CLASSIFICATION SYSTEM OF GIS-OBJECTS USING MULTI-SENSORIAL IMAGERY FOR NEAR-REALTIME DISASTER MANAGEMENT

Daniel Frey and Matthias Butenuth

Remote Sensing Technology Technische Universität München Arcisstr. 21, 80333 München, Germany daniel.frey@bv.tum.de, matthias.butenuth@bv.tum.de

KEY WORDS: System, Classification, Statistics, Multisensor, Integration, GIS, Disaster

ABSTRACT:

In this paper, a near-realtime system for classification of GIS-objects is presented using multi-sensorial imagery. The system provides a framework for the integration of different kinds of imagery as well as any available data sources and spatial knowledge, which contributes information for the classification. The goal of the system is the assessment of infrastructure GIS-objects concerning their functionality. It enables the classification of infrastructure into different states as destroyed or intact after disasters such as floodings or earthquakes. The automatic approach generates an up-to-date map in order to support first aid in crisis scenarios. Probabilities are derived from the different input data using methods such as multispectral classification and fuzzy membership functions. The main core of the system is the combination of the probabilities to classify the individual GIS-object. The system can be run in a fully automatic or semi-automatic mode, where a human operator can edit intermediate results to ensure the required quality of the final results. In this paper, the performance of the system is demonstrated assessing road objects concerning their trafficability after flooding. By means of two test scenarios the efficiency and reliability of the system is shown. Concluding remarks are given at the end to point out further investigations.

1 INTRODUCTION

A significant increase of natural disasters such as floodings and earthquakes has been observed over the past decades (Kundzewicz et al., 2005). There is no doubt that the disasters' impact on the population has dramatically increased due to the growth of population and material assets. The regrettable death of people is accompanied by heavy economic damage, which leads to a longterm backslide of the regions hit by the disaster. This situation calls for the development of integrated strategies for preparedness and prevention of hazards, fast reaction in case of disasters, as well as damage documentation, planning and rebuilding of infrastructure after disasters. It is widely accepted in the scientific community that remote sensing can contribute significantly to all these components in different ways, in particular, due to the large coverage of remotely sensed imagery and its global availability.

However, time is the overall dominating factor once a disaster hits a particular region to support the fast reaction. This becomes manifest in several aspects: firstly, available satellites have to be selected and commanded immediately. Secondly, the acquired raw data has to be processed with specific signal processing algorithms to generate images suitable for interpretation, particularly for Synthetic Aperture Radar (SAR) images. Thirdly, the interpretation of multi-sensorial images, extraction of geometrically precise and semantically correct information as well as the production of (digital) maps need to be conducted in shortest timeframes to support crises management groups. While the first two aspects are strongly related to the optimization of communication processes and hardware capabilities, at least to a large extend, further research is needed concerning the third aspect: the fast, integrated, and geometrically and semantically correct interpretation of multi-sensorial images.

Remote sensing data was already used in order to monitor natural disasters in the year 1969 (Milfred et al., 1969). Particularly, in the case of flooding a lot of studies are carried out to infer information as flood masks from remote sensing data (Sanyal and

Lu, 2004). The flooded areas can be derived from optical images (Van der Sande et al., 2003) as well as from radar images (Martinis et al., 2009) via classification approaches. Zwenzner (Zwenzner and Vogt, 2008) estimates further flood parameter as water depth using flood masks and a very high resolution digital elevation model. Combining this results with GIS data leads to an additional benefit of information and simplifies the decision making (Brivio et al., 2002, Townsend and Walsh, 1998). The combination of the GIS and remote sensing data is often carried out by overlaying the different data sources. But, there are only few approaches which use the raster data from imagery to assess the given GIS data. In (Gerke et al., 2004, Gerke and Heipke, 2008) an approach for automatic quality assessment of existing geospatial linear objects is presented. The objects are assessed using automatically extracted roads from the images (Wiedemann and Ebner, 2000, Hinz and Wiedemann, 2004). However, in case of natural disasters the original roads are destroyed or occluded and, therefore, it is not possible to extract them using the original methods. Hence, new approaches have to be developed which assesses damaged and occluded objects, too. The integration and exploitation of different data sources, e.g. vector and image data, was discussed in several other contributions (Baltsavias, 2004, Butenuth et al., 2007). However, there is a lack of methods which assess the GIS data concerning its functionality using imagery (Morain and Kraft, 2003).

