized to a 1 km grid. For Switzerland, we used the Corine Land Cover from 1990 since no updated information was available. The CLC1990 classes for this country were adjusted to those described in the CLC2000 and both datasets were merged together and aggregated to 1 km resolution. The original 44 classes were simplified to 8 classes: (1) urban infrastructures; (2) agriculture; (3) forest; (4) natural vegetation; (5) beaches; (6) ice bodies, (7) wetlands and (8) water bodies. Class 5 (beaches) disappeared in the process of upscaling from 100m to 1km. Additionally data from the Global Land Cover classification (FRITZ *et al.* 2003) has been used for areas where no Corine classification has been available.

- *Land forms:* The land forms were calculated by a modified method proposed by IWAHASHI & PIKE (2007). Basically it combines the topographic variables slope gradient, surface texture and local convexity to create 16 classes of landforms from steep to gentle landforms with fine/coarse texture and low/high convexity.
- *Climatic variables*: Mean annual temperature and accumulated precipitation maps were obtained from the very high resolution raster layers created by HIJMANS *et al.* (2005) on a global scale at 1 km grid resolution. The annual potential evapotranspiration (PET) was calculated from monthly temperature data using the THORNTHWAITE method (1948). Runoff was calculated as the difference between annual accumulated precipitation and annual potential evapotranspiration.
- *Alkalinity release rates due to the weathering of primary minerals in soils*: The release of alkalinity by weathering of primary minerals in soils was calculated using the *"Simple Mass Balance Method"* as described in RODRIGUEZ-LADO *et al.* (2007).
- *Atmospheric deposition of contaminants:* We also used the estimated annual deposition and emission rates of cadmium, lead and mercury in Europe for year 2004 calculated within the European Monitoring Evaluation Programme (http://webdab.emep.int/). The original 50 km grids were downscaled to 1 km grids using ordinary kriging.

We further masked out non-soil surfaces such as water bodies (rivers, lakes, sea etc.) and permafrost areas. A consistent European wide water mask, indicating the percentage of water area inside a 1 km pixel, has been created based on the NASA SRTM V2 SWBDB dataset (RABUS *et al.* 2003), the CORINE land use classification, lakes contained in the GISCO data, base water reflection by the use of Image2000 dataset (DE JAGER *et al.* 2006), the GSHHS - Database (WESSEL & SMITH 1996), and the Global Lakes and Wetlands Database (LEHNER & DÖLL 2004).

Areas with permanent ice cover have been detected using the mean annual EVI derived from the MODIS 1 km images obtained for the years 2003 and 2004. In this case, ice cover, water bodies and bare rock areas were detected based on a negative VI index. The total soil-cover area for these 26 countries was estimated to be 4,217,241 km^2.

Each class within the categorical variables (geology, land use and land forms) was transformed to binary raster layers. Later all the auxiliary raster layers were standardized and finally converted to 54 Principal Components raster maps by Principal Component Analysis in ENVI v4.3in order to minimize collinearity between variables. The kriging of the residuals from the linear model was done directly in 5 km blocks. In addition we also obtained a measurement of the estimation errors associated to the kriging interpolation method.

Filtering the database of observations and building the regression model were performed in the statistical environment 'R v.2.6.0". The Principal Component Analysis of the standardized auxiliary variables was done using the image processing software ENVI. The raster linear model and the final regression kriging map were done in SAGA-GIS 2.0.

3 RESULTS AND DISCUSSION

Descriptive statistics: The database used in this study is a compilation of samples collected in 11 different surveys (Fig. 2). There is a high variability of pH values across Europe, ranging from 0.09 to 1340. About 51 % of the samples present pH ≤ 10 while 41 % present values higher than 50 ppb.

We observe that the big survey campaigns across Europe (Soveur, Spade and FOREGS) present similar boxplot diagrams. However, the differences in pH are very evident when comparing regional campaigns on acidic soils like "Galicia" (parent materials mainly of granitic nature) and on calcareous soils like that in "Puglia" (mainly limestones).

The higher pH measurements (pH =9) correspond to soils in Puglia (Italy) and high pH values (pH > 8) were also observed mainly in Bulgaria. Very low pH values (pH < 3) were observed in Bavaria (Germany), in the sandy soils from the "Landes" department in France, and in Belgium and UK.

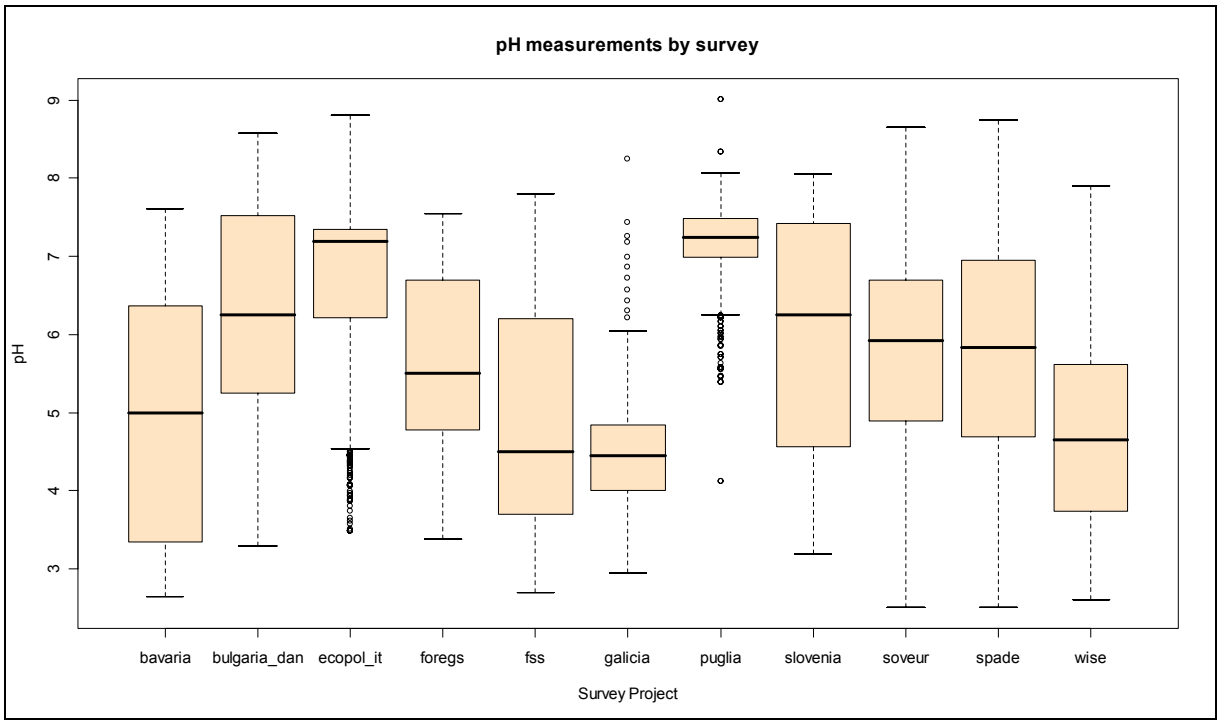


Fig. 2: Boxplot of the pH measurements by 11 different soil databases.

Geostatistics of the original data: The semi-variogram of the original data shows a moderate spatial dependency with a range of around 50-100 km (half of the total variance) and a slight linear trend up to 500 km (Fig. 3). Further analysis showed that this trend is due to a slight anisotropy in the east direction, whereas the north direction shows a periodic pattern. The semivariance shows two additional nested structures, which can be observed with a range of 600 km (reaching variance 1.6) and 800 km (reaching the total variance ~1.85). These two additional structures are due to anisotropy in the north –south direction, whereas the west-south direction shows only a prolonged trend as described before. Automated approaches fitting a semi variogram based on the range and total semivariance exist (see HENGL 2007). However, as we discovered these would have missed the 50-100 km scale variation structures.

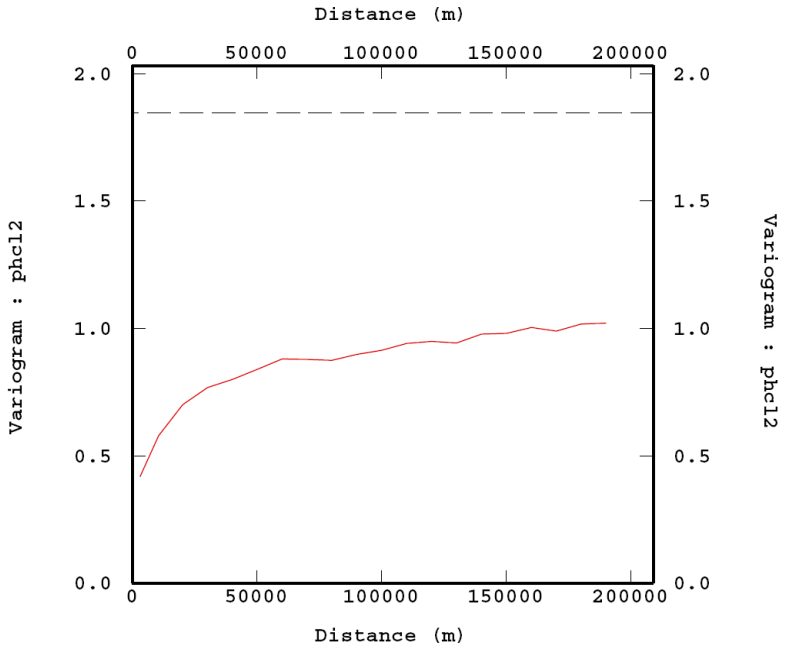


Fig. 3: Variogram of original pH_{CaCl2}

Linear Model results: The regression model obtained explains 43% of the variability (R^2_{adj}=0.433) and was significant at the $p<0.05$ level. Fifty one principal components contributed significantly ($p<0.05$) to the model (Tab. 1).

