Physics and Chemistry of the Earth 36 (2011) 241–252

Contents lists available at ScienceDirect

Physics and Chemistry of the Earth

journal homepage: www.elsevier.com/locate/pce

Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies

P. Matgen [a,c,*], R. Hostache [a], G. Schumann [b], L. Pfister [a], L. Hoffmann [a], H.H.G. Savenije [c]

[a] Centre de Recherche Public – Gabriel Lippmann, Rue du Brill, L-4422 Belvaux, Luxembourg
[b] School of Geographical Sciences, Bristol BS81SS, United Kingdom
[c] Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, NL-2600 GA Delft, The Netherlands

ARTICLE INFO

Article history:
Received 15 October 2010
Received in revised form 10 December 2010
Accepted 14 December 2010
Available online 21 December 2010

Keywords:
Synthetic aperture radar
Flood mapping
Remote sensing
ENVISAT
RADARSAT

ABSTRACT

This paper aims at contributing to the elaboration of new concepts for an efficient and standardized Synthetic Aperture Radar (SAR) based monitoring of floods. Algorithms that enable an automatic delineation of flooded areas are an essential component of any SAR-based monitoring service but are to date quasi non-existent. Here we propose a hybrid methodology, which combines radiometric thresholding and region growing as an approach enabling the automatic, objective and reliable flood extent extraction from SAR images. The method relies on the calibration of a statistical distribution of 'open water' backscatter values inferred from SAR images of floods. A radiometric thresholding provides the seed region for a subsequent region growing process. Change detection is included as an additional step that limits overdetection of inundated areas. Two variants of the proposed flood extraction algorithm (with and without integration of reference images) are tested against four state-of-the-art benchmark methods. The methods are evaluated through two case studies: the July 2007 flood of the Severn river (UK) and the February 1997 flood of the Red river (US). Our trial cases show that considering a reference pre- or post-flood image gives the same performance as optimized manual approaches. This encouraging result indicates that the proposed method may indeed outperform all manual approaches if no training data are available and the parameters associated with these methods are determined in a non-optimal way. The results further demonstrate the algorithm's potential for accurately processing data from different SAR sensors.

1. Introduction

Given the persistent cloud cover during flooding events and the rapid flood recession in small to medium sized river basins, systematic flood detection with visible satellite imagery is problematic and is realistically feasible only in a limited number of geographical regions. Space-borne microwave remote sensing with its nearly all-weather, day and night capabilities, seems to meet the demand for a worldwide, near real-time flood monitoring system. The use of passive microwave systems, however, is difficult given the large angular beams of such systems (Rees, 2001) resulting in spatial resolutions as large as 20–100 km. Due to the specular backscattering characteristics of active radar pulses on plain water surfaces and the resulting low signal return, the use of SAR data for high-resolution flood mapping is comparatively straightforward. Spatial resolutions down to 1 m can be achieved with available TerraSAR-X, COSMO SkyMed and RADARSAT-2 SpotLight modes. Space-borne SAR data have shown to be efficient tools not only in flood monitoring (Badji and Dautrebande, 1997; Schumann et al., 2009), but also in flood forecasting through near real-time assimilation of remote sensing-derived flood extent data in coupled hydrologic–hydraulic models (Matgen et al., 2010). In recent years, research has been focusing on the development of SAR-based flood extent mapping techniques (e.g. Martinis et al., 2009), on methods for deriving water levels from SAR data fused with digital elevation data (Schumann et al., 2010; Zwenzner and Voigt, 2009) and on ways to integrate such added-value data in prediction models, either in near real-time through assimilation (e.g. Neal et al., 2009; Matgen et al., 2010; Hostache et al., 2010), or *a posteriori* through re-calibration (e.g. Pappenberger et al., 2007; Hostache et al., 2009; Montanari et al., 2009; Di Baldassarre et al., 2009a,b). Schumann et al. (2009) provide a contemporary and comprehensive review of relevant research activities.

Commonly used SAR flood extent mapping techniques include simple visual interpretation (e.g. Oberstadler et al., 1997), supervised classification (e.g. De Roo et al., 1999; Townsend, 2002), image texture algorithms (Schumann et al., 2005), histogram thresholding (Schumann et al., 2010), and various multi-temporal change detection methods (e.g. Calabresi, 1995; Delmeire, 1997; Bazi et al., 2005). Image statistics-based active contour models have been used by Bates et al. (1997), Horritt (1999), and Matgen et al. (2007a).

* Corresponding author at: Centre de Recherche Public – Gabriel Lippmann, Rue du Brill, L-4422 Belvaux, Luxembourg.
E-mail addresses: matgen@lippmann.lu (P. Matgen), guy.schumann@bristol.ac.uk (G. Schumann), H.H.G.Savenije@tudelft.nl (H.H.G. Savenije).

1474-7065/$ - see front matter
doi:10.1016/j.pce.2010.12.009

Many SAR image processing techniques thus exist to derive more or less successfully flood area or extent. However, the parameters of these algorithms are very often determined through a visual inspection of the image histograms and are subsequently fine tuned by the operator based on the subjective impression of the result. Fully automated image processing algorithms are surprisingly scarce. In one of the rare studies applied to SAR data, Martinis et al. (2009) described an automatic split-based thresholding and classification refinement process as a computationally efficient approach that provides reliable results in a rapid mapping context. Schumann et al. (2010) computed a global threshold value from the radiometric histogram using Otsu's method (Otsu, 1979). The method applies a criterion measure to evaluate the between-class variance (i.e. separability) of a threshold at a given level computed from a normalized image histogram.

