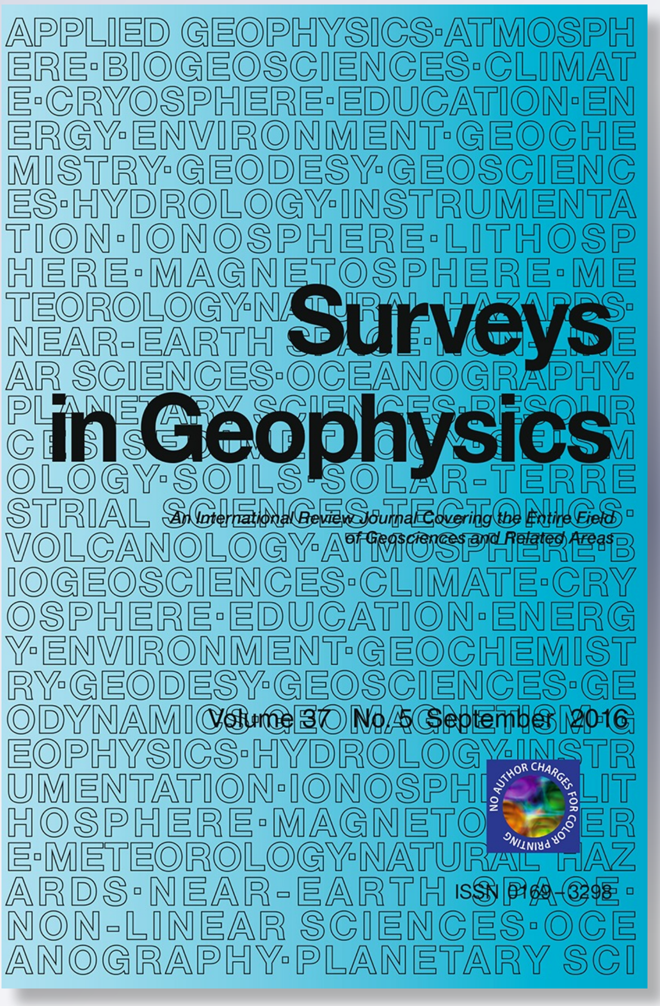
Remote Sensing-Derived Water Extent and Level to Constrain Hydraulic Flood Forecasting Models: Opportunities and Challenges

Stefania Grimaldi, Yuan Li, Valentijn R. N. Pauwels & Jeffrey P. Walker

Surveys in Geophysics
An International Review Journal Covering the Entire Field of Geosciences and Related Areas

ISSN 0169-3298
Volume 37
Number 5

Surv Geophys (2016) 37:977-1034
DOI 10.1007/s10712-016-9378-y



Springer

Surv Geophys (2016) 37:977–1034
DOI 10.1007/s10712-016-9378-y

Remote Sensing-Derived Water Extent and Level to Constrain Hydraulic Flood Forecasting Models: Opportunities and Challenges

Stefania Grimaldi[1] · **Yuan Li**[1] · **Valentijn R. N. Pauwels**[1] · **Jeffrey P. Walker**[1]

Received: 7 April 2016 / Accepted: 13 July 2016 / Published online: 2 August 2016

Abstract Accurate, precise and timely forecasts of flood wave arrival time, depth and velocity at each point of the floodplain are essential to reduce damage and save lives. Current computational capabilities support hydraulic models of increasing complexity over extended catchments. Yet a number of sources of uncertainty (e.g., input and boundary conditions, implementation data) may hinder the delivery of accurate predictions. Field gauging data of water levels and discharge have traditionally been used for hydraulic model calibration, validation and real-time constraint. However, the discrete spatial distribution of field data impedes the testing of the model skill at the two-dimensional scale. The increasing availability of spatially distributed remote sensing (RS) observations of flood extent and water level offers the opportunity for a comprehensive analysis of the predictive capability of hydraulic models. The adequate use of the large amount of information offered by RS observations triggers a series of challenging questions on the resolution, accuracy and frequency of acquisition of RS observations; on RS data processing algorithms; and on calibration, validation and data assimilation protocols. This paper presents a review of the availability of RS observations of flood extent and levels, and their use for calibration, validation and real-time constraint of hydraulic flood forecasting models. A number of conclusions and recommendations for future research are drawn with the aim of harmonising the pace of technological developments and their applications.

Keywords Hydraulic modelling of floods · Remote sensing · Flood extent and level · Data assimilation · Real-time forecast

✉ Stefania Grimaldi
stefania.grimaldi@monash.edu

[1] Department of Civil Engineering, Monash University, Clayton, VIC 3800, Australia

1 Introduction

Operational flood forecasting systems usually consist of a cascaded modelling framework. A hydrologic model computes the amount of water entering the river system. A hydraulic model then uses this input to compute water level and velocity in the river network, and when the storage capacity of the river is exceeded, in the floodplain. A significant amount of work has already been performed to improve hydrologic model forecasts (e.g., Liu et al. 2012; Li et al. 2016). However, this is much less the case for hydraulic models. This discrepancy can be explained by the lack, up until the last few decades, of spatially distributed input and calibration/validation data (e.g., Bates et al. 1997, 2004; Horritt 2000; Werner et al. 2005).

One significant problem for all environmental modelling is that the network of river gauging stations is declining globally (Stokstad 1999; Sivapalan et al. 2003; Schumann et al. 2015). Even though historical and current gauges provide useful data in the developed world, the number of gauges in developing and emerging economies is very small. Furthermore, measurement stations are often installed very far apart and in remote locations, thus making inference of processes and data collection difficult (Schumann et al. 2015). Nevertheless, even when available, measurements from gauging stations tend to consist only of data for a small number of points. For instance, gauging stations in Europe are usually installed every 10–60 km (Neal et al. 2009).

During the last two decades, an increasing awareness of the potential for remote sensing (RS) techniques to monitor floods and thus alleviate some of the field data limitations has led to a consensus among space agencies to strengthen flood monitoring from space. For instance, the International Charter on Space and Major Disasters (www.disasterscharter.org), operational from November 2000, aims to provide a unified system of space data acquisition and delivery to those affected by disasters, such as flooding, via its member space agencies. Flood inundation research has thus recently shifted from being a "data-poor" to a "data-rich" science (e.g., Bates 2004; Di Baldassarre and Uhlenbrook 2012; Schumann et al. 2015). This increased data availability has stimulated more efforts in fostering understanding of the ways in which RS can support flood modelling (Mason et al. 2010).

The remote sensing-derived (RS-D) digital elevation models (DEM), river width and land cover databases are now consolidated and provide indispensable support for the implementation of hydraulic models; reviews on RS data sets for the implementation of hydraulic models have been recently compiled by Di Baldassarre and Uhlenbrook (2012), Musa et al. (2015), Schumann et al. (2015) and Yan et al. (2015).

Moreover, the integration of hydraulic models with RS-D observations of flood extent and water level for model calibration, validation and real-time constraint has the potential to be a powerful approach for augmenting process understanding and prediction (e.g., Bates et al. 1997; Schumann et al. 2015). This concept has only begun to be investigated during the last decade; flood monitoring from space is a recent development with only 15–20 consistent RS-D flood extent and water level data sets currently available worldwide (Bates et al. 2014b). Despite this small number of consistent RS-D flood extent and water level data sets, pioneering studies have shown that such information has the potential to become critically important for the calibration, validation and real-time constraint of hydraulic models for flood forecasting. However, the potential of these RS-D observations to support flood models has not yet been widely explored nor adequately utilised. There is also a need to develop improved frameworks to incorporate such RS-D data into hydraulic models (Di Baldassarre and Uhlenbrook 2012; Schumann et al. 2015).

A review of the progress on integrating RS-D observations of flood extent and water level data and hydraulic models was completed by Schumann et al. (2009a); however, substantial advances have been made in the past few years and merits a new review. More recent reviews by Di Baldassarre et al. (2011) and Yan et al. (2015) only focused on the possibility of using coarse-resolution, low-cost RS data.

This paper provides an updated analysis of the state of the art on the use of coarse-, medium- and high-resolution RS-D observations of flood extent and water level to improve the accuracy of hydraulic models for flood forecasting. Section 2 reviews the present and planned availability of coarse-, medium- and high-resolution sensors for the observation of floods from a remote location. The protocols for the retrieval of maps of flood extent and spatially distributed water levels from these RS observations are then discussed in Sect. 3. Section 4 compares the main features of the traditionally available field data and the recently available RS-D data in order to highlight the relative merits and flaws when used to support the hydraulic modelling of floods. Section 5 presents an overview of the study sites and RS-D data sets discussed in Sect. 6 and in Sect. 7 to investigate the opportunities and challenges of using RS-D observations of flood extent and water level for the calibration and validation (Sect. 6) and real-time constraint (Sect. 7) of hydraulic models for flood forecasting. This review of the fit-for-purpose imaging characteristics and their effective use in calibration/validation and real-time constraint algorithms is meant to set the scene for a strategic and routine use of RS-D observations for flood forecasting in the near future. Technological improvements, new satellites and constellation of satellites are deemed to quickly enhance the global capability of flood monitoring from space in the coming decades. Improved knowledge and computational tools are essential to make full and optimal use of these large and timely data sets. Section 8 draws a number of conclusions and recommendations for future research with the aim of harmonising the pace of technological developments and their application.