In this paper, a classification system using remote sensing data and additionally available information is developed to assess GISobjects. The main goal of the system is the automatic classification and evaluation of infrastructure objects, for example the trafficability of the road network after natural disasters. However, the presented system can be transferred to other scenarios, such as changes in vegetation, because its design is modular. A focus is the integrated utilization of any available information, which is important to ease and speed up the classification process with the aim to derive complete and reliable results (Reinartz et al., 2003, Frey and Butenuth, 2009). In comparison to the manual interpretation of images the presented systems is very efficient, which is essential in crisis scenarios. Depending on the type and complexity of the input data, the system can be run in a fully automatic or semi-automatic mode, where a human operator can edit intermediate results to ensure the required quality of the final results.

Section 2 describes the generic near-realtime classification system with the objective to classify and evaluate objects using remote sensing and other available data. In Section 3 the system is applied to road objects in case of natural disasters. Two test scenarios of flooded areas are used to verify the system. By means of manually generated reference data, the applicability and efficiency of the system is evaluated in Section 4. Finally, further investigations in future work are pointed out.

2 CLASSIFICATION SYSTEM

The goal of the developed classification system is the assessment of GIS-objects using up-to-date remote sensing data. The system is designed in a general and modular way to provide the opportunity to label GIS-objects into different states. Typical states describe the functionality of infrastructure objects as roads or buildings. The generic system embeds different kinds of image data: multi-sensor as well as multi-temporal data. Additionally, any kinds of available data sources and spatial knowledge, which contributes information for the assessment, can be embedded. Typical examples are digital elevation models (DEM) and further GIS information, e.g. land cover or waterways. The minimum requirement of the system are the objects to be assessed and one up-to-date image which provides the information for the assessment.



Figure 1: Classification system

The classification system depicted in Figure 1 can be subdivided into different components. Starting point are the GIS-objects to be assessed. Secondly, the input data as imagery or digital elevation models which contribute the information for the assessment. In the following this information is called *data*. Thirdly, the classification system by itself and, finally, a resulting up-to-date map. The fusion of multi-sensor images is an important issue, because the corregistration between optical and radar images is still a current research topic (Pohl and Van Genderen, 1998). Methods such as mutual information can be applied for the system (Inglada and Giros, 2004). The system has to deal with multi-temporal images having the possibility to derive important information on time. This leads to an even more complex corregistration process. Change detection algorithms can provide information about the variation of assessed objects. In this article the temporal factor is neglected, but will be an essential part in future research.

The main core of the system represents the classification. The goal is to classify each object into a different state S_i . For each object probabilities are derived belonging to a certain state. The methods estimating the probabilities depends on the data: typical examples are multispectral classification or fuzzy membership functions (Figure 2).



Figure 2: Derivation of probabilities from data using various methods

Beside the derivation of the individual probabilities from each data source the combination plays a decisive role:

$$p_{S_1} = p_{d_1,S_1} \otimes p_{d_2,S_1} \otimes \cdots \otimes p_{d_n,S_1}$$

$$p_{S_2} = p_{d_1,S_2} \otimes p_{d_2,S_2} \otimes \cdots \otimes p_{d_n,S_2}$$

$$\vdots$$

$$p_{S_i} = p_{d_1,S_i} \otimes p_{d_2,S_i} \otimes \cdots \otimes p_{d_n,S_i}.$$
(1)

The variable p_{d_n,S_i} denotes the probability that the state S_i occurs given data d_n . The indices i and n describe the number of available states and data, respectively. The result p_{S_i} shows the probability that a GIS-object belongs to the state S_i . For each type of data weights w_n can be introduced in order to cope with the different influence of information content. Hence, Equation 1 for one state i leads to:

$$p_{S_i} = w_1 \cdot p_{d_1, S_i} \otimes \dots \otimes w_n \cdot p_{d_n, S_i}.$$
 (2)

Finally, the object is assigned to the state S_i with the largest probability p_{S_i} . A basic characteristic of the whole system is the combination at the probability level in order to remain flexible concerning the available data.