Depending on the satellite system parameters (i.e. wavelength, polarization, spatial resolution and incidence angle), the accuracy of the SAR-derived flood area can be very different. In addition, the image processing algorithm and its operator impact the result. Environmental factors, most notably the meteorological conditions at the time of the acquisition, the topography, the vegetation and infrastructure in the region of interest, have a more or less significant impact on the accuracy of the result. In the transient shallow water zone between the flooded and the non-flooded part of the floodplain, in particular, the radar signal only gradually increases (with protruding vegetation producing increased signal returns) which complicates the delineation of flooded areas. Also the roughening of water surfaces due to wind or heavy rainfall produces significant pulse returns that render the correct imaging of the flooded area difficult. Extraction of flooded areas within urban settlements has long been thought to be unfeasible due to the double bounce reflection off buildings. However, results obtained with some recently deployed high-resolution SAR satellites show some first promising results in this respect (Mason et al., 2010). This illustrates that accurate flood segmentation remains subject to many uncertainties. In order to account for the uncertainty associated with the image processing algorithms Di Baldassarre et al. (2009b) apply five different procedures based on visual interpretation, histogram thresholding, active contour modelling, image texture variance and Euclidean distance (e.g. Kokare et al., 2003). They overlay the five flood extent maps to generate a 'possibility of inundation' map.

Considering the numerous sources of uncertainty, one may argue that it is important and necessary to bring forward automatic procedures that ensure a maximum of objectivity and traceability with respect to SAR-derived flooded areas. The brief review of the current status of SAR-based flood monitoring methods illustrates that further improvements in terms of reliability and repeatability of data extraction from SAR imagery are required if current and future SAR satellite missions are to be routinely used for flooding observations. Moreover, in a crisis management context, it is known that the value of remote sensing information decreases rapidly (Matgen et al., 2007b). Hence, the time delay between image acquisition and distribution of flood information needs to be substantially shortened. Finally, it has to be mentioned that temporal imaging frequencies associated with current radar satellites, which can be up to 35 days, hamper the global monitoring of floods with SAR.

In order to ensure a high temporal and spatial coverage, efficient flood observation services need to make use of and combine data sets stemming from different sensors. Automatic image processing procedures can help to overcome these problems. It is the objective of this study to contribute to these developments by proposing new ways to efficiently process SAR data for enhanced flood monitoring. In particular, we aim at developing a fully automatic image classification method based on image statistics that can be applied to all existing SAR data sets.

2. Methodology

Here we propose a methodology for the automatic and objective flood extent extraction from SAR images. The proposed method is composed of three processing steps:

- estimation of the statistical distribution of 'water' backscatter values from a SAR flood image;
- radiometric thresholding of the SAR image in order to extract the core of the water bodies;
- region growing to extract all water bodies.

Another optional processing step may be included:

- change detection with respect to a non-flood image (if available) is applied in order to remove areas with non-significant changes between the flood and non-flood acquisition.

The individual processing steps are explained in detail hereafter.

2.1. Estimation of open water backscatter statistical distribution

The aim of the first processing step is to estimate the probability density function of water backscatter on SAR flood images. Ulaby et al. (1986) have shown that the backscatter variability on a homogeneous surface is mainly due to speckle and that the probability density function of its backscatter is of gamma type (1). Since open water may be considered as a homogeneous surface acting as a specular reflector with low backscatter values, it can be assumed that the statistical distribution of the backscatter values follows a gamma distribution.

The gamma probability density function f has two parameters k and θ as shown in equation:

$$f(\sigma^0/k,\theta) = \frac{(\sigma^0 - \sigma^0_1)^{k-1}}{\theta^k \Gamma(k)} \cdot e^{-\frac{(\sigma^0-\sigma^0_1)}{\theta}} \qquad (1)$$

Since the gamma distribution can only be computed for positive values, in (1) the backscatter values need to be shifted in order to be positive for the entire range of observed values. As a consequence, σ^0_1 is the minimum backscatter value in the SAR image.

To estimate k and θ for the SAR image histogram under investigation, we apply the non-linear fitting algorithm proposed by Seber and Wild (2003). Moreover, to facilitate the fitting procedure, we propose to make use of the formula of the gamma distribution mode. The latter is given by (2), when $k \geqslant 1$:

$$\sigma^0_m = (k-1)\cdot\theta + \sigma^0_1 \qquad (2)$$

For a given value of σ^0_m and according to (2), only the k parameter has to be optimized for determining $f_{\sigma^0_m}$. Eq. (1) thus becomes:

$$f_{\sigma^0_m}(\sigma^0/k) = \frac{(\sigma^0 - \sigma^0_1)^{k-1}}{\left(\frac{(\sigma^0_m-\sigma^0_1)}{k-1}\right)^k \Gamma(k)} e^{-\frac{(\sigma^0-\sigma^0_1)\cdot(k-1)}{(\sigma^0_m-\sigma^0_1)}} \qquad (3)$$

In the backscatter diagram, due to the overlapping pdf of the dry land, only the values below σ^0_m represent the pdf of the water backscatter. Hence, for all plausible values of σ^0_m, a non-linear fit is accomplished and the root mean squared error (RMSE) of the theoretical density function f with respect to the empirical density function h (image histogram) is calculated for all the backscatter values lower than σ^0_m. Finally, σ^0_m and the associated k value providing the lowest RMSE are considered as optimal (see Fig. 1). The tested values of σ^0_m have to be representative of possible open water backscattering values. Although these values depend on the