2 Remote Sensing-Derived Observations of Flood Extent and Level

Optical, passive microwave and radar instruments are increasingly being used for inundation detection and monitoring. The global availability of RS-D observations of flood extent and level has been previously discussed in Smith (1997), Bates et al. (2014a), Musa et al. (2015) and Schumann et al. (2015). Reviews by Schumann et al. (2009a, 2012), Mason et al. (2010), Di Baldassarre and Uhlenbrook (2012), Schumann and Moller (2015) and Yan et al. (2015) have focused exclusively on radar sensors. This section provides an integrated review of the current and future RS data availability for the observation of flood extent and spatially distributed water level.

2.1 Remote Sensing-Derived Observations of Flood Extent

Satellite and airborne imagery used in floodplain mapping can be characterised by different spatial resolutions, which, relative to the typical length scales of physical flow process during floods, can be broadly defined as: high (1–2 m), fine/medium (10–25 m) and coarse/low (about 100 m) (e.g., Di Baldassarre et al. 2011). These RS observations can be obtained using (1) optical sensors, (2) passive microwave instruments or (3) synthetic aperture radars.

2.1.1 Optical Imagery

The most straightforward mode of acquisition of a flood image is with visible and thermal bands (e.g., Schumann et al. 2009a). Aerial photography provides high-resolution data, and it is considered the most reliable source of remotely sensed flood extent data (e.g., Yu and Lane 2006; Yan et al. 2015). However, its acquisition requires the set-up of dedicated flights and has a high cost. Satellite imagery acquired by multi-purpose space missions offers a relatively cheaper alternative, with spatial resolutions varying from high to low, and the coarse-resolution data often provided by space agencies for free. Examples of high-, medium- and coarse-resolution satellite optical data are provided in Table 1. Flood mapping using optical imagery has proved to be relatively successful (e.g., Wang 2004; Marcus and Fonstad 2008; Faruolo et al. 2009; Proud et al. 2011). Historical databases of observed inundated areas are provided, for instance, by the Dartmouth Flood Observatory, and Geoscience Australia. The Dartmouth Flood Observatory (http://floodobservatory.colorado.edu/, see Adhikari et al. 2010) uses 250-m resolution MODIS and other data, such as the Shuttle Radar Topographic Mission (SRTM) Water Body Data Set, to map flooding in near real time and compile an archive of large floods at global scale. The web service Water Observation from Space (WOfS) provided by Geoscience Australia (http://www.ga.gov.au/scientific-topics/hazards/flood/wofs) displays a statistical analysis of historical surface water observations derived from LANDSAT-5 and LANDSAT-7 satellite imagery for all of Australia from 1987 to present day (Mueller et al. 2016).

However, the systematic application of optical techniques for the monitoring of specific events is hampered by their daylight-only application and their inability to map flooding beneath clouds and vegetation (e.g., Wilson et al. 2007). On the contrary, microwaves penetrate cloud cover, and they are capable to acquire data during day and night and also to some extent map flooded vegetation. Consequently, radar remote sensing offers an opportunity to routinely acquire flood information (e.g., Schumann et al. 2009a).

2.1.2 Passive Microwave Imagery

Passive microwave radiometers measure naturally emitted thermal radiation (i.e. brightness temperature). Due to the different thermal inertia and emission properties of land and water, the observed microwave radiation has a lower brightness temperature for water than for land. This effect can be used for detection of the flooded areas, and detailed techniques have been described by many authors (e.g., Schmugge 1987; Choudhury 1989; Hamilton et al. 1996; Smith 1997). However, the large angular beams of such systems result in spatial resolutions as large as 20–100 km (e.g., Rees 2012), and the interpretation of the wide range of materials with many different emissivities is very difficult (e.g., Papa et al. 2006; Schumann et al. 2009a). The potential of using passive microwave imagery for flood monitoring is thus limited to very large catchments (i.e. catchments having area larger than $\sim 10^3$ km^2). For instance, Sippel et al. (1998) and Jin (1999) mapped the flooded area in the Amazon catchment (Brazil) and in the Wuhan and Wuyuan regions (China), respectively. De Groeve (2010) demonstrated the existence of a good correlation between passive microwave-based flood extents and gauged water levels, when river overtopping occurs in the gently sloping floodplains of the Zambezi River (Namibia). The experimental Global Flood Detection System hosted by the Global Disaster Alert and Coordination System

Table 1 Satellite missions featuring SAR, optical, passive microwave and radar altimeter sensors

SAR

Mission/satellite	Agency	Years of operation	Ground resolution (m)	Revisit interval (days)	Polarisation	Band	Cost/scene [US $] (*)
ERS-1	ESA	1991–2000	25	35	VV (single)	C	Free (**)
ERS-2	ESA	1995–2011	25	35	VV (single)	C	Free (**)
JERS 1	JAXA	1992–1998	18	44	HH (single)	L	Free (**)
RADARSAT-1	CSA	1995–2013	8–100	24	HH (single)	C	1155 (prior to 2008) 2770–3465 (from 2008)
RADARSAT-2	CSA	2007	3–100	24	Full	C	2770–6000
ENVISAT-ASAR	ESA	2002–2012	30–1000	35	Single or dual	C	Free (**)
ALOS-PALSAR	JAXA	2006–2012	10–100	46	Single or dual	L	44/Free
ALOS-PALSAR 2	JAXA	2014	3–100	14	Single or dual	L	1335–4450 *archive* 2670–5785 *new*
COSMO-SkyMed	ASI	2007	15–100	16 (4 to <1 with 4 satellites)	Single or dual	X	730–2655 *archive* 1460–5310 *new*
TerraSAR-X	DLR	2007	1–16	11	Full	X	1216–2985 *archive* 2430–5970 *new*
TANDEM-X	DLR	2010	3	11	Full	X	1216–2985 *archive* 2430–5970 *new*
KOMPSAT-5	KARI	2013	1–20	28	Full	X	800–1650 *archive* 1600–3300 *new*
SENTINEL-1 *A* SENTINEL-1 *B*	ESA	2014 2016	5–100	12 (1–6 with 2 satellites)	Dual	C	Free

OPTICAL (examples)

Mission/satellite	Agency	Years of operation	Ground resolution (m)	Revisit interval (days)	Cost/km^2 [US $]
World-View 2; 3	DIGITAL GLOBE	2009, 2014	1.84; 1.24	1.1	$ 20 (min 25 km^2) *archive* $ 31.5 (min 100 km^2) *new*

Table 1 continued

OPTICAL (examples)

Mission/satellite	Agency	Years of operation	Ground resolution (m)	Revisit interval (days)	Cost/km^2 [US $]
SPOT 5; 6; 7	AIRBUS defence and space	2002, 2012, 2014	10; 6; 6	26–5 (1 with SPOT 6–7 in constellation)	$ 1.20 (min 500 km^2) *archive* $ 1.75 (min 100 km^2) *new*
LANDSAT 5; 7; 8	NASA	1984–2013, 1999, 2013	30	16	Free
MODIS (AQUA; TERRA)	NASA	1999, 2002	500	0.5	Free

PASSIVE MICROWAVE (examples)

Mission/satellite	Agency	Years of operation	Ground resolution (km)	Revisit interval (days)	Cost/km^2 [US $]
TRMM–TMI	NASA/JAXA	1997–2014	5 × 7 up to 72 × 43	1–2	Free
AMSR-E on board of AQUA	NASA	2002	6 × 4 up to 74 × 43	1–2	Free

RADAR ALTIMETERS (examples)

Mission/satellite	Agency	Years of operation	Ground resolution (km)	Inter track distance (km)	Revisit interval (days)	Cost [US $]
ERS-2 RA	ESA	1995–2011	3.4	80	35	Free (**)
ENVISAT RA-2	ESA	2002–2012	3.4	80	35	Free (**)
JASON 1; 2; 3	CNES/NASA	2001–2013, 2008, 2016	1	315	10	Free
SENTINEL 3 *A–B–C*	ESA	2016, 2017, 2020	1.6	104 (*A*)–52 (*A* and *B*)	27	Free

(*) The listed prices are the retail prices and are not the full cost of data acquisition. Prices are estimates based on online sources; actual prices may be lower with certain sales outlets and academic discounts can range from 20 to 30 %. Archive data are generally older than 90 days. SAR cost per scene depends on the imaging mode, and it is generally proportional to the spatial resolution; highest and lowest prices are reported here. Optical imagery cost refers to multispectral images; the minimum area of purchase is specified

(**) A project proposal might be required (http://www.gdacs.org/flooddetection/) uses passive microwave remote sensing (i.e. AMSR-E and TRMM sensors; the details of these products are listed in Table 1) to monitor floods worldwide (Kugler and De Groeve 2007; Revilla-Romero et al. 2015). Despite the successful detection of inundated areas in several kilometres wide river systems, several authors (e.g., Schumann et al. 2009a; Schumann and Moller 2015) highlighted that much high spatial resolution imagery is required to support flood monitoring and modelling in the large worldwide number of small- to medium-sized catchments.

