Remote Sensing of Environment 124 (2012) 705–716

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Automatic near real-time selection of flood water levels from high resolution Synthetic Aperture Radar images for assimilation into hydraulic models: A case study

D.C. Mason [a,*], G.J.-P. Schumann [b], J.C. Neal [b], J. Garcia-Pintado [a], P.D. Bates [b]

[a] National Centre for Earth Observation, University of Reading, Harry Pitt Building, 3 Earley Gate, Whiteknights, Reading RG6 6AL, UK
[b] School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK

ARTICLE INFO

Article history:
Received 16 January 2012
Received in revised form 19 June 2012
Accepted 23 June 2012
Available online 20 July 2012

Keywords:
Flood detection
Synthetic aperture radar
Water level
Assimilation
Hydraulic modelling

ABSTRACT

Flood extents caused by fluvial floods in urban and rural areas may be predicted by hydraulic models. Assimilation may be used to correct the model state and improve the estimates of the model parameters or external forcing. One common observation assimilated is the water level at various points along the modelled reach. Distributed water levels may be estimated indirectly along the flood extents in Synthetic Aperture Radar (SAR) images by intersecting the extents with the floodplain topography. It is necessary to select a subset of levels for assimilation because adjacent levels along the flood extent will be strongly correlated. A method for selecting such a subset automatically and in near real-time is described, which would allow the SAR water levels to be used in a forecasting model. The method first selects candidate waterline points in flooded rural areas having low slope. The waterline levels and positions are corrected for the effects of double reflections between the water surface and emergent vegetation at the flood edge. Waterline points are also selected in flooded urban areas away from radar shadow and layover caused by buildings, with levels similar to those in adjacent rural areas. The resulting points are thinned to reduce spatial autocorrelation using a top-down clustering approach. The method was developed using a TerraSAR-X image from a particular case study involving urban and rural flooding. The waterline points extracted proved to be spatially uncorrelated, with levels reasonably similar to those determined manually from aerial photographs, and in good agreement with those of nearby gauges.

1. Introduction

Flood extents caused by fluvial floods in urban and rural areas may be predicted by hydraulic models, given knowledge of the topography of the floodplain and channel together with other boundary conditions that may include the input flow rate at the upstream boundary of the reach being modelled and the water stage at the downstream boundary. Uncertainty in the flood extents predicted may be reduced by using data assimilation to combine the model state variables with observations. Assimilation may be used to correct the model state and to improve the estimates of the model parameters (e.g. channel friction) or external forcing (e.g. input flow rate).

One common observation that may be assimilated is the water level at various points along the modelled reach. Water levels may be obtained from river gauges, and assimilation of gauge water levels into models has been considered by Romanowicz et al. (2006) and Neal et al. (2007). In the UK as in many other places, a difficulty is that gauges are typically sited only every 20kms or so, thus giving little information on the spatial variations in the flood level, which may be particularly important in urban areas. Much more spatial information is contained in the flood extents captured in satellite SAR images. SARs are generally used for flood detection rather than visible-band sensors because of their all-weather day-night capability. Spatially distributed water levels may be estimated indirectly along the flood extents in SAR images by intersecting the extents with a floodplain Digital Elevation Model (DEM) (Horritt et al., 2003; Hostache et al., 2009; Lane et al., 2003; Raclot, 2006; Schumann et al., 2007). Assimilation of water levels derived from SAR images of flood extent into hydraulic models has been investigated by Matgen et al. (2007, 2010), Neal et al. (2009) and Giustarini et al. (2011).

Given that 50% of the world's rivers contain no gauges, and that the number that exist is actually declining (Vorosmarty et al., 1996), a further advantage of measuring water levels from SAR flood extents is that the method will work in un-gauged catchments. Direct space-borne measurement of surface water level has been made in the past by the Shuttle Radar Topography Mission (SRTM) (Alsdorf et al., 2007), ICESAT (Frappart et al., 2006) and altimeters such as RA-2 on Envisat, and can currently be made by altimeters such as Poseidon 2 on JASON-1, though the altimeter footprints are such that they are limited to level measurement in rivers ~1 km wide. In the future, NASA's Surface Water and Ocean Topography (SWOT) Mission will use K_a-band radar interferometry to measure surface water levels to 10 cm accuracy on smaller rivers ~100 m wide such as are found in the UK when in flood (Biancarmaria

* Corresponding author. Tel.: +44 118 378 8743; fax: +44 118 378 6413.
E-mail address: d.c.mason@reading.ac.uk (D.C. Mason).

0034-4257/$ – see front matter
doi:10.1016/j.rse.2012.06.017

et al., 2010). Assimilation of simulated SWOT water levels into hydraulic models has been considered by Andreadis et al. (2007) and Biancarmaria et al. (2011). As SWOT is not scheduled for launch until 2020 and will not measure levels for floods less than 100 m wide, the water levels from SAR flood boundaries should continue to be an important source of data for assimilation into models, especially in the near future. It is worth noting that the water levels used in conjunction with the hydraulic model/assimilation system provide an indirect method of measuring river discharge from space.

Although models run in hindcasting mode can provide useful information for minimising the effects of future floods, the ultimate goal must be to use SAR water levels in a forecasting model, which means that they have to be estimated in near real-time. It might be questioned whether it is possible, having acquired a raw SAR image, to perform the processing required to extract a set of water levels in near real-time, given the substantial number of tasks involved. It is necessary to download the image to the ground station, process the raw SAR data to a multi-look SAR image, perform automatic geo-registration using the spacecraft orbit parameters, extract the flood extent from the image automatically, and select a distributed subset of water levels for assimilation. It appears that there are reasons for optimism on this front. ESA has already developed the FAIRE system for ASAR data, which while Envisat was functioning was able to provide processed geo-registered ASAR images only 3 hours after acquisition of the raw data (Cossu et al., 2009). While such systems still have to be developed for newer high resolution SARs such as TerraSAR-X and COSMO-SkyMed, they do at least appear technically feasible. In addition, algorithms have been developed for extracting a flood extent from a SAR image automatically and in near real-time, for flooding in rural areas by Martinis et al. (2009, 2011), and in both urban and rural areas by Mason et al. (2012).