3 MODEL FOR ROAD OBJECTS

After describing the generic system, a model is shown which assesses linear objects as roads after flooding. However, this model is transferable to other linear objects like railways and further natural disasters such as avalanches, landslides or earthquakes. In case of natural disasters the GIS-object can be divided into the state *intact/usable* or *not intact/destroyed*. Furthermore, a state between these extrema is possible. Hence, a third state *possibly not intact/destroyed* is introduced, if the automatic approach can not provide a reliable decision. In order to assess roads after a flood disaster following states can be used:

- trafficable
- flooded
- possibly flooded

For every available data source the probability for each state has to be derived. The methods which are employed to the different data are shown in the following section.

3.1 Methods

A multispectral classification is accomplished in order to derive different classes from the input imagery. The goal is to assess each linear object individually without taking adjacent linear objects into account, because such kind of topological knowledge about the connectivity of a road network is no more valid in case of road networks hit by a natural disaster. Every linear object is a polyline, which consists of several line segments. A line segment is a straight line, which can be defined with two points. Every line segment is assigned to a class using an segment-based multispectral classification. To this end, a buffer is defined around each line to investigate the radiometric image information. In many cases additional information as the width of the line object can be used in order to generate the size of the buffer region.

For the multispectral classification various classes have to be defined depending on the underlying imagery in order to classify the road segments into the three states *trafficable*, *flooded* and *possibly flooded*. In case of optical imagery the classes road, water, forest and clouds are convenient, because the class road corresponds to the state *trafficable*, the class water to *flooded* and the classes forest and clouds describe occlusions and therefore belong to the state *possibly flooded*. If radar images are available the class clouds can be neglected. Beside the assignment to a class each individual line segment consists of a probability belonging to a class ω_i , which is derived from the k-sigma error ellipsoid. The probability can be formulated as $p_{\omega_i}(\vec{g})$, whereas \vec{g} defines the gray values. The length of the vector is equivalent to the number of channels.

Beside the imagery additional information such as digital elevation models or GIS data can be integrated in the system. The methods to derive probabilities depend on the data. One method are membership functions of fuzzy sets (Zadeh, 1965). Membership functions do not describe the likelihood of some event, but they only characterize a degree of truth in vaguely defined sets. Since it is often difficult to derive sound probabilities from GIS data, membership functions provide an opportunity to infer confidence values. To emphasize the distinction the membership function is labeled as μ instead of p.

The membership functions $\mu_t(a)$, $\mu_f(a)$ are introduced if a digital elevation model is given. The function $\mu_t(a)$ denote the belonging to the state *trafficable t* depending on the altitude *a*. Similarly $\mu_f(a)$ represents the state *flooded f*. Both functions are depicted in Figure 3. There are two thresholds a_1 and a_2 which determine the height of very likely flooded or trafficable areas, respectively. The current water level lies between these thresholds, which can be calculated by

$$a_1 = l_l - b_1 a_2 = l_h + b_2,$$
(3)

in which l_l is the lowest and l_h is the highest water level in the scene. In order to involve variations due to flows and barriers additional buffers b_1, b_2 are added.



Figure 3: Membership functions for flooded roads and trafficable roads derived from DEM

3.2 Combination of Probabilities

The core of the classification system is to combine probabilities resulting from a multispectral classification with the degree of truth of membership functions. In this section, an example is shown which combines the derived probabilities from optical images with membership functions inferred from a digital elevation model. By means of multispectral classification for each class (water w, road r, forest o, cloud c) the corresponding probability p_{ω_i} for $i = \{w, r, o, c\}$ can be derived. On the other side, the membership function provide the degree of truth $\mu_t(a)$ and $\mu_f(a)$. Utilizing the knowledge that roads higher than a_2 are definitely trafficable and roads lower than a_1 are very likely flooded a case differentiation is carried out:

$$\mu_f(\vec{g}, a) = \begin{cases} \mu_f(a) = 1 & a \le a_1 \\ \mu_f(a) \cdot p_{\omega_w}(\vec{g}) & a_1 < a < a_2 \\ \mu_f(a) = 0 & a \ge a_2 \end{cases}$$
(4)

$$\mu_t(\vec{g}, a) = \begin{cases} \mu_t(a) = 0 & a \le a_1 \\ \mu_t(a) \cdot p_{\omega_r}(\vec{g}) & a_1 < a < a_2 \\ \mu_t(a) = 1 & a \ge a_2. \end{cases}$$
(5)