2.1.3 Active Microwave Imagery

Active microwave imagery from synthetic aperture radar (SAR) is deemed to be the only reliable source of information for monitoring floods on rivers less than 1 km in width (e.g., Schumann et al. 2009a; Mason et al. 2010). The applications of SAR remote sensing of flood inundation in a variety of natural and man-made environments have been recently reviewed by Schumann and Moller (2015).

SARs are active systems that emit microwave pulses at an oblique angle towards the target. The amount of microwave energy scattered off an object or feature is a function of its surface texture, shape and dielectric properties (Woodhouse 2005). Open water, in the absence of roughness due to weather causing waves, has a relatively smooth surface which causes radar radiation to be reflected away from the sensor, resulting in low backscatter (Henderson and Lewis 2008). Rough terrestrial land surfaces, by contrast, reflect the energy in many directions, including back towards the sensor, and therefore appear as high backscatter zones. These differences allow flood extent to be mapped using a variety of techniques (Sect. 3.2) to an accuracy of approximately one pixel (Bates et al. 2014b).

Airborne SAR provides high-resolution flood extents and has been used recently to better understand floodplain inundation processes and to compare different hydraulic models with satisfactory results (e.g., Bates et al. 2006; Horritt et al. 2007; Wright et al. 2008). However, satellite SAR images are distributed by space agencies at lower or no cost and are often regarded as an attractive alternative for flood monitoring. Table 1 provides a list of the main past, current and planned satellite missions featuring SAR sensors with high potential for flood propagation and inundation studies. SAR sensors offer a number of acquisition modes, and each acquisition mode is directly linked with the resolution of the resulting image and the size of the scene area covered. For instance, RADARSAT, TerraSAR-X and CosmoSkyMed can acquire in SpotLight, StripMap or ScanSAR mode. The highest resolution but the smallest scene area are achieved with the SpotLight mode; the StripMap mode provides a good trade-off between the size of the scene area and the resolution; the ScanSAR mode is intended for use in applications requiring large area coverage. Technical details on the SAR acquisition modes can be found, for instance, in Lacomme et al. (2001) and Parker (2012).

It must be stated that two difficulties may be encountered using this kind of sensor. The first challenge, discussed here, is related to the technical possibility of acquiring useful data during the flood events. The second challenge, discussed in Sect. 3, is related to the interpretation of the RS images in order to extract information of the flood event.

The strong inverse relationship between spatial resolution and revisit time for satellites (e.g., Di Baldassarre et al. 2011) makes monitoring floods from space in near real time currently only possible through either low-resolution SAR imagery or satellite constellations (e.g., Schumann et al. 2012). Revisit times for SAR imagery with spatial resolutions in the order of 100 m (usually referred to as wide swath mode) are in the order of 3 days, and the data can be obtained within 24 h at relatively low cost (e.g., Schumann et al. 2012).

In smaller basins with shorter flood wave travel times, the probability of imaging a flood decreases proportionately and the available acquisitions from coarse-, medium- or fine-resolution sensors become increasingly opportunistic. Moreover, when monitoring and modelling urban areas, very fine spatial resolution (lower than 5 m) imagery is necessary (Schumann et al. 2012), despite not being often sufficient due to the uncertainties in water surface identification in urban areas (see Sect. 3.2 for a detailed discussion). Constellations of satellites are likely to be the only way to achieve a suitable combination of resolution and revisit frequency (García-Pintado et al. 2013). For instance, the COSMO-SkyMed constellation is composed by four satellites in a Sun-synchronous orbit, providing a 3-m image sequence with a time from request to acquisition of the first image of 26–50 h, and then subsequent images at 12- and 24-h revisit times. An example of flood monitoring using multi-temporal COSMO-SkyMed data has been provided by Pulvirenti et al. (2011a). Further, a new constellation of SAR satellites will be shortly available. The ESA mission Sentinel-1 is composed of two satellites, Sentinel-1A (launched in 2014) and Sentinel-1B (launched in April 2016), operating in four imaging modes with resolution down to 5 m and coverage up to 400 km. The orbit configuration optimises coverage, offering a global revisit time of just 6 days. At the equator, however, the repeat frequency will be just 3 days and less than 1 day over the Arctic. Europe, Canada and main designed shipping routes will be covered in less than 3 days. Product delivery times of 24 down to 3 h have been planned for two types of data sets: a predefined fixed large data collection for specific applications and a flexible, on-demand data collection covering last minute requirements, like observations over emergency events. Experience of fast product delivery has been previously provided by the ESA Grid Processing On-Demand (GPOD) Fast Access to Imagery for Rapid mapping Exploitation (FAIRE) system (http://gpod.eo.esa.int) which allowed ENVISAT-ASAR end ERS images to be available 3 h after acquisition for emergency responses applications such as the International Charter Space and Major Disasters. Based on these premises, as observed by Schumann et al. (2012), the rapid delivery of fine-resolution image is technically feasible and might be a common form of dissemination in the near future.

2.2 Remote Sensing-Derived Observations of Flood Level

Flood level can be obtained through remote sensing in a direct or indirect manner.

Radar altimeters, LiDAR, SAR Interferometry or wide swath altimetry allows the direct observation of flood level. The main features and applications of these direct monitoring techniques are described here. The indirect retrieval of water levels from RS data can be performed by intersecting RS-D inundation extent maps with digital terrain models; in particular, the technical details of the flood level retrieval from SAR images are discussed in Sect. 3.3.

Radar altimeters (RALT) emit a short, nadir-directed radar pulse to the Earth's surface. The two-way return time is used to calculate the distance between the instrument and the target. A standard altimeter product represents the average water level over the altimeter footprint (Smith 1997). Water surface elevation data provided by satellite-based RALT instruments on board of past, active and planned missions such as TOPEX/Poseidon (1992–2005), ERS-2 (1995–2011), ENVISAT (2002–2012), Jason-1 (2001–2013), 2 (2008) and 3 (January 2016) and Sentinel-3A-B-C (February 2016 planned, 2017, 2020) have centimetre accuracy. The revisit time of 10–35 days and the global coverage have allowed the development of global databases such as the LEGOS (Laboratoire d'Etudes en Géophysique et Océanographie Spatiales) database (http://www.legos.obs-mip.fr/en/soa/

hydrologie/hydroweb/) and the ESA's Rivers and Lakes project database (http://tethys.eaprs.cse.dmu.ac.uk/RiverLake). However, their kilometre footprint size, their limited orbital coverage due to an orbit-to-orbit spacing of several kilometres and an along-track spacing from 500 m to 6–7 km (e.g., Fu 2001) may impede their application for riverine flood modelling (e.g., Schumann et al. 2015; see Table 1). In small rivers, the surrounding topography can have a major impact on the echo shape returned to the altimeter and consequently the observed water surface elevation is potentially affected by significant uncertainties (e.g., Bercher and Kosuth 2012). For instance, Birkett et al. (2002) found that the minimum river width to be used with TOPEX/Poseidon, to avoid footprint contamination, is approximately 1 km. RALT water levels have thus been successfully used for many years to monitor the ocean surface and the water level of several kilometres wide inland water bodies (the Ob River in Kouraev et al. 2004; e.g., the Amazon River in Frappart et al. 2006; the Congo River in Lee et al. 2011). Nevertheless, the number of applications of RALT water levels for the study of small- to medium-sized rivers has recently increased. For instance, Tarpanelli et al. (2013a) used RALT water levels in combination with measured and modelled discharge data to estimate rating curves where the satellite track intersects the river reach along the Po River (Italy). We note here that in this study the Po River main channel width varied from 200 to 300 m and two lateral banks had an overall width from 400 m to 4 km. Using the same case study, Domeneghetti et al. (2014, 2015) showed that the use of a time series of RALT data, or an ensemble of RALT data from different satellite tracks, can compensate for the low accuracy of a single measurement over a medium-sized river and efficiently integrate in situ observations to improve the knowledge of the streamflow regime. Furthermore, some studies used RALT data to calibrate and validate hydraulic models (e.g., Wilson et al. 2007; Biancamaria et al. 2009; Siddique-E-Akbor et al. 2011).