It would be useful to complete the chain of automation by developing an automatic near real-time method of selecting a subset of water levels from a SAR flood extent (Schumann et al., 2011). Assimilation techniques such as the Ensemble Kalman Filter (EnKF) assimilate water levels from a subset of points along a flood extent by generating an ensemble of model runs in which the levels are varied about their observed values by an amount governed by their variance. It is necessary to select a subset of levels because adjacent levels along the flood extent will be strongly correlated and add little new information, while a large number of levels will increase the computational cost unnecessarily. The subset of points selected should be at positions at which the water level can be accurately determined, with the points distributed uniformly over the flood extent, sufficiently sparsely that adjacent water levels are spatially uncorrelated. This could be viewed as an extension of an automatic near real-time algorithm for SAR flood extent delineation. Without such an algorithm, it is not possible to perform near real-time assimilation of SAR-derived flood water levels into a flood forecasting model. The objective of this paper is to develop and test a suitable algorithm satisfying the above requirements.

2. Study area and data set

In common with a number of previous studies, the data set used for this study was acquired during the 1-in-150-year flood that took place on the lower Severn around Tewkesbury, U.K., in July 2007 (Mason et al., 2010; Schumann et al., 2011). This resulted in substantial flooding of urban and rural areas, about 1500 homes in Tewkesbury being flooded. Tewkesbury lies at the confluence of the Severn, flowing in from the northwest, and the Avon, flowing in from the northeast. The peak of the flood occurred on July 22, and the river did not return to bank-full until July 31. On July 25, TerraSAR-X acquired a 3 m-resolution StripMap image of the region (Fig. 1), showing considerable detail in the flooded urban areas (Fig. 2). The TerraSAR-X incidence angle was 24°, and the image was HH polarisation multi-look ground range spatially enhanced. At the time of overpass, there was relatively low wind speed and no rain. Aerial photos of the flooding were acquired on July 24 and 27, and these were combined to validate the flood extent and candidate water level points extracted from the TerraSAR-X image (Mason et al., 2010). The data set also included airborne scanning laser altimetry (LiDAR) data (2 m resolution, 0.1 m height accuracy) of the un-flooded area, with coincident LiDAR and aerial photography covering the two regions identified in Fig. 1. Rectangular region A covers the Tewkesbury urban area (2.6×2 km) (Fig. 2), while region B covers a larger more rural area along the Severn (with north-south extent 12.3 km, east-west extent 6 km). The TerraSAR-X and LiDAR data in region A were re-sampled to 1 m pixel size to maintain resolution in the urban flood detection procedure (Mason et al., 2012), while the data in region B were sampled at a lower resolution (2.5 m pixel size).

3. Flood extent extraction algorithm

The input to the method for selecting a subset of candidate water levels is a flood extent extracted from a high resolution SAR image. Although it would be possible to detect candidate waterline points in the image directly, there are significant advantages in selecting these from the waterline of a flood extent extracted using a sophisticated algorithm based on object segmentation and classification, which takes into account, for example, object heights as well as SAR backscatter, and the presence of radar shadow and layover in urban areas. Previous work has involved the development of such an algorithm for the extraction of flood extent in both urban and rural areas from a high resolution SAR image automatically and in near real-time. This is described in (Mason et al., 2012) and only a summary is given here.

The algorithm first detects the flood in the rural areas. Instead of using per-pixel classification, the image is segmented into homogeneous regions, which are then classified on the basis of their spectral, textural, shape and contextual features. Classification is performed by assigning all segmented regions with mean SAR backscatter less than a threshold to the 'flood' class. To determine the threshold, training regions for 'flood' are automatically selected from regions giving no return in the LiDAR data (i.e. water that has acted as a specular reflector), and for 'non-flood' from un-shadowed areas well above the flood level. The initial segmentation is refined using a variety of rules e.g. flood regions having mean heights significantly above the local flood height are reclassified as non-flood.

A simpler region-growing technique is used in the urban areas, guided by knowledge of the local waterline heights in adjacent rural areas. A SAR simulator is used in conjunction with LiDAR data to estimate regions of the image in which water would not be visible due to shadow or layover caused by buildings and taller vegetation. A set of seed pixels having backscatter less than the threshold, and heights less than or similar to the adjacent rural waterline heights, is identified. Seed pixels are clustered together provided that they are close to other seeds. Regions of shadow and layover are masked out in the processing.

The algorithm was developed using the TerraSAR-X image and associated data acquired for the Tewkesbury 2007 flood. The algorithm proved capable of detecting flooding in rural areas using TerraSAR-X with good accuracy, classifying 89% of flooded pixels correctly, with an associated false positive rate of 6%. Of the urban water pixels visible to TerraSAR-X, 75% were correctly detected, with a false positive rate of 24%. Fig. 3 shows the flood extents extracted in urban and rural areas.

4. Method of candidate water level selection

4.1. Overview

The method consists of five stages, as shown in Fig. 4:

(a) Candidate waterline point selection in rural areas.
(b) Correction of rural waterline positions and levels due to the presence of emergent vegetation at the flood edge.