The residuals of the regression model showed spatial structure and they were incorporated to the final model by ordinary-kriging using the software ISATIS V8.1 (GEOVARIANCES 2008). A spherical variogram was fitted to the observed semivariances with a nugget of 0.55, a range of 20,000 and a sill of 1.22. The nugget/sill ratio (0.45) indicates that the residuals have a good spatial dependence.

Tab. 1: Regression coefficients for the multiple linear model.

VARIABLES	Estimate	VARIABLES	Estimate	VARIABLES	Estimate
(Intercept)	1.20E+03	[PCA18]	4.83E+01	[PCA36]	-7.49E-03
[PCA1]	2.28E-02	[PCA19]	1.26E+00	[PCA37]	8.70E-02
[PCA2]	2.69E-04	[PCA20]	-1.52E-01	[PCA38]	-1.16E-01
[PCA3]	9.64E-02	[PCA21]	-9.00E-01	[PCA39]	-9.67E-02
[PCA4]	-2.82E-01	[PCA22]	-5.82E-01	[PCA40]	2.88E-02
[PCA5]	-2.38E-01	[PCA23]	-4.20E+01	[PCA41]	6.24E-02
[PCA6]	2.89E-01	[PCA24]	4.21E-02	[PCA42]	6.13E-02
[PCA7]	4.17E-01	[PCA25]	-1.02E-01	[PCA43]	-8.06E-02
[PCA8]	-8.06E-01	[PCA26]	2.06E-02	[PCA45]	-1.14E-01
[PCA9]	-7.90E-01	[PCA27]	-5.87E-03	[PCA46]	4.67E-02
[PCA10]	-6.15E-01	[PCA28]	-3.65E-01	[PCA47]	-1.32E-02
[PCA11]	8.02E-02	[PCA29]	2.38E-01	[PCA48]	5.07E-02
[PCA12]	2.43E-01	[PCA30]	8.85E-03	[PCA50]	1.65E-02
[PCA13]	6.33E-01	[PCA31]	9.82E-02	[PCA51]	-9.83E-03
[PCA14]	5.59E-02	[PCA32]	-3.23E-03	[PCA53]	6.89E-03
[PCA15]	1.31E-01	[PCA33]	6.42E-02	[PCA54]	-2.77E-01
[PCA16]	4.30E-01	[PCA34]	-1.39E-01		
[PCA17]	-8.86E-01	[PCA35]	2.49E-01		

As suspected, the following map (Fig. 4) shows that the spatial distribution of soil pH is highly dependent on the nature of the parent material. We can observe low pH values in the granitic areas all over the Hesperic massif (Portugal and north of Spain), in the Vosges mountains, in the Pyrenees, and in the shallow soils from Scandinavia, mainly developed on acid materials. The higher pH values (pH > 7) are mainly present in the sedimentary areas of the Mediterranean countries (Spain south of France, Italy, Albania and Greece) because of the calcareous nature of the parent material. We also observe differences due to land use patterns and large scale climatic differences (e.g. the Mediterranean area versus Scandinavia). These results might be useful in guiding further research - for example – as to why specific stream water catchments in the Czech republic showed differences in pH time series in periods of increasing acidic deposition (VESELY *et al.* 2002).

According to our computations 16.7% of the territory has pH values lower than 4.2 and only 1.9 % of the area present values of pH > 8. The higher pH estimates are located throughout all European countries, mainly located close to major cities, or in arid areas with intensive agriculture area (southeast of Spain) or Deltas. The distribution of pH values in relation to land use classes (Fig. 5) showed that forest soils have lower pH values, while agricultural and urban areas present the higher mean values probably due to liming and the influence of the dissolution of cement in buildings.

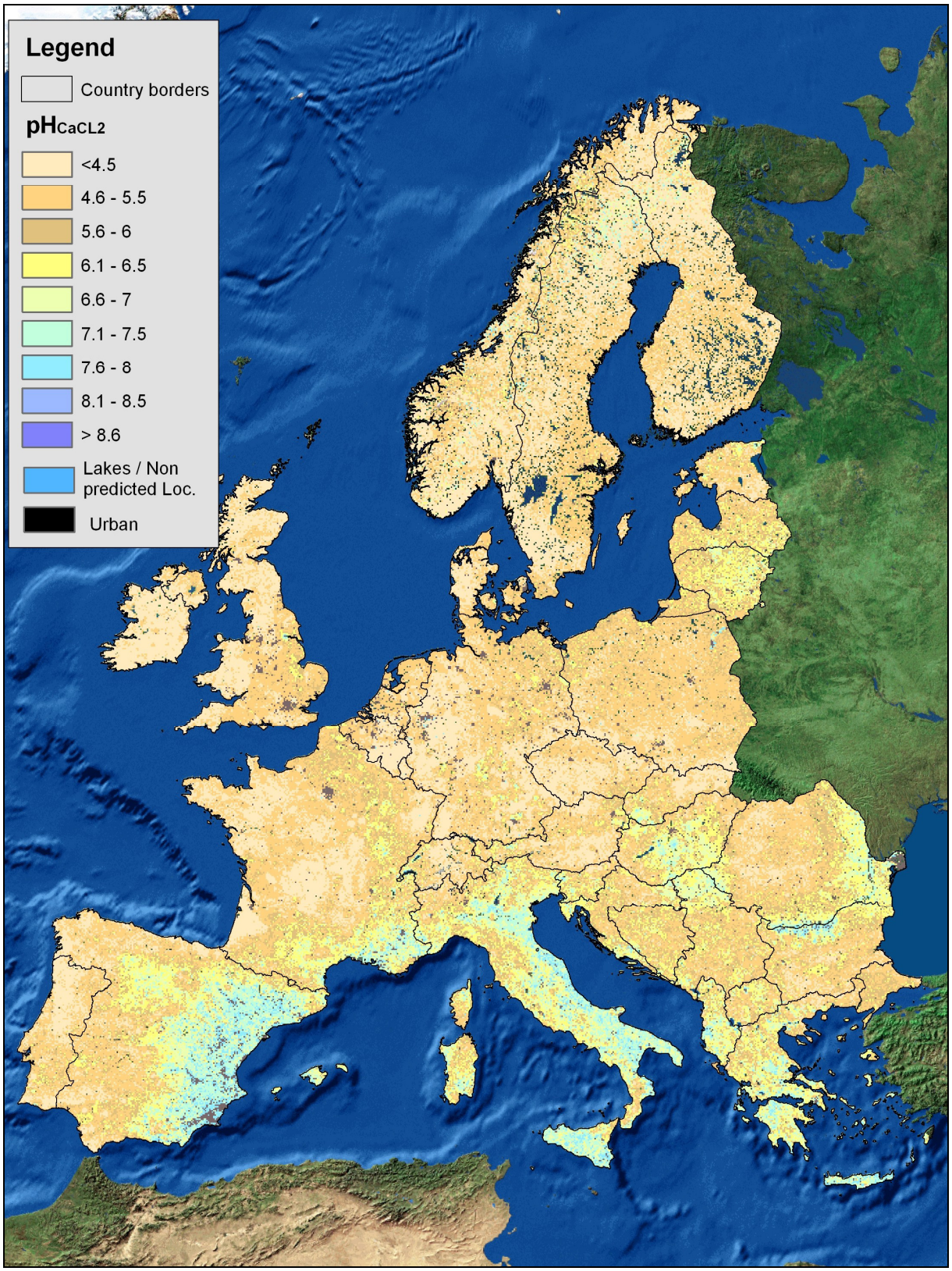


Fig. 4: Estimated values of pH_{CaCl2} for the EU27 MS and some adjacent countries.

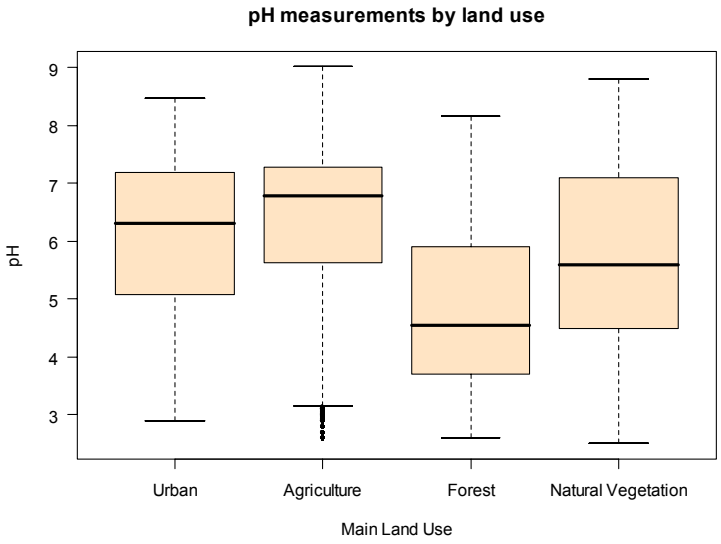


Fig. 5: Estimated pH_{CaCl2} boxplots for different major land uses

Validation: The accuracy of the model was assessed by comparing the pH measurements and their estimates in the 2,362 samples from the validation dataset. We performed a linear regression between both variables to check their relationship and we obtained a significant correlation between them (R^2_{adj}=0.56, α=0.05; Fig. 6). The equation of the regression model writes:

$$pH_{measured} = 1.34 + 0.779 * pH_{predicted}$$

Most of the samples fall within the limits of the 95% prediction bands. This relationship demonstrates that in general, our model slightly underestimates the real pH measurements. The Root Mean Squared Error of the predictions is low (RMSE=0.9). A Student's t-test for paired samples ($p < 0.05$) shows that there are not significant differences between measurements and the predictions. Our model does show a good fit in the prediction of soil pH values.