The space-borne LiDAR GLAS (Geoscience Laser Altimeter System) on board of ICESat (2003–2009) had the potential to overcome the RALT footprint shortcoming (e.g., Schumann et al. 2015). This satellite mission was designed primarily to monitor changing elevations of ice sheets (Zwally et al. 2002), and GLAS pulses illuminate footprints 70 m in diameter, spaced at 170-m intervals along the Earth's surface with a revisit time interval of 91 days. Recent studies demonstrated the ability of ICESat's GLAS to retrieve water levels on large to medium rivers. In particular, Hall et al. (2012) showed that the mean difference between ICESat and gauged water levels from the Mississippi and Danube rivers was −16 cm (with a standard deviation of 73 cm); this value reduced to −10 cm (with a standard deviation of 27 cm) when a more restrictive observations selection criterion was used. This proof of concept study allowed the use of ICESat data to geodetically level gauges over a 400-km-long and 3.4-km-wide reach of the Amazon River. O'Loughlin et al. (2013) successfully retrieved the spatial and temporal dynamics of water slopes of the 500-m to 13-km-wide Congo River, while Neal et al. (2012) and Schumann et al. (2013) used ICESat's GLAS water levels to calibrate and validate hydraulic models of the 750-m-wide Niger River (Mali) and of the 1-km-wide Zambesi River (Mozambique). A follow-on ICESat-2 mission is planned for 2017 and might provide the opportunity for the analysis of flood events in medium to small catchments.

Spatial and temporal changes in water levels can also be estimated by radar interferometry (InSAR). This technique requires two SAR images from slightly different viewing geometries. Co-registration of the two images to a sub-pixel accuracy and subtraction of the complex phase (time delay) and amplitude (intensity) for each SAR image pixel allows changes in surface topography or displacements to be mapped (Bamler and Hartl 1998). For changes in water level retrieval, the specular reflection of smooth open water that causes most of the return signal to be reflected away from the antenna results in a complete loss of temporal coherence between SAR images acquired at different times, rendering interferometric retrieval difficult if not impossible (Alsdorf 2002). However, for inundated floodplains where there is emergent vegetation, the implementation of numerical algorithms such as the two step protocol suggested by Massonnet et al. (1993) allows reliable interferometric phase signatures of water level changes to be observed. Alsdorf et al. (2000, 2001, 2005, 2007) were able to measure water surface elevation changes to centimetric accuracy and with a spatial resolution of approximately 200 m across the roughly 50-km-wide Amazon floodplain. Jung et al. (2010) used SAR interferometry to contrast the spatial and temporal dynamics of river and floodplain water level connectivities of the Amazon, Congo and Brahmaputra river systems.

The possibility of using interferometric SAR for the calibration of two-dimensional hydraulic models was highlighted by Jung et al. (2012) in their study of the 2008 flood in the Atchafalaya River (Louisiana). Despite this promising result, the number of practical applications of InSAR hydrologic products is still sparse, most likely because of the complex numerical procedures that need to be applied for the detection of water level variations (Domeneghetti et al. 2014).

Finally, water level observations will be provided by the upcoming Surface Water Ocean Topography (SWOT) wide swath altimetry satellite mission, expected in 2020.

The SWOT mission is designed to observe all rivers wider than 100 m and water bodies (lakes, reservoirs, ponds, continuous wetlands) with an area greater than 250 m × 250 m (i.e. 62,500 m^2) (Rodrıguez 2015). Water level observations will be provided with 10-cm accuracy over 1 km^2 (e.g., a 10-km reach for a river of 100 m width) and 25-cm accuracy over 250 m^2 (Biancamaria et al. 2015). The orbit period will be approximately 21 days with the number of revisits per repeat period ranging from a minimum of two at the equator to more than ten above 70° N/S (Biancamaria et al. 2015). A key attribute of swath altimetry is the possibility of measuring spatial fields of surface water elevation rather than transects as provided by “traditional” radar altimetry. Furthermore, the Ka-band radar interferometer on board of SWOT is expected to be able to penetrate vegetation through canopy openings. Consequently, many researchers are enthusiastically looking at the SWOT satellite mission as an opportunity to sensibly improve our understanding of hydrology (e.g., Bates et al. 2014a). Andreadis et al. (2007), Biancamaria et al. (2011) and Andreadis and Schumann (2014) showed that the assimilation of synthetic SWOT water levels into two-dimensional hydraulic models can improve the accuracy of flood forecasts in the 30-m-wide Ohio River (USA) and in the 2- to 20-km-wide Ob River (Russia). Yet, it must be stated that, despite the targeted decimetre accuracy, considerable water level variation might exist over 1 km^2 in many catchments and investigation of other case studies is required.

3 SAR-Derived Observations of Flood Extent and Level

SAR is currently the most viable technique of observation of flood extent and level. Yet the retrieval of flood information from SAR imagery is not straightforward, and interpretation errors and inaccuracies impact the outcomes of the flood monitoring and modelling exercise. The following paragraphs discuss the challenges of the interpretation protocol, which is mainly composed by four steps: (1) pre-processing, (2) image classification, (3) retrieval of water levels from classified images and a DEM and (4) comparison with auxiliary data (when available).

3.1 Pre-processing

Geo-referencing, ortho-rectification and speckle removal are essential for the success of the retrieval of flood-related data from a SAR image. Geo-referencing locates the image on Earth; ortho-rectification corrects systematic sensor and platform-induced geometry errors which introduce terrain distortions when the sensor is not pointing directly at the Nadir location.

Speckle occurs where distributed targets are imaged and the pixel is therefore representative of the contributions coming from many scatterers with random phase. These contributions cause interference (i.e. a "salt-and-pepper" effect on the resultant image, e.g., O'Grady and Leblanc 2014) and result in a noise-like multiplicative modulation of the true backscatter (Oliver and Quegan 2004). Speckle characterisation and removal has been commonly attempted using filtering techniques (Frost et al. 1981, 1982; Lee 1983; Durand et al. 1987; Lee et al. 2009). More recently, Giustarini et al. (2015) proposed a nonparametric, fully automatic bootstrap method to account for the influence of speckle.

3.2 Image Classification for the Retrieval of Flood Extent Maps

Image classification is the interpretation process that aims to produce a map consisting of dry and flooded pixels. Surface roughness is considered to be the main factor affecting SAR backscattering: smooth surfaces such as flooded areas cause a sharp drop in the backscatter intensity (see Sect. 2.1.3). Nevertheless, a number of event-related and catchment-related meteorological and geometric factors can alter the backscatter characteristic causing errors in the detection of the flooded area. For instance, multiple reflections due to emerging vegetation and the roughening of the water surface due to rain and wind cause an increased backscatter that can lead to an underestimation of the flood extent (Mason et al. 2009; Zwenzner and Voigt 2009). Conversely, smooth surfaces such as roofs, tarmac and car parks act as specular reflectors and may lead to an overestimation of the flood extent (e.g., Giustarini et al. 2013). Moreover, radar shadow and layover caused by buildings and tall vegetation can hide relevant flooded areas.

These effects are function of the sensor characteristics, i.e. geometric spatial resolution, wavelength, incidence angle and polarisation. In a review on microwave remote sensing of flood inundation, Schumann and Moller (2015) denoted the lack of a detailed investigation on the sensitivities of the sensor characteristics for mapping flooded surfaces. Here, we summarise the few notable exceptions.

Wavelength is the distance over which the wave's shape repeats. Commonly used radar wavelengths for monitoring flood inundation processes range from 30 to 4 cm and include the bands L, C and X. Long L-band wavelengths can penetrate vegetation canopy and are less sensitive to wind and roughness on water surfaces. The capacity for canopy penetration is less at C- and X-band, and the total backscatter is predominantly a first surface return from the small vegetation components (Lang and Kasischke 2008; Whitcomb et al. 2009). Consequently, data from L-band sensors are preferable for the detection of floods in forested areas (Hess et al. 1990; Horritt et al. 2003; Schumann et al. 2012). Shorter-wavelength C- and X-band SAR data have proven useful for detecting floods in herbaceous and wooded wetlands (Evans et al. 2010; Hong et al. 2010; Brisco et al. 2011). The use of L-band data in combination with C- or X-band data has been suggested as the optimal solution for discrimination of flooded forest and other wetland surfaces (e.g., Mitchell et al. 2015).

The microwave polarisation describes the orientation of the electric field vector of the transmitted and reflected signals with respect to the horizontal direction. Henry et al. (2006) compared different polarisations for flood mapping purposes and concluded that the image mode in HH polarisation (horizontal transmit–horizontal receive), with its reduced sensitivity to water surface roughness due to wind, is most efficient in distinguishing flooded areas. However, Schumann et al. (2007b) pointed out the utility of VV and VH polarisations as VV polarisation data highlights vertical features like vegetation and VH polarisation data reflect the horizontal nature of the smoothed flood water. In fact, as previously discussed by Smith (1997), the optimal discrimination of various inundated vegetation cover types would require the analysis of multiple frequencies and polarisations.

Incidence angle refers to the angular deviation of the incident signal from nadir. In one of the rare studies, Lang and Kasischke (2008) found that the incidence angle had little impact on C-band backscatter at HH polarisation from flooded and non-flooded forest.