Variable *a* denotes the height of a road object. The road is assigned to the state *flooded* S_F if the degree of truth $\mu_f(\vec{g}, a)$ exceeds an threshold t_1 , which can be pre-estimated via the standard deviation of the likelihood function resulting from the training data for water. The road is assigned to the state *possibly flooded* S_{PF} , if $\mu_f(\vec{g}, a)$ is less than t_1 . The probability $\mu_t(\vec{g}, a)$ is treated in an analogous manner. The road is assigned to the state *trafficable* S_T if $\mu_t(\vec{g}, a)$ exceeds a pre-determined threshold t_2 . Otherwise, the road is again assigned to the state *possibly flooded* S_{PF} . The road segments which are classified as forest ω_o or clouds ω_c are assigned to the states in the following way:

$$\begin{array}{ll} a < a_1 & \Rightarrow \text{ flooded } S_F \\ a_1 < a < a_2 & \Rightarrow \text{ possibly flooded } S_{PF} \\ a > a_2 & \Rightarrow \text{ trafficable } S_T \end{array}$$
(6)

In Figure 4 a schematic overview of the used classification system is depicted. A multispectral classification is carried out to assign the road objects to the different classes. The results of the multispectral classification combined with the membership function leads to the assignment of the road objects to the different states.



Figure 4: Schematic overview of the classification system

4 RESULTS AND EVALUATION

The presented system has been exemplarily tested with two scenarios representing flood disasters. In both cases roads are assessed concerning their trafficability. The first scenario is the Elbe flood in the year 2002 near Dessau, Germany. Three different data sources are used for the assessment: Firstly, an IKONOS-Image with four channels (red, green, blue and infrared), cf. Figure 5. The ground-sampling distance of the panchromatic channel is 1 meter and the color-channels is 4 meter. As second source a digital elevation model with a resolution of 10 meters is used. Finally, the objects to be assessed are taken form the ATKIS (German Official Topographic Cartographic Information System) database. The test scene covers an area of 33 km², which contains 5484 line segments. In the following investigations only the road objects are studied.

The second study area is located in Gloucesterhire Region in Southeast England. In July 2007 the record flood level at Tewkesbury was measured. During the flooding a TerraSAR-X scene in StripMap mode with a spatial resolution of 3 meter was acquired. The polarization is HH, which is more efficient than HV or VV to distinguish flooded areas (Henry et al., 2003). The test scene covers an area of of 9,5 km². Additionally, linear membership functions from the original rivers are derived and an automatically extracted flood mask is used. As GIS-objects 522 roads from OpenStreetMap are assessed.

The test scenarios are very appropriate to test the classification system due to their diverse global context and the different kinds of roads. The roads vary from paths to highways. Both test scenarios are evaluated using manually derived reference data. The availability of reference data describing the real status of roads during the flooding is very difficult caused by the fast changes of the water level and the accessibility of the roads. One possibility is to derive the reference data from the image itself, which is done for the Elbe scenario. This kind of reference data does not describe the ground truth, but the information which is possible to get from the studied image. In the case of the Gloucesterhire scenario high resolution airborne image with a resolution of 20 cm are available. This imagery which was acquired half a day later than the studied TerraSAR-X scene was used to infer the exact ground truth. To draw conclusions from the following results, it is important to consider the kind of used reference data.

The result of the Elbe scene is visualized in Figure 5. The red lines refer to flooded roads, green lines to trafficable roads and the yellow lines point out, that no decision is possible by the automatic system. In Figure 6 a detail of the original IKONOS image and the assessed roads is shown.



Figure 5: Automatic assessment of roads using the classification system: flooded roads (red), trafficable roads (green) and possibly flooded roads (yellow)



Figure 6: Detail of original and assessed IKONOS scene

Comparing the result with the manually generated reference leads to the numerical results shown in Table 1. "Correct assignment" means that the manually generated classification is identical with the automatic approach. In the case of "Manuel control necessary" the automatic approach leads to the state *possibly flooded* whereas the manual classification assigns the line segments to *flooded* or *trafficable*. The other way around denotes the expression "Possibly correct assignment". "Wrong assignment" means that one approach classifies the line segment to *flooded* and the other to *trafficable*. With the current implementation of the system the approach achieves a correct assignment for 78% of the road objects. Only a very small value of false assignments is obtained. This result is deteriorated due to the 5% of "Possibly wrong assignments". Less than 1/5 of all road segments (17%) should be controlled manually in order to reach a correctness of 95%.