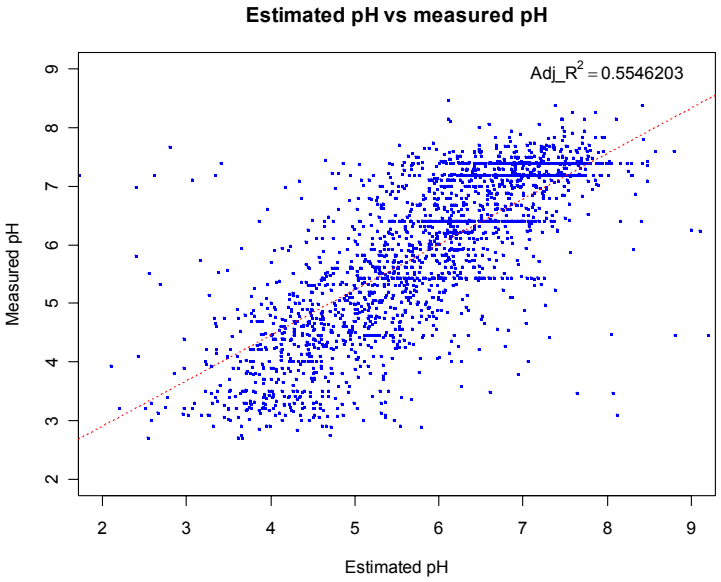


Fig. 6: Plot of measured vs. estimated pH in the validation dataset.

4 CONCLUSION

The study revealed that it was possible in a limited time and data frame to predict soil pH_{CaCl2}. This might serve as a first proxy to the sensitivity of soils to the process of acidification. Our linear model was satisfactory and the residuals showed clear auto-correlation structure. If new data are available to be included in the DSM process, the procedure can be recomputed and uncertainties reduced.

Any further assessment of sensitivity of soils to acidification has to take into account not only the pH but the soil organic matter content, soil texture and Al and Fe contents (KOPTSIK & ALEWELL 2007) as well as the sulphate sorption capacity of soils. One possible approach would be to generate a buffering capacity map using Digital soil mapping techniques, therefore providing an additional data set to the already existing critical load acidity maps (HETTELINGH *et al.* 1991, 1993, 1995). However, the author's first impressions are that such DSM approaches would need a dedicated sampling campaign as well as standardized measurements for these kinds of parameters that are generally missing in the databases used in the study.

A further geostatistical exercise would include simulations to predict probability density functions of the pH_{CaCl2} distribution. A disadvantage of our study is the lack of a temporal component. Most of the observations in the generated database contain no information on the time or timeframe of the sampling. If such information were available, maps for different times/time steps could be created, allowing results to be compared with changes in time as provided by e.g. de SCHRIJVER *et al.* (2006) or FÖLSTER *et al.* (2003).

REFERENCES

ALCAMO J., SHAW R. & L. HORDIJK (1990): The RAINS Model of Acidification. – Science and Strategies in Europe. Dordrecht, Netherlands: Kluwer Academic Publishers.

BATJES, N.H. (2002): Revised soil parameter estimates for the soil types of the world. – Soil Use and Management 18 (3):232–235.

CONRAD, O. (2001): Tools for (grid based) digital terrain analysis V1.0. – Program module for SAGA GIS 2.0, Göttingen.

DE SCHRIJVER, A., MERTENS, J., GEUDENS, G., STAELENS, J., CAMPFORTS, E., LUYSSAERT, S., DE TEMMERMAN, L., DE KEERSMAEKER, L., DE NEVE, S. & K. VERHEYEN (2006): Acidification of forested podzols in North Belgium during the period 1950-2000. – Science of the Total Environment, 361, 1-3, 189-195.

DE JAGER, A., RIMAVIČIŪTĖ, E. & P. HAASTRUP (2006): A Water Reference for Europe. – In: HAASTRUP, P. & WÜRTZ, J. (Eds.): Environmental Data Exchange Network for Inland Water (pp. 259-286), Elsevier.

DOLL, C.H., MULLER, J.P. & C.D. ELVIDGE (2000): Night-time Imagery as a Tool for Global Mapping of Socioeconomic Parameters and Greenhouse Gas Emissions. – AMBIO: A Journal of the Human Environment 29: 157-162.

FRITZ, S., BARTHOLOME, E., BELWARD, A., HARTLEY, A., STIBIG, H.-J. & H. EVA (2003): Harmonization, mosaicking, and production of the Global Land Cover 2000 database. – Joint Research Center (JRC), Ispra, Italy.

GEOVARIANCES (2008): ISATIS V8.1.1 Software for calculation of geostatistics.

HENGL, T. (2007): A Practical Guide to Geostatistical Mapping of Environmental Variables. – EUR 22904 EN Scientific and Technical Research series, Office for Official Publications of the European Communities, Luxemburg, 143 pp.

HENGL, T. (2006): Finding the right pixel size. – Computers & Geosciences 32 (9): 1283-1298.

HETTELINGH, J.P., DOWNING, R.J. & P.A.M. DE SMET (1993): Maps of critical loads, critical sulphur deposition and exceedances. – In: DOWNING, R.J., HETTELINGH, J.P. & P.A.M. DE SMET (Eds.): Calculation and mapping of critical loads in Europe. – Status report 1993. The Netherlands7 Coordination center for effects, RIVM; p. 6 – 18.

HETTELINGH, J.P, POSCH, M. & P.A.M. DE SMET (1995): Analysis of European maps. – In: POSCH, M., DE SMET, P.A.M., HETTELINGH, J.P. & R. DOWNING (Eds.): Calculation and mapping of critical thresholds in Europe. – Status report 1995. Bilthoven, The Netherlands7 Co-ordination Centre for Effects, National Institute of Public Health and the Environment; p. 5 – 22.

HETTELINGH, J.P., DOWNING, R.J. & P.A.M. DE SMET (1991): European critical loads maps. – In: HETTELINGH, J.P., DOWNING, R.J. & P.A.M. DE SMET (Eds.): Mapping critical loads for Europe. – CCE technical report 1. Bilthoven, the Netherlands7 Coordination Center for Effects, National Institute for Public Health and Environmental Protection; p. 5 – 30.

HIJMANS, R.J., CAMERON, S.E., PARRA, J.L., JONES, P.G. & A. JARVIS (2005): Very high resolution interpolated climate surfaces for global land areas. – International Journal of Climatology 25: 1965-1978.

HIEDERER, R., JONES, R.J.A. & J. DAROUSSIN (2006): Soil Profile Analytical Database for Europe (SPADE): Reconstruction and Validation of the Measured Data (SPADE/M). – Geografisk Tidsskrift, Danish Journal of Geography 106(1): 71-85.

FÖLSTER, J., BISHOP, K., KRAM, P., KVARNAS, H. & A. WILANDER (2003): Time series of long-term annual fluxes in the streamwater of nine forest catchments from the Swedish environmental monitoring program (PMK 5). – The Science of The Total Environment, 310, Issues 1-3: 113-120.

IWAHASHI, J. & R.J. PIKE (2007): Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature, Geomorphology 86: 3-4 and 409-440.

INSAM, H. & A. PALOJÄRVI (1995): Effects of forest fertilization on nitrogen leaching and soil microbial properties in the Northern Calcareous Alps of Austria. – Plant and Soil, 168-169,1: 75-81.

KOPTSIK, G. & C. ALEWELL (2007): Sulphur behaviour in forest soils near the largest SO2 emitter in northern Europe. – Applied Geochemistry 22,(6): 1095-1104.

LEHNER, B. & P. DÖLL (2004): Development and validation of a global database of lakes,reservoirs and wetlands. – Journal of Hydrology 296: 1-22.

MOORE, I.D., GRAYSON, R.B. & A.R. LADSON (1991): Digital Terrain Modelling: A Review of Hydrological, Geomorphological and Biological Applications. – Hydrological Processes 5: 3-30.

NACHTERGAELE F.O., VAN LYNDEN, G.W.J. & N.H. BATJES (2002): Soil and terrain databases and their applications with special reference to physical soil degradation and soil vulnerability to pollution in Central and eastern Europe. – In: PAGLIAIA, M. (Ed.): Advances in GeoEcology *35*. CATENA Verlag GMBH, Reiskrirchen.

NAVRÁTIL, T., KURZ, D., KRÁM, P., HOFMEISTER, J. & J. HRUSKA (2007): Acidification and recovery of soil at a heavily impacted forest catchment (Lysina, Czech Republic) – SAFE modeling and field results. – Ecological Modelling 205: 3-4 and 464-474.

ODEN, S. (1968): The acidification of air and precipitation and its consequences in the natural environment. – Ecological Committee Bulletin No. 1. Swedish Natural Science Research Council, Stockholm. Translations Consultants, Ltd. Arlington, Va.

PAWLEWICZ, M.J., STEINSHOUER, D.W. & D.L. GAUTIER (2003): Map Showing Geology, Oil and Gas Fields, and Geologic Provinces of Europe including Turkey.