In order to deal with the listed sources of uncertainty and achieve an accurate detection of the flooded area, a number of image processing algorithms have been proposed. The most common procedures are visual interpretation (e.g., MacIntosh and Profeti 1995; Oberstadler et al. 1997), supervised classification (e.g., De Roo et al. 1999; Townsend 2002), image texture algorithms (e.g., Schumann et al. 2005), histogram thresholding (e.g., Brivio et al. 2002; Schumann et al. 2010), image statistic-based active contour models (e.g., Bates et al. 1997; Horritt 1999; Matgen et al. 2007a) and various multi-temporal change detection methods (e.g., Calabresi 1995; Delmeire 1997; Bazi et al. 2005).

Schumann et al. (2009a) and Di Baldassarre et al. (2011) pointed out that each image processing technique has advantages and disadvantages; consequently, no technique can be identified as optimal for specific image characteristics and no technique performs equally well for all the images.

The understanding of flood dynamics in urban areas is essential. However, radar layover, foreshortening, shadows and double backscatter due to buildings and man-made structures may impede the accurate mapping of flooded areas, even when SAR image resolutions of 3 m and higher are available, and enhanced image processing algorithms are used (e.g., Giustarini et al. 2013).

A possible approach includes the use of SAR simulators. SAR simulators are mathematical models of the overall SAR system chain that can be used to generate synthetic reflectivity maps of an urban area. Site-specific SAR effects such as layover, foreshortening, shadowing, double backscatter and speckle can thus be modelled for comparison with the RS data. For instance, Mason et al. (2012a) used a SAR simulator (Speck et al. 2007) in conjunction with LiDAR terrain data to estimate regions of the image in which water would not be visible due to shadow or layover caused by buildings and tall vegetation in the Severn River catchment (UK). An algorithm combining image segmentation, spectral and textural classification and region-growing techniques was then used to classify a TERRASAR-X image acquired during the July 2007 flood. The assessment of the algorithm flood detection accuracy was carried out using aerial photographs mosaics. In the urban area, 75 % of the urban water pixels visible to TerraSAR-X were correctly detected; even better results were achieved in rural areas, with almost 90 % of water pixels being correctly detected. Mason et al. (2014) showed that it is possible to detect flooding in radar layover regions by using a SAR simulator (Franceschetti et al. 2002) in conjunction with a high-resolution DEM to analyse the backscatter due to double bouncing of the radar signal between a flooded road and wall.

Since high-resolution DEMs are not globally available, Giustarini et al. (2013) formulated a change detection algorithm based on the calibration of a statistical distribution of open water backscatter values from images of floods. Images acquired during dry conditions enabled the identification of areas that are not visible to the sensor (i.e. affected by shadow and layover) and that systematically behave as specular reflectors (e.g., smooth tarmac, permanent water bodies). The algorithm was applied to classify the same TERRASAR-X image as Mason et al. (2012a) and yielded a classification accuracy in the urban area of around 81 % when compared with aerial photography-derived flooded areas.

Although this latter approach overcomes the need of a high-resolution DEM and a SAR simulator, it requires a reference image with the same imaging characteristics as the flood image. As such an image might be rarely available, Westerhoff et al. (2013) and Schlaffer et al. (2015) proposed an approach that takes advantage of the full information from a time series containing a number of images acquired over an area. Probability distributions of water and non-water backscatter are derived from multi-temporal SAR imagery; using these histograms, the probability of a "new" measurement belonging to either one or the other population can be derived. Westerhoff et al. (2013) showed that inundation maps of the Pakistan 2010 and Thailand 2011 events retrieved by using a reference data set derived from the analysis of ENVISAT-ASAR images had a strong resemblance with MODIS-derived results. Schlaffer et al. (2015) analysed over 500 ENVISAT-ASAR scenes with a spatial resolution of 150 m to characterise the seasonality in backscatter under non-flooded conditions of the river Severn catchment (UK). This reference data set was then used for the interpretation of an ENVISAT-ASAR image acquired during the 2007 flood. When validated against airborne photography and a TERRASAR-X-derived inundation map (Giustarini et al. 2013), flood detection accuracy was larger than 80 % in rural areas but lower than 50 % in urban areas. This large misclassification is consistent with the expectation that large portions of floods within urban areas remain unseen by coarse-resolution ASAR data due to layover and radar shadow.

In addition to the challenges of obtaining an accurate and reliable image interpretation, an ideal flood mapping system should be fully automatic in an operational crisis management context. Notable examples of research into automatic near real-time flood detection algorithms are provided by Martinis et al. (2009, 2011, 2015), Matgen et al. (2011) and Pulvirenti et al. (2011a, b). In a global context, Matgen et al. (2011) developed an automated flood mapping tool for different SAR image modes and resolutions based on a region-growing algorithm refined by change detection. Martinis et al. (2009, 2011, 2015) developed automatic, near-real-time flood detection algorithms based on a region-growing iterated segmentation/classification approach for single-polarisation high-resolution TerraSAR-X imagery, while Pulvirenti et al. (2011a) developed an algorithm that integrates backscatter analysis with simple hydraulic considerations and contextual information for the interpretation of COSMO-SkyMed SAR imagery.

Further, ESA is developing a free open-source toolbox named Sentinel-1 Toolbox (S1TBX) for the processing of SAR images acquired by ESA SAR missions including SENTINEL-1, ERS-1, 2 and ENVISAT, as well as third-party missions such as ALOS-PALSAR, TerraSAR-X, COSMO-SkyMed and RADARSAT-2. S1TBX includes tools for speckle filtering, orthorectification, mosaicking and data analysis.

In summary, uncertainty in flood mapping from SAR images stems from both the image input to the algorithm and the algorithm itself. Classification accuracies of flooded areas vary considerably and only in rare cases exceed 90 % (Schumann et al. 2012). A more extensive testing of the proposed algorithms on a number of case studies is necessary. For instance, as highlighted above, the same TerraSAR-X of the 2007 flood in the river Severn

(UK) was successfully analysed using different algorithms for the mapping of floods in urban areas (i.e. Mason et al. 2012a; Giustarini et al. 2013). However, it is worth noting that the low housing density and height in the area of acquisition may lead to optimistic conclusions on the efficiency of the algorithms proposed. A general prerequisite to estimate and eventually reduce uncertainties is to first identify and understand their sources (e.g., Schumann et al. 2012; Schumann and Moller 2015). Further research on automatic image classification in near real time is required. Additionally, in order to ensure a higher temporal and spatial coverage, flood observation services may need to make use of data sets stemming from different sensors (Matgen et al. 2011) and several different interpretation problems may have to be tackled in the study of a single flood event.

Finally, it is worth noting here that even small inaccuracies in the maps of the flooded areas can significantly reduce the quality of higher level products such as maps of inundation depths (see Sect. 3.3).

3.3 Assessment of Water Level from Maps of Flood Extent

The water level estimation methodology is composed of two compulsory steps, sometimes followed by three additional steps: (1) SAR image processing in order to extract the flood extent limits; (2) estimation of water levels by merging the flood extent limits and a digital elevation model (DEM); (3) application of constraining protocols to guarantee the hydraulic coherence of the data set of RS-D water levels (e.g., Mason et al. 2007; Schumann et al. 2007a; Hostache et al. 2009); (4) verification against field data (e.g., water/debris marks; Schumann et al. 2008c); (5) analysis of the data set for the retention of the data that could meet some pre-defined quality criteria (e.g., Hostache et al. 2009; Neal et al. 2009; Schumann et al. 2011; Stephens et al. 2012).

The implementation of hydraulic coherence protocols is deemed to account for most of the "noise" in water levels from a DEM at the SAR-derived flood boundaries. In particular, Schumann et al. (2007a) presented a steady-state regression analysis (REFIX—Regression and Elevation based Flood Information eXtraction) of water level values retrieved along the river centre line. Mason et al. (2007) suggested the need for ensuring that the water line varies smoothly in elevation along the reach. Hostache et al. (2009) suggested constraining the RS-D water levels using the hydraulic coherence algorithm by Raclot and Puech (2003) and Raclot (2006). The latter algorithm states that hydraulic energy decreases from upstream to downstream. Under the assumption of low velocity, this statement can be simplified into a decrease in water level in the flow direction. Notwithstanding the successful implementation of these strategies, the accuracy of the water level data set highly depends on the quality of the RS-D flooded area and the accuracy and resolution of the DEM. Taking advantage of the wide spatial coverage offered by remote sensing, many studies (e.g., Hostache et al. 2009; Neal et al. 2009; Schumann et al. 2011; Mason et al. 2012b; Stephens et al. 2012) suggested the retention of data that could meet some predefined quality criteria. For instance, Hostache et al. (2009) suggested removing steep areas, areas having dense vegetation and urban structures at the flood boundary. Mason et al. (2012b) and Stephens et al. (2012) described methods for selecting spatially independent subsets of water levels, with the aim to reduce the impact of spatially clustered errors.