Possible assignment	Result
Correct assignment	76.99%
Manual control necessary	17.87%
Possibly correct assignment	4.96%
Wrong assignment	0.18%

Table 1: Results Scenario: Elbe

The results are obtained with the threshold parameters $t_1 = 0.5$ and $t_2 = 0.001$. The variations of the parameters are depicted in Figure 7. The parameters are responsible for the amount of road segments which are assigned to the state *possibly flooded* on condition that they are classified to the classes water or road. The decrease of "Wrong assignment" comes along with the decrease of "Correct assignments" and an increase of manual control.



Figure 7: Results dependent on parameter t_1 and t_2 (red = Wrong assignment, orange = Possibly correct assignment, yellow = Manual control necessary, green = Correct assignment)



Figure 8: Combination of probabilities and impact of the parameter t_1

In Figure 8 the combination of the probabilities $\mu_f(a)$ and $p_{\omega_w}(\vec{g})$ is shown. The grayscale bar indicates the combined probability $\mu_f(\vec{g}, a)$. Every star defines a road segment assigned to the class water by multispectral classification, the color shows the state assigned in the reference. Many road segments which are assigned to the state *trafficable* in the reference are wrongly classified by the system to the class water. The reason is the high standard deviation of the probability densitiy function for the class road

and, therefore, the overlapping of the class road and water. Road segments in urban areas occluded by shadows are responsible for this effect. The threshold t_1 is depicted in blue which devide the assignment of the roads to the state *flooded* and *possible flooded* (Figure 8). Shifting this parameter leads to the results illustrated on the right plot in Figure 7. Furthermore, the improvement of the combined probability is shown in Figure 8. If only one probability is available, the threshold t_1 would be depicted as a straight horizontal or vertical line. The total required time to generate the manual reference is about three hours. Compared to the time needed for the automatic classification (less than one minute) points out the efficiency of the approach.

The results of the second test scenario are depicted in Figure 9. A detail of the original TerraSAR-X scene and the assessed road segments is shown in Figure 10.



Figure 9: Automatic assessment of roads using the classification system: flooded roads (red), trafficable roads (green) and possibly flooded roads (yellow)



Figure 10: Detail of original and assessed TerraSAR-X scene

In the second test scenario the real ground truth is available. Hence, the assignment *possibly flooded* is not existing in the reference data. The comparison with the automatic classification system leads to the result shown in Table 2. After controlling 5% manually, altogether over 86% are correctly assigned. The value of 14% of wrong assignment is caused by mainly two reasons: Firstly, the resolution of the StripMap mode hardly enables to

detect flooded roads in urban areas. Secondly, the geometric accuracy of the used OpenStreeMap road objects are in many cases not accurate enough for a correct assignment.

Possible assignment	Result
Correct assignment	81.22%
Manual control necessary	4.60%
Wrong assignment	14.18%

Table 2: Results Scenario: Gloucesterhire

CONCLUSIONS

This article presents a classification system to assess GIS-objects concerning their functionality. The system is evaluated by means of two test scenarios with the goal to derive the trafficability of roads during a flooding. Both test scenarios show the good performance and especially the efficiency of this approach. In future work, the whole system will be evaluated using real ground truth to identify the reliability in disaster scenarios. Moreover, the additional benefit combining different image data types such as optical and radar will be part of further study. Currently, the combination of the probabilities is accomplished with a simple multiplication. It has to be investigated, if the combination of different probabilities could be realized better using a Dampster-Shafer framework. In addition, future work comprises the development of multi-temporal models to better exploit different image acquisition times including different data types. A further point is the preprocessing of the used GIS-objects to impove the spatial accuracy of the used infrastructure objects.

ACKNOWLEDGEMENTS

This work is part of the IGSSE project "SafeEarth" funded by the Excellence Initiative of the German federal and state governments, and part of the project "DeSecure". The author would like to thank the Federal Agency for Cartography and Geodesy Sachsen-Anhalt to provide the DEM and the ATKIS road data.