RABUS, B., EINEDER, M., ROTH, A. & R. BAMLER (2003): The shuttle radar topography mission – a new class of digital elevation models acquired by spaceborne radar. – Photogrammetric Engineering and Remote Sensing 57: 241-262.

RASMUSSEN, P.E. & C.R. ROHDE (1989): Soil Acidification From Ammonium-Nitrogen Fertilization in Moldboard Plow and Stubble-Mulch Wheat-Fallow Tillage. – Soil Sci Soc Am J. 53: 119-122.

RODRIGUEZ LADO, L., HENGL, T. & H.I. REUTER (2008?): Heavy metals in European soils: a geostatistical analysis of the FOREGS Geochemical database. – Geoderma (in revision).

RODRIGUEZ LADO, L. (2008): personal communication.

RODRIGUEZ-LADO, L., MONTANARELLA, L. & F. MACÍAS (2007): Evaluation of the sensitivity of European soils to the deposition of acid compounds: Different approaches provide different results. – Water, Air and Soil Pollution 185: 293-303.

SALMINEN, R., BATISTA, M.J., BIDOVEC, M., DEMETRIADES, A., DE VIVO, B., DE VOS, W., DURIS, M., GILUCIS, A., GREGORAUSKIENE, V., HALAMIC, J., HEITZMANN, P., LIMA, A., JORDAN, G., KLAVER, G., KLEIN, P., LIS, J., LOCUTURA, J., MARSINA, K., MAZREKU, A., O'CONNOR, P.J., OLSSON, S.Å., OTTESEN, R.-T., PETERSELL, V., PLANT, J.A., REEDER, S., SALPETEUR, I., SANDSTRÖM, H., SIEWERS, U., STEENFELT, A. & T. TARVAINEN (2005): Geochemical Atlas of Europe. – Part 1/2 - Background Information, Methodology and Maps.

SCHEFFER, F., SCHACHTSCHABEL, P., BLUME, H.P., BRÜMMER, G., HARTGE, K.H., SCHWERTMANN, U., FISCHER, W. R., RENGER, M. & O. STREBEL (1992): Lehrbuch der Bodenkunde. – Enke, Stuttgart, 1-491.

SUTTON, P. (1997): Modeling population density with night-time satellite imagery and GIS. – Computers, Environment and Urban Systems 21: 227-244.

SUTTON, P., ROBERTS, D., ELVIDGE, C. & H. MEIJ (1997): A comparison of nighttime satellite imagery and population density for the continental United States. – Photogrammetric Engineering and Remote Sensing 63: 1303–1313.

THORNTHWAITE, C.E. (1948): An approach towards a rational classification of climate. – Geographical Review 38: 55-94.

ULRICH, B. (1981): Theoretische Betrachtungen des Ionenkreislaufs in Waldökosystemen. – Zeitschrift für Pflanzenernährung und Bodenkunde 144: 647-659.

ULRICH., B. (1983): Soil acidity and its relation to acid deposition. – In: ULRICH, B. & J. PANKRATH (Eds.): Effects of Accumulation of Air Pollutants in Forest Ecosystems. – Proc. Workshop,16-19 May, 1982, Gijttingen.Reidel, Dordrecht, The Netherladds, pp. 127 146.

VESELÝ, J., MAJER, V., & S.A. NORTON (2002): Heterogeneous response of central European streams to decreased acidic atmospheric deposition, Environmental Pollution 120(2): 275-281.

WESSEL, P. & W.H.F. SMITH (1996): A Global Self-consistent, Hierarchical, High-resolution Shoreline Database. – J. Geophys. Res. 101: 8741-8743.

DIGITAL TERRAIN MODEL GENERATION FROM AIRBORNE LASER SCANNING POINT DATA AND THE EFFECT OF GRID-CELL SIZE ON THE SIMULATION RESULTS OF A DEBRIS FLOW MODEL

Volker Wichmann [12]*, Martin Rutzinger* [13] *& Michael Vetter* [3]

[1]alpS – Centre for Natural Hazard Management, Grabenweg 3, 6020 Innsbruck (Austria)

[2]Laserdata GmbH, Technikerstr. 21a, 6020 Innsbruck (Austria)

[3]Institute for Geography – University of Innsbruck, Innrain 52, 6020 Innsbruck (Austria)

Abstract: Airborne laser scanning technology allows a rapid and cost-effective measurement of topography at high spatial resolutions over large areas. In this paper we present the generation of digital terrain models (DTMs) with different grid-cell sizes from a classified point cloud including several postprocessing steps such as morphological filtering and surface depression filling. A qualitative analysis is carried out to investigate the effect of grid-cell size on the simulation results of a debris flow model. Like most of the available debris flow models for natural hazard assessment, this model was originally developed for application on regional scales. So far, only few studies address possible consequences arising out of the application of such models on high-resolution DTMs. The results of this study suggest, that the debris flow model is only applicable at a certain range of scales and that flow path routing algorithms, not taking into account the local flow depth, have to be used with care on DTMs, which preserve a high level of topographic detail.

1 INTRODUCTION

Recent computational methods to predict landscape evolution by simulating geomorphic and hydrologic processes rely heavily on the accuracy of digital elevation models (WALKER & WILLGOOSE 1999). Airborne laser scanning (ALS), an active remote sensing technology, is determining the distance from the sensor to a reflecting surface by measuring the travel time of an emitted laser pulse. The ALS measurements result in a three-dimensional cloud of points (x, y, z, i) with irregular spacing. Because of the smaller footprint of the emitted signal, the spatial resolution is better than with radar measurements. Current systems in operational use record the first and last echo of the reflected beam. Newer systems are able to record also intermediate echoes. Classification of the point cloud into ground and non-ground points is called filtering, which is crucial to obtain digital terrain models (DTMs). So far, several approaches for filtering have been proposed (cf. SITHOLE & VOSSELMANN 2004, PFEIFER & BRIESE 2007), for example: morphological filtering (e.g. LOHMANN *et al.* 2000, VOSSELMANN 2000, ZANGH *et al.* 2003), progressive densification (e.g. AXELSSON 2000, VON HANSEN & VÖGTLE 1999), iterative (and hierarchical) interpolation (e.g. KRAUS & PFEIFER 1998, PFEIFER *et al.* 2001) and segmentation based filtering (e.g. JACOBSEN & LOHMANN 2003, TÓVARI & PFEIFER 2005). Irrespective of the filtering method used, the classification results show two basic errors. Commission error that classifies non-ground points as ground measurements and omission error that removes ground points mistakenly. An optimum filter setting needs to be determined to minimize both errors. In many cases such an unique setting does not exist. Especially in steep and wooded areas, terrain and vegetation show similar characteristics (large elevation differences at small horizontal distances) which make it difficult to classify terrain points (e.g. KOBLER *et al.* 2007). Besides these methodological problems, the processing of high-resolution and thus storage intensive data is, irrespective the continually improving computational power of modern hardware, still problematic and requires special attention.

In general, the use of ALS data in natural hazard management is manifold. The applications range from flood and avalanche simulation, rockfall modeling, to object and infrastructure protection. For example, CAVELLI & MARCHI (2008) use a high-resolution DTM and derived datasets such as an index of topographic roughness to investigate and classify alluvial fan morphology with regard to debris flow deposit occurrence. An overview of further applications of ALS data in natural hazard management is given by GEIST *et al.* (in press).

The use of ALS data for process modeling is very promising because, especially in forested areas, the resolution of DTMs derived from laser scanning are more accurate than any other conventional available DTMs (e.g. from photogrammetry, KRAUS & PFEIFER 1998). However, so far only few studies have addressed the consequences arising out of the use of such high-resolution DTMs for simulation models originally developed on scales where less topographic detail is preserved (e.g. WALKER & WILLGOOSE 1999, SCHMIDT *et al.* 2003). Such models usually use flow path routing algorithms that require a depressionless DTM and that route the flow to the lower neighboring cells of the currently processed grid-cell (e.g. ZIMMERMANN *et al.* 1997, GAMMA 2000, WICHMANN & BECHT 2004).

In this study we investigate the effect of grid-cell size on the simulation results of such a debris flow model. The debris flow model used is described in Section 2. Section 3 introduces the study area and the used data sets. The workflow of DTM generation from ALS point data is explained in Section 4. Next the calibration of the used model is presented in Section 5. Afterwards the variation of parameter settings with grid-cell size is analyzed and discussed.

2 DEBRIS FLOW MODEL

In the following the model is described briefly. The reader is referred to WICHMANN & BECHT (2005) and WICHMANN (2006) for a more detailed description of the model. The model is based on two coupled algorithms. The first one is a grid-based random walk to determine flow directions. The second one is used to calculate the velocity change between two grid-cells of the flow path and thus the run-out distance. The flow direction algorithm uses the *mfdf* (*multiple flow directions for debris flows*) criterion established by GAMMA (2000) to determine potential successors of the currently processed grid-cell. The set of potential successor cells N is given by

$$\mathbf{N} = \left\{ n_i \mid \begin{cases} \gamma_i \geq (\gamma_{max})^a & \text{if } 0 < \gamma_{max} \leq 1 \\ \gamma_i = \gamma_{max} & \text{if } \gamma_{max} > 1 \end{cases}, i \in \{1,2,...8\}, a \geq 1 \right\} \quad [1]$$

with

$$\gamma_i = \frac{\tan\beta_i}{\tan\beta_{thres}}, \qquad \beta_i \geq 0, \qquad i \in \{1,2,...8\} \quad [2]$$

where γ_{max} is max(γ_i), β_i is the slope to neighbor cell i, β_{thres} is a slope threshold, and the exponent a is a parameter controlling divergent flow. The set N is reduced to the neighbor of steepest descent if $\gamma_{max} > 1$. In case γ_{max} is in the range 0 to 1, the probability for each neighbor i to be selected as successor is given by

$$p_i = \begin{cases} \frac{\tan\beta_i \cdot p}{\sum_j \tan\beta_j} & \text{if } i' \in \mathbf{N} \\ \frac{\tan\beta_i}{\sum_j \tan\beta_j} & \text{if } i' \notin \mathbf{N} \end{cases} \quad i,j \in \mathbf{N} \quad [3]$$

where i' denotes the previous flow direction selected. If the set contains i', abrupt changes in direction can be reduced by a higher weighting of this successor. Therefore a persistence factor p is introduced, which is also contained in the calculation of the sum. These transition probabilities are scaled to accumulated values between 0 and 1, and a pseudo random number generator is used to select one grid-cell from the set. Because of the selection of a successor by chance, each model run results in varying flow paths. This property is used to calculate several model runs (Monte Carlo simulation) from each debris flow initiation site. A high enough number of runs is necessary to reproduce the whole extent of the process area.