The accuracy of the RS-D data set of water levels thus depends on a plethora of factors, whose impacts are related to the specific case study. A review by Di Baldassarre et al. (2011) showed that reported root mean square errors between the RS-D water levels and in situ data ranged from below 20 cm (Schumann et al. 2007a) to 2 m (Oberstadler et al. 1997; Schumann et al. 2008b).

Finally, similar techniques can be used to observe river widths, since it essentially requires the observation of a water mask and possibly prior information on river topology (Pavelsky and Smith 2008).

4 Information Content of Remote Sensing Data for Flood Forecast

4.1 A Comparison of the Information Content of Remote Sensing-Derived Observations and Traditional Field Data

Continuous time series of gauged water levels and/or discharge recorded at discrete locations have traditionally been used for model calibration and validation. More recently, they have been used for real-time constraint of hydraulic flood forecasting models (e.g., Madsen and Skotner 2005; Neal et al. 2007). Gauged data at discrete locations represent the aggregate response of the catchment to that point and might hinder relevant features of the flooding behaviour at the spatially distributed scale (e.g., Bates et al. 1998). The value of anecdotal yet spatially distributed high water marks or debris marks has been acknowledged by many studies (e.g., Hunter et al. 2005; Werner et al. 2005). Nevertheless, these kinds of field data are quite rare. The increasing availability of highly spatially distributed RS-D observations of flood extent and levels offer new opportunities for investigation and analysis (e.g., Bates 2004; Bates et al. 2014a; Schumann et al. 2009a). Per contra, time series of gauged water levels and/or discharge values are normally continuous in time, while RS-D observations of flood extent and level availability for a single flood event are often reduced to one, or just a few images. Despite the fact that the frequency of acquisition is likely to increase in the near future, the temporal coverage of RS-D observations of flood extent and level is inevitably discrete. Questions on the effectiveness of discrete-time-instant coverage and on the optimal timing and frequency of acquisition need to be answered. It is commonly agreed that an assessment of the uncertainty of observed data is pivotal to any meaningful calibration, validation or data assimilation exercise. A large number of studies are available for the assessment of the uncertainty in gauged data (e.g., Di Baldassarre and Montanari 2009; Domeneghetti et al. 2012; Tomkins 2014; Coxon et al. 2015). Errors in high water marks and wrack marks are related to the quality of the instruments and can be easily assessed. In contrast, the description of the uncertainty in RS-D observations of flood extent and levels is a current scientific challenge. The need for a thorough understating and an effective description of RS-D data uncertainty is exacerbated by their discrete-time coverage: with just one (or a few) images available during a flood event, a poor description of its uncertainty is deemed to negatively impact any calibration/validation/data assimilation protocol. The following paragraph presents a review of the approaches listed in the literature.

Table 2 presents a summary of the aforementioned field and RS-D data features used for flood monitoring and modelling.

4.2 Uncertainty in the Remote Sensing-Derived Observations

Notwithstanding the relevant technological and methodological progresses, RS-D observations of flood extent and levels are deemed to be intrinsically susceptible to sources of

Table 2 Main features of field and RS-D data used for flood monitoring and modelling

	Field data		RS-D flood extent and levels
	Gauged water levels/discharge	High water marks/debris marks	
Spatial coverage	Discrete	Distributed	Distributed
Temporal coverage	Continuous	Discrete but unreliable due it being anecdotal evidence	Discrete
Description of uncertainty	Mature topic	Mature topic	Young topic

uncertainty due to (1) imaging characteristics (e.g., imaging geo-referencing, look angle and resolution); (2) atmospheric and ground perturbations (e.g., rain, wind, trees, buildings, terrain geometry); and (3) image processing procedures (i.e. the algorithm selected, the quality of the DEM). A methodology for the description of the uncertainty in RS-D observations of flood extent and levels should be able to account for this long and varied list of (1) system-specific, (2) event and catchment-specific and (3) algorithm and data availability-specific sources of errors.

An almost straightforward approach used in many studies consists of the definition of intervals of values of RS-D water level data that implicitly represent a number of uncertainties. For instance, the maximum and minimum values defined by Schumann et al. (2008d) accounted for image noise, medium image ground resolution, image position errors, wind roughening and protruding vegetation (that is the sources of uncertainty (1) and (2) in the list above). Matgen et al. (2004) and Pappenberger et al. (2007) focused on the uncertainties in the image processing (3); Hostache et al. (2009) considered the uncertainties stemming from the image geo-referencing, the DEM altimetric uncertainty and the image processing (1, 3).

Many authors suggested the analysis of an ensemble of flood extent/water level maps derived from the application of a number of image processing techniques and parameters (3). More specifically, Schumann et al. (2009b) aggregated ten plausible flood maps derived by combining two contemporaneous SAR images with five different flood mapping procedures to build a "possibility of inundation map", which is a map of inundation that expresses a degree of belief that a given pixel (or area) is wet. Despite being a remarkable approach, it has been observed (e.g., Schumann et al. 2009b; Stephens et al. 2012) that the number of ensemble members and the procedure tend to be subjective and the "possibility of inundation map" cannot represent a probability in strict terms. Schumann et al. (2008c) suggested performing cross section-specific statistical analysis on the values of water levels extrapolated from an ensemble of flood maps. This approach is appealing as it allows exploiting the full empirical distribution of RS-D water levels and facilitates the identification of problematic areas which require careful hydraulic investigation prior to the implementation of flood inundation models (Giustarini et al. 2011). Nevertheless, RS-D data uncertainty can be very high and the distribution functions often exhibit bias and skewness (e.g., Schumann et al. 2010), hampering the definition of a statistical distribution. In order to limit the negative impact of bias in RS data, Neal et al. (2009), Giustarini et al. (2011) and Mason et al. (2010, 2012b) suggested to take advantage of the vast spatial coverage offered by the RS imagery by retaining only the measurements that meet some form of quality control and follow a Gaussian distribution. As an alternative to the implementation of a quality control routine, Giustarini et al. (2012) proposed the use of empirical, cross section-specific histograms of water levels which were derived by addressing the uncertainties of the entire SAR image processing chain. A bootstrap methodology was used to define 100 possible threshold values for delineation of the flooded area based on a hybrid methodology combining backscatter thresholding and region growing. The boundaries of the consequent 100 flood extent maps were then fuzzified to include the horizontal geo-referencing uncertainty. Each cross section was consequently assigned an ensemble of possible water levels and the vertical error associated with the DEM included.

5 Overview of the Study Sites

Flood monitoring from space is a recent development; although the amount of observations is expected to increase with the launch of new satellites, studies investigating the benefits for hydraulic flood models calibration/validation and real-time constraint had to rely on a limited number of RS-D flood extent and water level data sets currently available worldwide. RS-D observations information content for flood monitoring and forecast is the outcome of a combination of factors including sensor characteristics (mainly spatial resolution, wavelength, polarisation and look angle), timing of acquisition, image classification algorithm, catchment geometry, land cover/use and flooding behaviour. Knowledge of these details is required for a critical understanding of the results of any calibration/validation and data assimilation exercise. This section provides a schematic overview of the study sites and data sets used in Sects. 6 and 7. In particular, Table 3 details the geometry of the river network (main channel and floodplain), the implementation data, the available RS-D and field data for the specific flood event/period. Figure 1 shows the uneven spatial distribution of the study sites worldwide. The quality and quantity of the available data led the attention of the scientific community to small, European catchments (e.g., the Alzette River, Grand Duchy of Luxemburg). However, there is a need to test the use of RS-D observations for flood monitoring and forecast in a larger variety of catchments, having different geometries, different land cover/use characteristics, different flooding dynamics and where only low resolution or scarce implementation data set is available.

6 Remote Sensing-Derived Observations for the Calibration and Validation of Hydraulic Flood Forecasting Models

Calibration is the process by which model parameters are conditionally ranked based on their ability to make the model results match observed data. Validation requires that predictions of a model are compared to observed data, ideally independent from the calibration data set, to demonstrate the accuracy and reliability of the model.