REFERENCES

Baltsavias, E., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. ISPRS Journal of Photogrammetry and Remote Sensing 58(3-4), pp. 129–151.

Brivio, P., Colombo, R., Maggi, M. and Tomasoni, R., 2002. Integration of remote sensing data and GIS for accurate mapping of flooded areas. International Journal of Remote Sensing 23(3), pp. 429–441.

Butenuth, M., Gösseln, G., Tiedge, M., Heipke, C., Lipeck, U. and Sester, M., 2007. Integration of heterogeneous geospatial data in a federated database. ISPRS Journal of Photogrammetry and Remote Sensing 62(5), pp. 328–346.

Frey, D. and Butenuth, M., 2009. Analysis of road networks after flood disasters using multi-sensorial remote sensing techniques. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation 18, pp. 69 – 77.

Gerke, M. and Heipke, C., 2008. Image-based quality assessment of road databases. International Journal of Geographical Information Science 22(8), pp. 871–894.

Gerke, M., Butenuth, M., Heipke, C. and Willrich, F., 2004. Graph-supported verification of road databases. ISPRS Journal of Photogrammetry and Remote Sensing 58(3-4), pp. 152 – 165. Henry, J., Chastanet, P., Fellah, K. and Desnos, Y., 2003. EN-VISAT multipolarised ASAR data for flood mapping. Proceedings of Geoscience and Remote Sensing Symposium, IGARSS 2, pp. 1136–1138.

Hinz, S. and Wiedemann, C., 2004. Increasing efficiency of road extraction by self-diagnosis. Photogrammetric Engineering and Remote Sensing 70(12), pp. 1457–1464.

Inglada, J. and Giros, A., 2004. On the possibility of automatic multisensor image registration. IEEE Transactions on Geoscience and Remote Sensing 42(10), pp. 2104–2120.

Kundzewicz, Z., Ulbrich, U., Brücher, T., Graczyk, D., Krüger, A., Leckebusch, G., Menzel, L., Pińskwar, I., Radziejewski, M. and Szwed, M., 2005. Summer floods in central europe - climate change track? Natural Hazards 36(1), pp. 165–189.

Martinis, S., Twele, A. and Voigt, S., 2009. Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. Natural Hazards and Earth System Science 9(2), pp. 303– 314.

Milfred, C., Parker, D. and Lee, G., 1969. Remote sensing for resource management and flood plain delineation. 24th Midwestern States Flood Control and Water Resources Conference.

Morain, S. and Kraft, W., 2003. Transportation lifelines and hazards: Overview of remote sensing products and results. Proceedings of Remote Sensing for Transportation 29, pp. 39 – 46.

Pohl, C. and Van Genderen, J., 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. International Journal of Remote Sensing 19, pp. 823–854.

Reinartz, P., Voigt, S., Peinado, O., Mehl, H. and Schroeder, M., 2003. Remote sensing to support a crisis information system: Mozambique rapid flood mapping system, river elbe flood: Germany 2002. Proceedings of Remote Sensing of Environment pp. 10–14.

Sanyal, J. and Lu, X., 2004. Application of remote sensing in flood management with special reference to monsoon Asia: a review. Natural Hazards 33(2), pp. 283–301.

Townsend, P. and Walsh, S., 1998. Modeling floodplain inundation using an integrated GIS with radar and optical remote sensing. Geomorphology 21(3-4), pp. 295–312.

Van der Sande, C., De Jong, S. and De Roo, A., 2003. A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment. International Journal of Applied Earth Observations and Geoinformation 4(3), pp. 217–229.

Wiedemann, C. and Ebner, H., 2000. Automatic completion and evaluation of road networks. International Archives of Photogrammetry and Remote Sensing 33(B3/2; PART 3), pp. 979– 986.

Zadeh, L., 1965. Fuzzy sets. Information and Control 8(3), pp. 338–353.

Zwenzner, H. and Vogt, S., 2008. Improved estimation of flood parameters by combining space based SAR data with very high resolution digital elevation data. Hydrology and Earth System Sciences Discussions 5(5), pp. 2951 – 2973.