The parameters of the algorithm allow to adjust the flow divergence to different relief. In general, a tendency towards the steepest decent is achieved by weighting the transition probabilities by slope. The slope threshold determines if divergence is simulated at all or, in combination with the divergence exponent, to which degree. The persistence factor can be used to increase the probability to retain a chosen flow direction.

A 2-parameter, center-of-mass, friction model, originally developed for calculating the run-out distance of snow avalanches (PERLA *et al.* 1980), is used to calculate the change in velocity from grid-cell to grid-cell. Assuming that the motion is mainly governed by a sliding friction coefficient (μ) and a mass-to-drag ratio (M/D in [m]), the velocity on grid-cell i is given by

$$v_i = \sqrt{\alpha_i * \left(\frac{M}{D}\right)_i (1 - \exp^{\beta_i}) + \left(v_{(i-1)}\right)^2 \exp^{\beta_i}} \quad [4]$$

with

$$\alpha_i = g\left(\sin\theta_i - \mu_i \cos\theta_i\right) \quad [5]$$

$$\beta_i = \frac{-2L_i}{\left(\frac{M}{D}\right)_i} \quad [6]$$

where v is the velocity [m/s], g is the acceleration due to gravity [m/s²], θ is the slope and L is the slope length [m] from predecessor to successor cell. At concave transitions in slope the predecessor velocity $v_{(i-1)}$ is adjusted due to the momentum change by

$$v^*_{(i-1)} = \begin{cases} v_{(i-1)} \cos\left(\theta_{(i-1)} - \theta_i\right) & \text{if } \theta_{(i-1)} \geq \theta_i \\ v_{(i-1)} & \text{if } \theta_{(i-1)} < \theta_i \end{cases} \quad [7]$$

In order to avoid the mathematical redundancy that different combinations of μ and M/D result in the same run-out distance, it has become common practice to set M/D to a value that results in realistic maximum velocities for the process under study (e.g. ZIMMERMANN *et al.* 1997; Gamma 2000), and to adjust only μ to reproduce the observed run-out distance.

The model, implemented as SAGA module, requires at least a DTM and a grid with initiation sites as input. Additional input grids can be provided to, for example, allow for spatially distributed friction coefficients or to automatically determine objects at risk. Further functionality, especially an extension of the model to simulate material deposition (which is not used in this study), is described in WICHMANN (2006).

In order to investigate the parameter settings of the debris flow model on different grid-cell sizes, DTMs with different resolutions are prepared. In a first step, a raster dataset with a cell size of 1 m is aggregated from the point cloud classified as ground. This dataset is refined by removing residual surface features like trees with a progressive morphological filter and high-frequency noise with a Gaussian filter. In the following this dataset is resampled to coarser grid-resolutions. In a final preprocessing step, depression-free, hydrological correct DTMs are calculated.

Based on true color orthophotos and additional datasets derived from the DTM, four well distinguishable debris flows of different size are mapped. Next, the debris flow model is calibrated to match the process areas as far as possible. Finally, the model parameter variation is analyzed and discussed.

3 STUDY AREA AND DATA SETS

The study area is located in the high mountain region of Vorarlberg, Austria (Gargellen Tal, Fig. 1), with altitudes rising from 1295 m to 2770 m a.s.l. The land cover in the valleys is characterized by coniferous and mixed forests, shrubs and meadows. For the study, two tests sites are selected. The first one is located above the tree line and comprises slopes with pioneer vegetation and hiking trails. The second one comprises some buildings of a small village, forested areas as well as pasture and farming land, roads, and hiking trails.

The ALS data was collected for an area of 137 km² by two flight-campaigns in December 2002 (valleys, leaf-off conditions, ALTM 1225, pulse repetition rate 25 kHz) and July 2003 (mountainous regions, ALTM 2050, 50 kHz) by the company TopScan[1] in the course of the ALS data acquisition campaign of the Federal State of Vorarlberg (cf. Rieger et al. 2005). The average point density on the ground varies from 0.9 points per m² (winter campaign) to 2.7 points per m² (summer campaign). The point cloud was classified into ground and non-ground points by hierarchical robust filtering using SCOP++[2] (HOLLAUS *et al.* 2005).

4 DTM GENERATION AND DERIVED DATASETS

Ground points from the classified ALS point cloud form the basis for the raster DTM computation. The point cloud is stored in a PostgreSQL/POSTGIS[3] database (HÖFLE 2007, PETRINI-MONTEFERRI *et al.* 2008) and can be queried spatially. We implemented a special SAGA module to retrieve the data from the database and bin it into a raster map with a grid-cell size of 1 m. From the variety of statistical methods available for binning the data, the calculation of a trimmed mean from all points falling into a grid-cell is favored. By discarding 5% of the smallest and largest observations, pronounced outliers are removed. This first preprocessing stage is finalized by closing small data gaps (i.e. grid-cells containing no points) with SAGA's CloseGaps module.

The calculation of shaded relief maps from this initial DTM shows errors arising from incorrectly classified points within the database like trees and other non-ground features. These non-ground points are detected with a progressive morphological filter (ZANGH *et al.* 2003), which we implemented as a 2-D version in SAGA. The major component of the morphological filter is an opening operation. To detect non-ground objects of various size, the window size of the filter is gradually increased. In each iteration, the filtered minimum elevation surface and an elevation difference threshold is used to separate ground and non-ground grid-cells. The difference threshold can be determined based on the slope of topography in the study area and may vary with filter window size. Please refer to ZANGH *et al.* (2003) for a comprehensive description of the algorithm.

We use an initial window size of 3 grid-cells and a constant elevation difference threshold of 2 m. The window size is increased linearly, resulting in a final window size of 13 grid-cells after six iterations. Common to all filtering methods for removing non-ground features is that problems occur at steep terrain. Especially at ridges omission errors (ground points are removed mistakenly) are usual. These errors are handled by applying an elevation threshold (tree line at about 2000 m a.s.l.) and a mask digitized along ridges in lower altitudes to remove these grid-cells from the non-ground mask. After applying the non-ground mask to the initial DEM, small data gaps are again filled up with the CloseGaps module. This raw version of the DTM is then analyzed by calculating shaded relief maps with varying angles of illumination. These maps show minor elevation differences in areas of overlapping flight strips indicating inaccurate strip adjustment. These irregularities are traced in biased flow paths calculated from the DTM. In order to remove the noise, the terrain is smoothed by applying SAGA's Gaussian filter module with a standard deviation of 1 and a radius of 2 m.

The smoothed DTM with a cell size of 1 m is used to prepare raster maps with coarser grid-cell sizes (DTM1, DTM2.5, DTM5, DTM10, DTM25 and DTM50; the numbers indicate the grid-cell size in meters) by bilinear interpolation with SAGA's resampling module. In a final preprocessing stage, hydrological correct DTMs are calculated. Therefore we implemented an algorithm proposed by WANG & LIU (2006) to efficiently fill surface depressions in large datasets. The method is enhanced to preserve a minimum downward slope (here 0.01°) instead of creating flat areas. This ensures valid flow directions for each grid-cell. The final DTM1 is used to derive further datasets (shaded relief, slope, catchment area, contour lines, etc.) for general use.

1 http://www.topscan.de
2 http://www.inpho.de; www.ipf.tuwien.ac.at
3 http://postgis.refractions.net

5 DEBRIS FLOW MAPPING AND MODEL CALIBRATION

Based upon colored orthophotos and the shaded relief map, four clearly distinguishable debris flows of different size were selected for this study. The outlines of the debris flows (erosion and sedimentation areas) were mapped based on orthophotos and the datasets derived from DTM1 (Fig. 1). The characteristics of the debris flows are listed in Table 1.