Flood inundation models are traditionally calibrated by tuning the channel and floodplain roughness coefficients (e.g., Aronica et al. 1998, 2002; Horritt and Bates 2002; Pappenberger et al. 2005, 2007). These coefficients are used to account for more sources of uncertainty than just roughness (e.g., Romanowicz and Beven 2003; Di Baldassarre et al. 2011). According to Bates et al. (2014b), the major sources of uncertainty in flood inundation modelling are: (1) roughness values; (2) errors in the model input data (mainly

Table 3 Study sites

Study site	Geometry	DEM	Flood event	RS data	Field data
I. Alzette River (Luxembourg)	RL = 19 km FP-W = 300 m S = 0.08 %	2 m LiDAR DEM $A = \pm 15$ cm	[a] 2003, 50-year ARI	ERS-2 GR = 25 m AT = 10 h before the flood peak	6 WLh $A = \pm 1$ cm
				ENVISAT-ASAR GR = 25 m AT = 4 h after the flood peak	84 high water marks $A = \pm 2$ cm
			[b] 2007, 2-year ARI		6 WLh
II. Atchafalaya floodplain (USA)	RL = 225 km TA = 230 km^2	1 m LiDAR DEM	1 April 2008–1 June 2008	2 ALOS-PALSAR GR = 90 m (InSAR)	1 WLh
III. Dee River (UK)	RL = 20 km FP-W = 0.1–2 km CW = 12–30 m S = 0.005 % TA = 40 km^2	2 m LiDAR DEM A = 0.15 m	2006	ERS-2 GR = 25 m AT = right after the flood peak ENVISAT-ASAR WSM GR = 150 m AT = right after the flood peak	3 WLh
IV. Fitzroy river (Australia)	RL = 730 km FP-W = 15 km TA = 94,000 km^2	90 m SRTM DEM 30 m SRTM DEM	2002	MODIS GR = 500 m AMSRE-2 GR = 10,000 m	7 WLh
V. Genna stream, Tiber River (Italy)	RL = 20 km FP-W = 50–350 m CW = 15 m	LiDAR DEM A = 3 m	2010	ENVISAT-ASAR WSM GR = 150 m	1 Qh
VI. Haiti	TA = 135 km^2		2004	2 Terra ASTER GR = 90 m AT = before and after the flood	Low-flow discharge measurements
VII. Meuse River (NDL)	RL = 35 km	5 m DEM A = 6 cm	1995 63-year ARI	ERS-1: 30 January, discharge at Borgharen 2631 m^3s^{-1} Aerial photograph imagery: 27 January, discharge at Borgharen 2645 m^3s^{-1}	4 WLh 84 high water marks

Table 3 continued

Study site	Geometry	DEM	Flood event	RS data	Field data
VIII. Mekong River (Cambodia, Vietnam)	TA = 55,000 km^2	90 m SRTM DEM	2008, June–November	14 ENVISAT-ASAR WSM GR = 150 m	12 WLh
IX. Mengwa area (China)	TA = 180 km^2	A = 100 m	2007, June–July	MODIS GR = 500 m	1 uprestam Qh
X. Moselle River (France–Germany)	RL = 28 km FP-W = 3 km TA = 35,000 km^2 S = 0.05 %	2 m LiDAR DEM	1997, February	RADARSAT-1 GR = 25 m 5 aerial photographs 1:15,000 scale	3 Qh
XI. Ohio River (USA)	RL = 50 km	30 m SRTM DEM	[a] April 1st to June 23rd 1995	Synthetic (SWOT)	
	RL = 1007 km FP-W = 500 m CW = 30 m TA = 40,000 km^2	30 m SRTM DEM 30 m National Elevation Dataset (both resampled to 600 m)	[b] Oct 1984–Sept 1985	Synthetic (SWOT)	6 WLh and 2 Qh
XII. Po River (Italy)	RL = 140 km FP-W = 0.4–4 km CW = 200–500 m	2 m LiDAR DEM; 90 m SRTM	[a] 1995–2011	ERS2–RALT at two locations (May 1995 to June 2003) Temporal resolution = 35 days ENVISAT–RALT at one location (October 2002–August 2010) Temporal resolution = 35 days	2 WLh 2 Qh
			[b] 2000, 60-year ARI		2 WLh 2 Qh 176 high water marks A = ±15–20 cm
			[c] 2008, 2-year ARI	ENVISAT-ASAR WSM GR < 150 m AT = 1 h before the flood peak	2 WLh 2 Qh TR = daily average

Table 3 continued

Study site	Geometry	DEM	Flood event	RS data	Field data
XIII. Severn and Avon River (UK)	TA = 1524 km^2	2 m LiDAR DEM	[a] 2007, 150-year ARI		7 upstrem Qh 7 internal WLh 1 downstream WL
			[b] 2012	7 COSMO-SkyMed StripMap Himage GR = 5 m AT = decreasing limb	
XIV. Upper Severn River (UK)	RL = 60 km FP-W = 500 m CW = 30 m	2 m LiDAR DEM	[a] 1998	RADARSAT GR = 25 m AT = decreasing limb	
			[b] 2000	ERS-2 GR = 25 m AT = decreasing limb	
			[c] 2007 150-year ARI	2 ENVISAT-ASAR WSM GR = 150 m AT = after the flood peak ENVISAT-ASAR IMG GR = 25 m AT = decreasing limb RADARSAT 1 GR = 25 m AT = decreasing limb TERRASAR-X StripMap GR = 3 m AT = decreasing limb 2 AERIAL PHOTOGRAPHS GR = 1 m AT = decreasing limb	
XV. Sure River (Germany)	RL = 22 km FP-W = 500 m CW = 30 m	2 m LiDAR DEM $A = \pm 15$ cm	2011 5 years ARI	RADARSAT 2 GR = 5 m AT = close to the flood peak	1 WLh

Table 3 continued

Study site	Geometry	DEM	Flood event	RS data	Field data
XVI. Thames River (UK)	RL = 12 km FP-W = 0.2–2 km CW = 30 m S = 0.005 % TA = 25 km^2	1 m LiDAR DEM	1992 5 years ARI	ERS-1 SAR GR = 25 m AT = just after the flood peak.	1 upstream Qh
XVII. Zambesi River (Mozambique)	TA = 170,000 km^2	90 m SRTM DEM	2007	Envisat-ASAR GR = 150 m AT = a few days after the peak	

A accuracy; *AT* acquisition time, *CW* average channel width, *FP-W* average floodplain width, *GR* ground resolution, *RL* reach length, *Qh* discharge hydrograph, *S* average slope, *TA* total area simulated, *TR* temporal resolution, *WLh* water level hydrograph

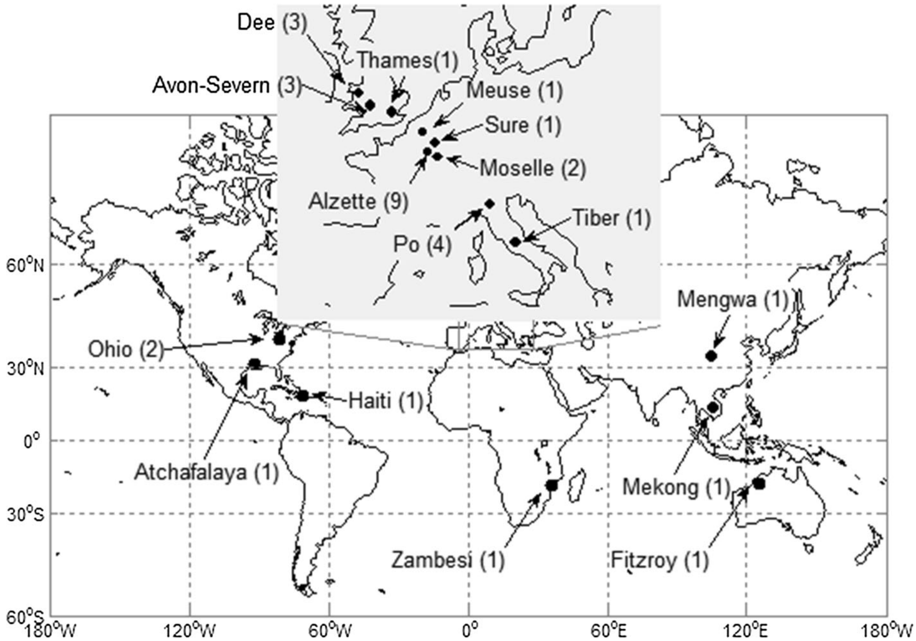


Fig. 1 Geographical distribution of the study sites

boundary and initial conditions, topography and bathymetric data); (3) model structural errors; (4) conceptual model uncertainty (e.g., three-dimensional processes are not represented by two-dimensional models); and (5) errors in the independent observed data used to estimate model parameters. Therefore, calibrated parameter values are not physically realistic and a large number of different parameter sets might be able to map model predictions to the observed data to an acceptable level of performance. This is the concept of equifinality in environmental modelling (Beven 2006). Methodologies acknowledging the uncertainty inherent in the inundation modelling process have gained favour in recent years as being preferable to generating single deterministic maps of flood extent/level (Aronica et al. 1998; Romanowicz and Beven 1998; Aronica et al. 2002; Romanowicz and Beven 2003; Bates et al. 2004; Hall et al. 2005; Pappenberger et al. 2005). The GLUE (generalised likelihood uncertainty estimation) methodology of Beven and Binley (1992) embraces the concept of equifinality and attempts to find those feasible models that provide acceptable fits to any available observational data.