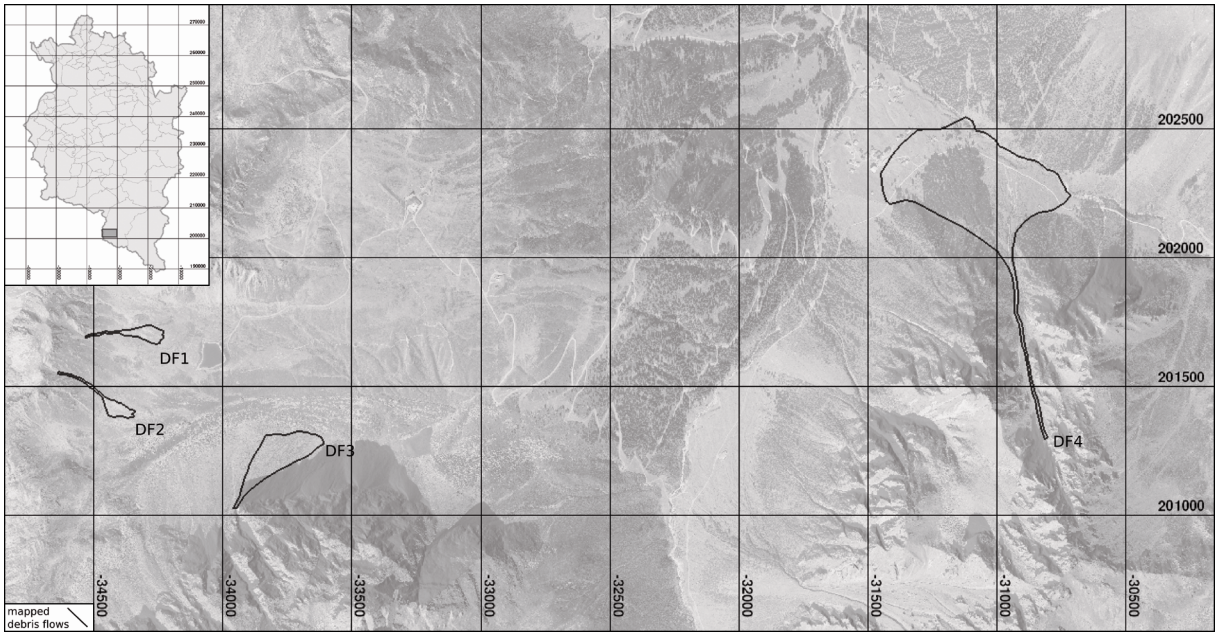


Fig. 1: Location of the study area in Vorarlberg, orthophoto draped over shaded relief, and mapped debris flows (DF1, DF2, DF3 and DF4; coordinate system GK M28, scale unit is meter).

Tab. 1: Characteristics of the mapped debris flows.

debris flow	perimeter [m]	area [m²]	min altitude [m]	max altitude [m]	mean slope [°]	number of cells DTM1	number of cells DTM50
DF1	681	8131	2244	2370	17.1	8141	6
DF2	751	6957	2226	2401	22.3	6957	6
DF3	997	34844	2010	2224	24.5	34848	16
DF4	3401	220834	1562	2117	15.9	220829	97

The three debris flows DF1 to DF3 are typical slope-type debris flows and are completely located above the tree line. The two smaller ones (DF1 and DF2) vary in mean slope. Debris flow DF3 has an area about five times larger than DF2. DF3 has also the highest mean slope. Debris flow DF4 is a torrent bed-type debris flow, whose sediment is mainly supplied by smaller slope-type debris flows in the upper part of the catchment. The debris flow has build up a huge fan, from which the largest part is mainly inactive. About half of the fan is forested. To the northwest of the mapped area, the fan reaches farther as mapped and visible in the orthophoto. The shaded relief map shows a small scale topography typical for debris flow deposits even in areas, which are now agriculturally used. Nevertheless, in this study we delimit the debris flow fan to the extent shown in Figure 1.

The model calibration is done stepwise: To facilitate the calibration, the persistence factor is kept at a value of 1 (no weighting) as long as the extent of the process area can be reproduced with this value. The slope threshold is varied until a realistic beginning of divergent flow is achieved. In some cases, the threshold is slightly increased during the next calibration step to increase the lateral flow on the debris fan. The exponent of divergent flow is increased until the extent of the fan is reproduced as accurate as possible. In the last step, the number of model runs is adjusted until reruns of the simulation model produce almost the same results. The run-out distance is approximated by setting the mass-to-drag ratio (M/D) to a value of 75 m and varying the sliding friction coefficient (μ).

6 MODEL PARAMETER ANALYSIS

The simulation results for different grid-cell sizes are shown in Figure 2 (DF1 to DF3) and Figure 3 (DF4). The three smaller slope-type debris flows DF1 to DF3 are well reproduced on grid-cell sizes up to 5 m. With greater grid-cell sizes, the small scale topography is not further preserved in the DTMs, resulting in inaccurate flow paths of the two smaller debris flows (DF1 and DF2) and finally in an overestimation of the process area of debris flow DF3.

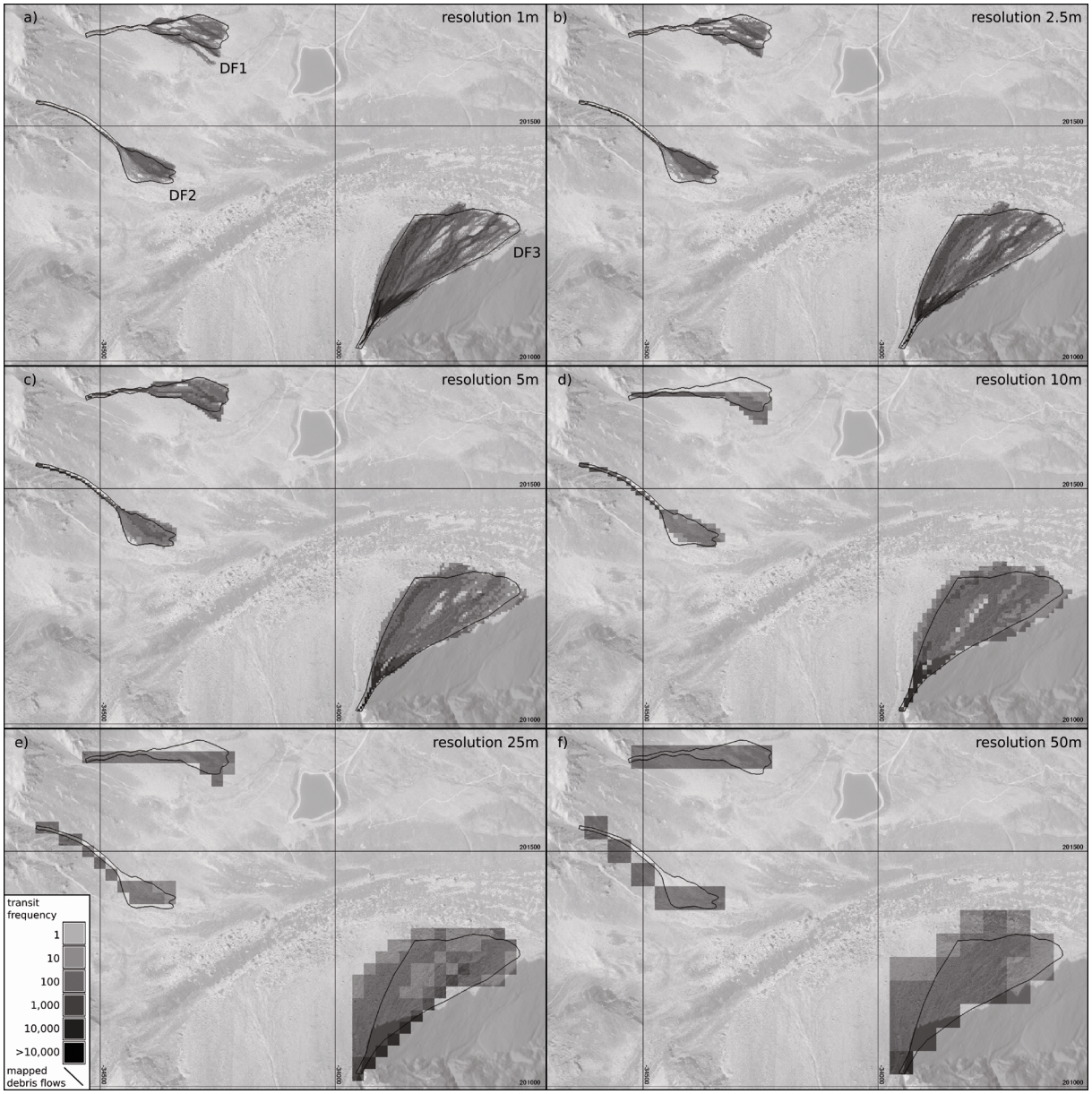


Fig. 2: Simulated process area for debris flows DF1 to DF3 on different grid-cell sizes: a) 1 m, b) 2.5 m, c) 5 m, d) 10 m, e) 25 m, f) 50 m.

The simulation results of debris flow DF4 show the opposite effect: the large debris fan is well reproduced on lower grid-resolutions and is only partially reproduced on the high-resolution DTMs. In this case, the small-scale topography (i.e. the terrain roughness) preserved at small grid-cells sizes prevents the model from reaching all parts of the debris fan. These parts of the debris fan were build up by former events and are currently inactive (see Fig. 1).

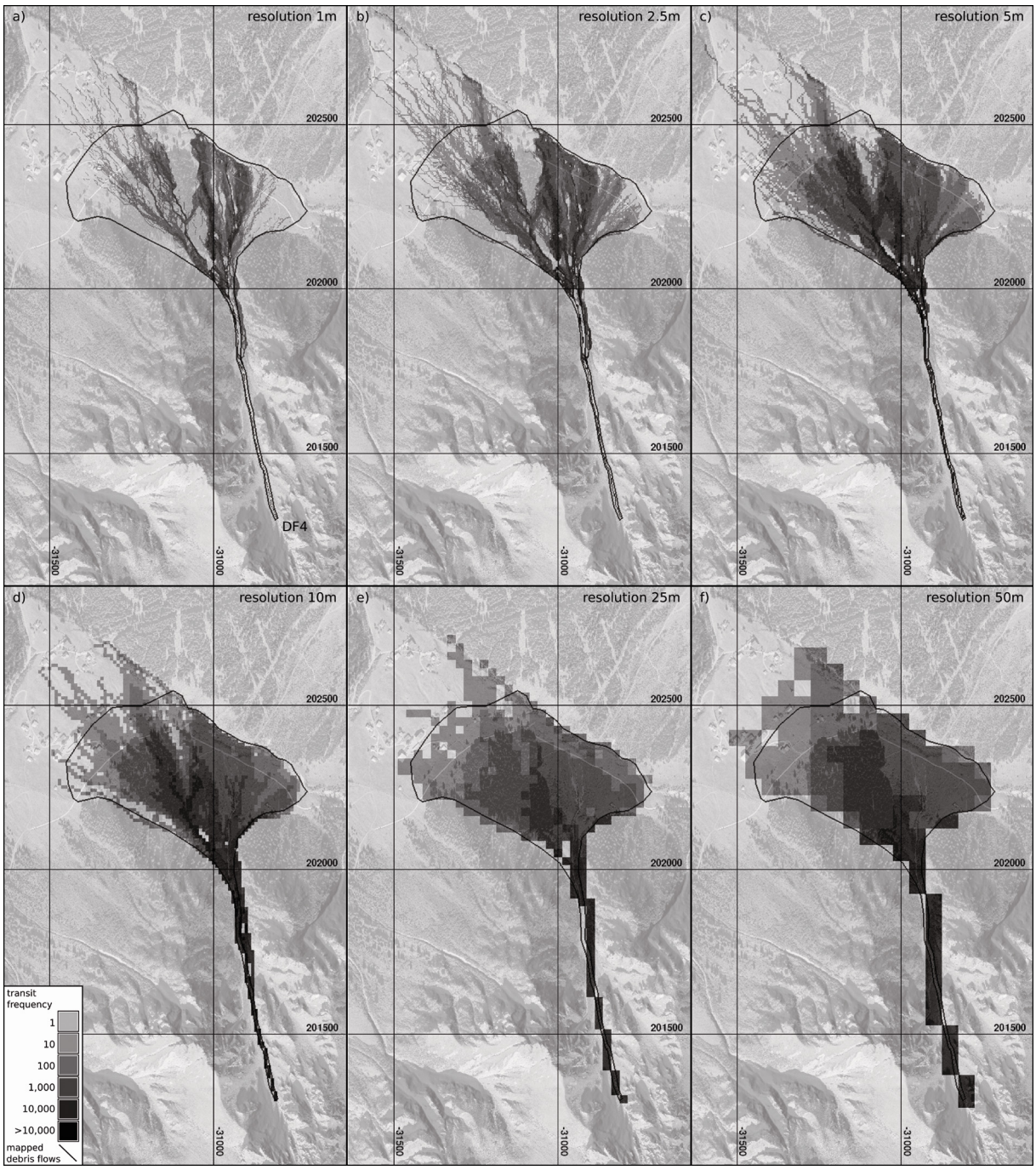


Fig. 3: Simulated process area for debris flow DF4 on different grid-cell sizes: a) 1 m, b) 2.5 m, c) 5 m, d) 10 m, e) 25 m, f) 50 m.

The number of debris flows investigated in this study is too small to analyze the parameter variation statistically. However, the parameter variation with grid-cell size shows general trends (Fig. 4) that allow a qualitative interpretation. The only parameter that shows no considerable trend is the slope threshold, although there is a tendency to smaller values with increasing grid-cell size. This is in accordance with the finding that the average slope is decreasing the greater the grid-cell size (DTM1: 31.3°, DTM10: 30.3°, DTM50: 27.9°). Therefore we tend to attribute the variation to the calibration procedure used. As stated above, in some cases it is necessary to slightly increase the slope threshold to reproduce the lateral flow on the debris fan.

The exponent controlling divergent flow is decreasing with grid-cell size. This can be attributed to the smaller number of grid-cells that form the debris fan at lower grid-resolutions and that need to be reached by the model. This is supported by the observation that a larger exponent is necessary to reproduce a larger debris fan.

The persistence factor shows a decrease with increasing grid-cell size. This is difficult to interpret, because, as already stated, this parameter is kept fixed as long as the whole process area can be reproduced. The higher values on small grid-cell sizes are necessary to retain the flow direction and thus force the model to select grid-cells from the successor set whose probabilities are very small. This is especially true for debris flow DF4, whose process area is only partially reproduced at higher grid-resolutions (see Fig. 3).

To reproduce the mapped process areas on high grid-resolutions, a higher number of model runs is necessary. This can be attributed to the higher number of grid-cells that form the debris fan. In addition, the model needs to select more grid-cells with lower probabilities from the successor set for larger process areas. In case of debris flow DF4 we use the same number of model runs on DTM2.5 and DTM1. At these grid-resolutions the roughness of the terrain is reproduced in such a detail that all potential successors already get selected (i.e. the calculation of a higher number of model runs has no effect on the results).

The friction parameter μ shows a slight increase with grid-cell size (Fig. 4), i.e. the friction must be increased to obtain the same run-out distance. This is in accordance with Gamma (2000), who shows that the slope length L has no effect on the simulated velocities but that the momentum change correction (Equ. 7) yields to lower velocities on rough terrain. At lower grid-resolutions, the small scale topographic features are not preserved and thus the slopes are smoother.

7 DISCUSSION

ALS technology allows a rapid and cost-effective measurement of topography over large areas. The measurements obtained allow to generate high-resolution DTMs of a quality unknown so far. However, this needs several processing steps, which takes considerably time and effort. In this study, a DTM was generated from an already classified point cloud. Still, it was necessary to filter this raw DTM again to eliminate residual misclassification and to smooth out inherent noise. Finally, small surface depressions had to be filled up before the debris flow model could be applied. Therefore, a special algorithm was used, which not only fills the depression but also retains a small slope to assure valid flow directions. Most of the pre- and postprocessing steps require special algorithms that are developed and adopted respectively to be able to deal with such large datasets. Such algorithms are not implemented in standard GIS packages so far.

The simulation results of the debris flow model and the analysis of the parameter settings indicate that the model is not applicable at all scales. The DTM resolution has an effect on the location and the extent of the debris flow deposit, which is especially true for large debris fans. The model was originally developed for natural hazard assessments on a regional scale. This is in accordance with the findings that the model works best on grid-cell sizes ranging from 5 m to 25 m. Although the small slope-type debris flows could be well reproduced using high resolution DTMs, the problems observed with the larger debris flow DF4 still hold for them too. If the DTM resolution reproduces the small-scale topography of the debris fan in great detail, local maxima such as ridges or peaks limit the number of potential successors cells and thus divergent flow. In this case, the model is not able to reproduce the whole process area.

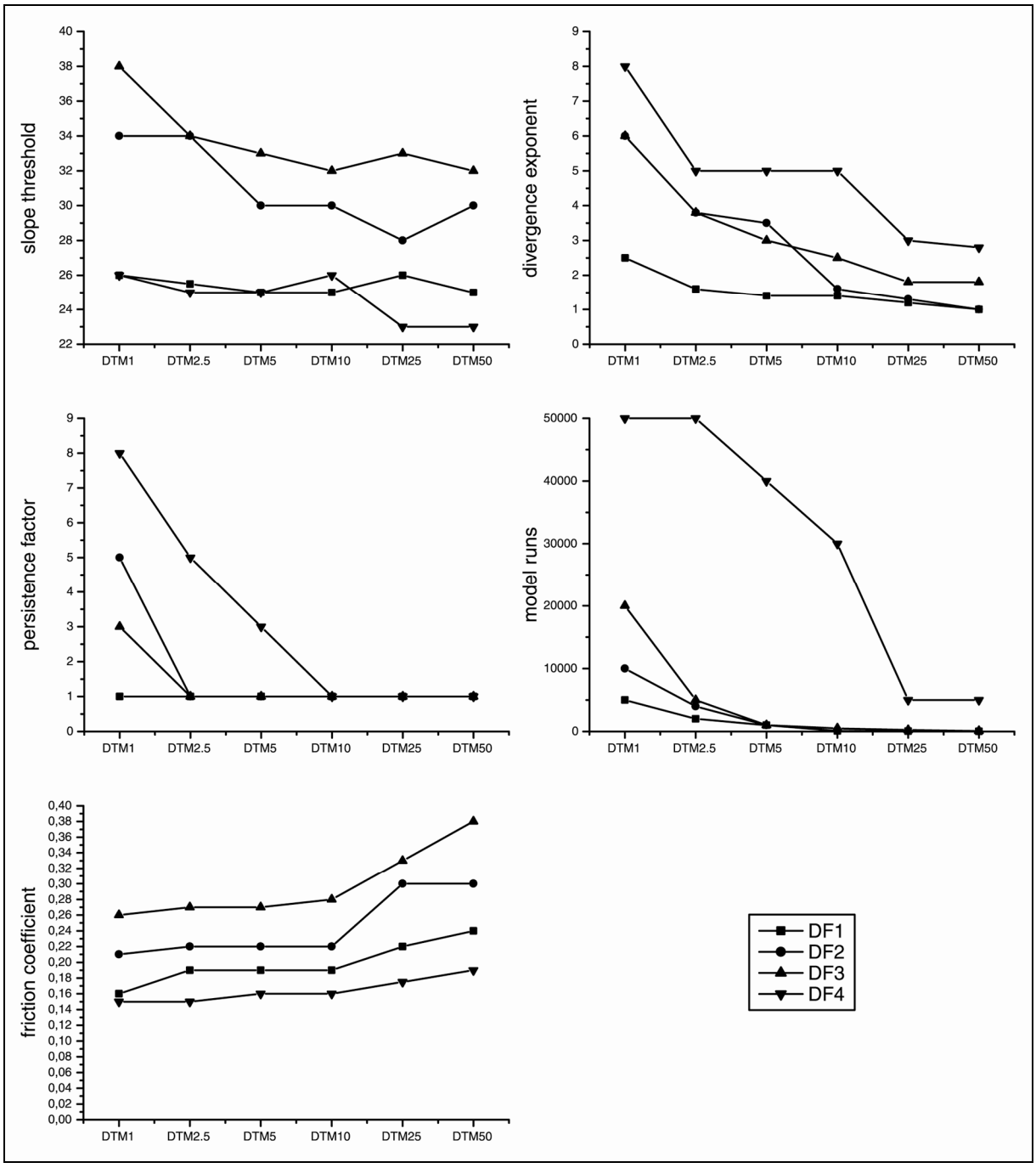


Fig. 4: Variation of model parameter settings with grid-cell size: slope threshold [°], divergence exponent, persistence factor, model runs, and friction parameter μ.

8 CONCLUSION AND OUTLOOK

Topography measurements with ALS technology provide the opportunity to generate derivatives like high-resolution raster DTMs. But the disk space required to store the point cloud and associated attribute data is enormous and still requires special algorithms to be developed for processing. Currently it is difficult to utilize all the information available within the point cloud and one will be restricted to use derived (raster) datasets, usually provided at lower resolutions, which can be processed with standard GIS software. This is especially true when full-waveform laser scanning data, providing additional attributes (WAGNER *et al.* 2004), will become available.

The application of a regional scale debris flow model on high-resolution DTMs showed up several problems. On the one hand it is difficult to reproduce smaller debris flows on low resolutions, on the other hand the small scale topography preserved in the high-resolution terrain models causes problems. Although such details should theoretically result in a better reproduction of the flow paths, local increases in elevation such as older debris flow deposits prevent the model from selecting these grid-cells as potential flow path. This problem is not limited to the debris flow model used in this study, but is inherent to all flow path routing algorithms that disregard the local flow depth when searching for potential flow path cells (which is the majority of flow path routing algorithms published so far, e.g. O'CALLAGHAN & MARK 1984, FREEMAN 1991, TARBOTON 1997). As a consequence such algorithms must be used with care if applied on high-resolution DTMs as long as these problems are not resolved.

Acknowledgments: The authors would like to thank the Federal State of Vorarlberg for providing the remote sensing data sets, especially P. Drexel (Landesvermessungsamt Feldkirch).

REFERENCES

AXELSSON, P. (2000): DEM generation from laser scanner data using adaptive TIN models. – International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences, Vol. 33, B4/1: 110-117.

CAVALLI, M. & L. MARCHI (2008): Characterisation of the surface morphology of an alpine alluvial fan using airborne LiDAR. – Natural Hazards and Earth System Sciences 8: 323-333.

FREEMAN, G.T. (1991): Calculating catchment area with divergent flow based on a regular grid. – Computers and Geosciences 17: 413-422.

GAMMA, P. (2000): *dfwalk* - Ein Murgang-Simulationsprogramm zur Gefahrenzonierung. – Geographica Bernensia G66, University of Bern, Bern, Switzerland.

GEIST, T., HÖFLE, B., RUTZINGER, M., PFEIFER, N. & J. STÖTTER (in press): Laser Scanning - a paradigm change in topographic data acquisition for natural hazard management. – In: VEULLIET, E., STÖTTER, J. & H. WECK-HANNEMANN (Eds.): Sustainable Natural Hazard Management in Alpine Environments. – Springer Verlag.

HÖFLE, B. (2007): Detection and utilization of the information potential of airborne laser scanning point cloud and intensity data by developing a management and analysis system. – Dissertation, Faculty of Geo- and Atmospheric Sciences, University of Innsbruck, Austria.

HOLLAUS, M., WAGNER, W. & K. KRAUS (2005): Airborne laser scanning and usefulness for hydrologic models. – Advances in Geosciences 5: 57-63.

JACOBSEN, K. & P. LOHMANN (2003): Segmented filtering of laser scanner DSMs. - International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences, Vol. 34, 3/W13: 87-93.

KOBLER, A., PFEIFER, N., OGRIBC, P., TODOROVSKI, L., OSTIR, K. & S. DZEROSKI (2007): Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. – Remote Sensing of Environment 108: 9-23.

KRAUS, K. & N. PFEIFER (1998): Determination of terrain models in wooded areas with airborne laser scanner data. – ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 53, No. 4: 193-203.

LOHMANN, P., KOCH, A. & M. SCHAEFFER (2000): Approaches to the filtering of laser scanner data. – International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences, Vol. 33, B3/1: 534-541.

O'CALLAGHAN, J.F. & D.M. MARK (1984): The extraction of drainage networks from digital elevation data. – Computer Vision, Graphics and Image Processing 28: 323-344.

PETRINI-MONTEFERRI, F., DREXEL, P., GEORGES, C. & V. WICHMANN (2008): Aufbau eines Laserscanning-Daten Informationssystems für das Landesvermessungsamt Feldkirch (Land Vorarlberg). – In: STROBL, J., BLASCHKE, T. & G. GRIESEBNER (Eds.): Angewandte Geoinformatik 2008, Beiträge zum 20. AGIT-Symposium, Salzburg, Austria.

PFEIFER, N. & C. BRIESE (2007): Geometrical aspects of airborne laser scanning and terrestrial laser scanning. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36, 3/W52: 311-319.

PFEIFER, N., STADLER, P. & C. BRIESE (2001): Derivation of digital terrain models in the SCOP++ environment. – OEEPE workshop on airborne laserscanning and interferometric SAR for detailed digital elevation models, Stockholm, Sweden.

RIEGER, W., SEEBACHER, M. WÜRLÄNDER, R. & C. BAUERHANSL (2005): Erstellung eines LaserDHM für Vorarlberg 2002-2005. – In: CHESI G. & T. WEINOLD (Eds.) Internationale geodätische Woche Obergurgl 2005, Obergurgl, Austria: 115-124.

SCHMIDT, R., HELLER A. & R. SAILER (2003): Die Eignung verschiedener digitaler Geländemodelle für die dynamische Lawinensimulation mit SAMOS. – In: STROBL, J., BLASCHKE, T. & G. GRIESEBNER (Eds.): Angewandte Geographische Informationsverarbeitung XV, Beiträge zum AGIT-Symposium 2003, Salzburg, Austria: 455-464.

SITHOLE, G. & G. VOSSELMAN (2004): Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. – ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 59, No. 1-2: 85-101.

TARBOTON, D.G. (1997): A new method for the determination of flow directions and upslope areas in grid digital elevation models. – Water Ressources Research, Vol. 33, No. 2: 309-319.

TÓVARI, D. & N. PFEIFER (2005): Segmentation based robust interpolation – a new approach to laser data filtering. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36, 3/W19: 79-84.

VON HANSEN, W. & T. VÖGTLE (1999): Extraktion der Geländeoberfläche aus flugzeuggetragenen Laserscanner-Aufnahmen. – Photogrammetrie Fernerkundung Geoinformation, Vol. 4: 229-236.

VOSSELMAN, G. (2000): Slope based filtering of laser altimetry data. – International Archives of Photogrammetry and Remote Sensing, Vol. 33, B3/2: 935–942.

WAGNER, W., ULLRICH, A., MELZER, T., BRIESE, C. & K. KRAUS (2004): From single-pulse to full-waveform airborne laser scanners: potential and practical challenges. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35, B/3: 201-206.

WALKER, J.P. & G.R. WILLGOOSE (1999): On the effect of digital elevation model accuracy on hydrology and geomophology. – Water Resources Reseach, Vol. 35, No. 7: 2259-2268.

WANG, L. & H. LIU (2006): An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. – International Journal of Geographical Information Science, Vol. 20, No. 2: 193-213.

WICHMANN, V. & M. BECHT (2004): Modellierung geomorphologischer Prozesse zur Abschätzung von Gefahrenpotenzialen. – Zeitschrift für Geomorphologie N. F., Suppl.-Vol. 135: 147-165.

WICHMANN, V. & M. BECHT (2005): Modeling of Geomorphic Processes in an Alpine Catchment. – In: ATKINSON, P.M., FOODY, G.M., DARBY, S.E. & F. WU (Eds.): GeoDynamics: 151-167.

WICHMANN, V. (2006): Modellierung geomorphologischer Prozesse in einem alpinen Einzugsgebiet - Abgrenzung und Klassifizierung der Wirkungsräume von Sturzprozessen und Muren mit einem GIS. – Eichstätter Geographische Arbeiten, Vol. 15, Profil Verlag, Munich/Vienna, pp. 231

ZHANG, K., CHEN, S.-C., WHITMAN, D., SHYU, M.-L., YAN, J. & C. ZHANG (2003): A Progressive Morphological Filter for Removing Nonground Measurements From Airborne LIDAR Data. – IEEE Transactions on Geoscience and Remote Sensing, Vol. 41, No. 4: 872-882.

ZIMMERMANN, M., MANI, P., GAMMA, P., GSTEIGER, P., HEINIGER, O. & G. HUNZIKER (1997): Murganggefahr und Klimaänderung - ein GIS-basierter Ansatz. – Schlussbericht NFP 31, vdf Hochschulverlag AG, Zurich, Switzerland.