6.1 Overview of the Studies

Continuous time series at discrete locations of gauged water levels/discharge have traditionally been used for model calibration and validation. However, many different spatial patterns of effective parameter values can lead to the same aggregate response. Schumann et al. (2008a) argued that in order to constrain model parameter uncertainty and at the same time increase parameter identifiability as much as possible, models need to satisfy some behavioural criterion at a large number of spatially distributed locations. Therefore, the frequent lack of distributed field data such as high water marks or wrack marks has so far been a significant contributor towards equifinality (Grayson and Blöschl 2001). Despite a substantial agreement on the potential value of spatially distributed RS-D water level/flood extent data for the calibration and validation of hydraulic models (e.g., Bates 2004; Schumann et al. 2015), the optimal use of this uncertain, low-frequency data set compared to traditional field data and the definition of the optimal calibration protocol are current scientific challenges (e.g., Yan et al. 2015).

In the Generalized Likelihood Uncertainty Estimation (GLUE) procedure, each member of a Monte Carlo ensemble of model simulations is assigned a weight according to how well it fits with observed calibration data. Model calibration and validation then require the use of performance measures to determine the accuracy of each parameter set to match the observed data. Threshold performance measure values are sometimes defined to reject the simulations that deviate too much from the observations. However, the use of a threshold performance measure has been criticised in the past (Gupta et al. 1998) for its potential lack of objectivity. A weighted sum of the predictions from a subsample, or all the simulations, then produces an uncertain model prediction (Romanowicz et al. 1996). In the case of traditional field data of time-series gauged water levels and discharge values, three quantitative performance measures have been generally recommended, more specifically the Nash–Sutcliffe efficiency (NSE), per cent bias (PBIAS) and ratio of the root mean square error (RMSE) to the standard deviation of measured data (RSR) (Moriasi et al. 2007). The RMSE has been traditionally used when water or debris marks were available. New performance measures have been required to fully use the wide potential information content embedded in RS imagery of flood events. The understanding that uncertainty should be viewed as central to the use of RS data resulted in a shift from deterministic to uncertain performance measures and the correlated adoption of the extended GLUE (Aronica et al. 2002) methodology.

Sections 6.2, 6.3 and 6.4 provide a review of the performance measures and the calibration and validation strategies used in the literature. Section 6.5 reviews the benchmarking analysis completed to compare the effectiveness of field and RS data in constraining the parameter space of hydraulic models. Sections 6.6, 6.7 and 6.8 focus on the impact of the characteristics of RS data on the effectiveness of the calibration and validation exercise.

Table 4 summarises the studies discussing the use of RS-D observations of flood extent and levels for the calibration and validation of hydraulic models for flood forecasting.

6.2 Calibration Strategies: Deterministic Performance Measures

Due to the difficulties in describing and quantifying RS-D data uncertainty, many studies still assume the RS data to be "perfect". Examples are the use of RS-D flooded area by Hunter et al. (2005), Horritt (2006), Mason et al. (2009) and Dung et al. (2011); the use of RS-D observations of inundation width by Di Baldassarre et al. (2009a) and Prestininzi et al. (2011); and the use of RS-D water levels by Mason et al. (2009) and Domeneghetti et al. (2014).

As the extraction of flood extent maps is the first step in the RS image processing chain, a number of performance measures to condition inundation models on flood extent data have been proposed in order to limit the sources of uncertainty in the observational data set. Observed and modelled data are divided into discrete categories of wet/dry cells separated by deterministic boundaries for the purpose of building a contingency table (see Table 5) which reports the number of pixels correctly and incorrectly predicted as wet or dry. The model performance is then assessed by the binary measures given in Table 6. Following the recommendations of Schumann et al. (2005) and Hunter et al. (2005), the Critical Success Index (CSI) has been most commonly used. However, the lack of

Table 4 Summary of the main studies featuring the use of RS-D observations of flood extent and level for hydraulic models calibration and validation

Reference	Topic (section)	Study site (Table 3)	RS-D product	RS data	DEM [m]	Hydraulic model	Approach	Performance measures (See Table 6)	Uncertainty in RS data (0 = RS-D data are assumed perfect)
Matgen et al. (2004)	6.5	I [a]	IW	ENVISAT-ASAR ERS-2	75	HecRas 1D	GLUE	Membership function	Interval of confidence
Hunter et al. (2005)	6.5	VII	FE	ERS-2 air photo	5	LISFLOOD 2D	GLUE	CSI (FE) Heteroscedastic likelihood measure (WLh); MAV (wrack marks)	0
Horrit (2006)	6.4	XIV [a]	FE	RADARSAT	250	LISFLOOD 2D	GLUE	$F^{<3>}$ (FE) Entropy Reliability diagram	0
Pappenberger et al. (2007)	6.3	I [a]	FE	ENVISAT-ASAR	50	LISFLOOD 2D	GLUE	Standard similarity function	Fuzzy category vectors
Schumann et al. (2008c)	6.3	I [a]	WL	ENVISAT-ASAR	2				Cross section-specific statistical distribution of RS-D WLs
Schumann et al. (2008a)	6.3	I [a]	WL	ENVISAT-ASAR	2	HecRas 1D	GLUE MR	Membership function	Interval of confidence
Brandimarte et al. (2009)	6.6	VI	FE	Terra ASTER					0
Di Baldassarre et al. (2009a)	6.5 6.6	XII [b, c]	IW	ENVISAT-ASAR	2	HecRas 1D	Det	MAE	0
Hostache et al. (2009)	6.3 6.5	I [a, b]	WL	ENVISAT-ASAR	2	HecRas 1D	GLUE MR	Membership function RMSE; NS	Interval of confidence
Schumann et al. (2009b)	6.3	III	FE	ENVISAT-ASAR WSM ERS-2				CSI	Possibility of inundation map
Di Baldassarre et al. (2009b)	6.3	III	FE	ENVISAT-ASAR WSM ERS-2	20	LISFLOOD 2D	GLUE	CSI reliability diagram	Possibility of inundation map

Table 4 continued

Reference	Topic (section)	Study site (Table 3)	RS-D product	RS data	DEM [m]	Hydraulic model	Approach	Performance measures (See Table 6)	Uncertainty in RS data (0 = RS-D data are assumed perfect)
Mason et al. (2009)	6.2 6.8	XVI	FE	ERS-1	1	LISFLOOD 2D	GLUE	CSI (FE)	0
			WL					Student t test (WL)	
Schumann et al. (2010)	6.5 6.6	XII [c]	WL	ENVISAT-ASAR WSM	90	HecRas 1D		Visual comparison	Interval of confidence
Hostache et al. (2010)	6.5	X	WL	RADARSAT-1	2	original 2D	Det, 4D-Var	Cost function	Interval of confidence
Dung et al. (2011)	6.5	VIII	FE	ENVISAT-ASAR WSM	90	Mike 1D	Det	CSI (FE); NS (WLh)	Statistic approach over the neighbouring cells
Prestininzi et al. (2011)	6.6	XII [c]	FE	ENVISAT-ASAR WSM	2	HecRas 1D	Det	MAE	0
						1D–2D	Det	"Adjusted" CSI	0
Stephens et al. (2012)	6.2 6.3 6.8	III	FE	ERS-2	5	LISFLOOD 2D	GLUE	CSI (FE)	Selected values
			WL					RMSE (WL)	
Tarpanelli et al. (2013a, b)	6.6	V	FE	ENVISAT-ASAR WSM	3	Mike 11	Det	PC (FE) NS (Qh)	6 maps of FE (6 image processing methods)
Karim et al. (2011)	6.5	IV	FE	MODIS AMSRE-2	30	Mike 21	Det	RMSE	0
Domeneghetti et al. (2014)	6.5	XII [a,c]	WL	ERS-2–RALT ENVISAT–RALT	2	HecRas 1D	Det	NS	0
Lai et al. (2014)	6.6	IX	FE	MODIS	100	original 2D	Det, 4D-Var	Cost function	Interval of confidence

Table 4 continued

Reference	Topic (section)	Study site (Table 3)	RS-D product	RS data	DEM [m]	Hydraulic model	Approach	Performance measures (See Table 6)	Uncertainty in RS data (0 = RS-D data are assumed perfect)
Schumann et al. (2014)	6.2	III	FE	ERS-2	20	LISFLOOD 2D	ROC	ROC	ROC
		XVII	FE	ENVISAT–ASAR WSM	90	LISFLOOD 2D			

Det deterministic calibration, *FE* flood extent, *GLUE* generalised likelihood uncertainty estimation, *IW* inundation width, *MR* model rejection, *Q* discharge hydrograph, *ROC* receiver operating characteristics space, *WL* water level, *WLh* water level hydrograph, *4D-Var* 4D-variational technique

Table 5 Contingency table

	Present in observation	Absent in observation
Present in model	A	B
Absent in model	C	D

Table 6 Flood extent-based performance measures (the values A to D are from Table 5)

Name	Equation	Description	Issues
Bias	$\frac{A+B}{A+C}$	Aggregate model performance highlighting over-/underprediction	(i) (iv)
Proportion correct (PC) or $F^{<1>}$	$\frac{A+D}{A+B+C+D}$	Proportion of cells whose wet/dry state has been correctly predicted over the total extent of the study area	(i) (ii) (v) (iv)
Critical success index (CSI) ore threat score ($F^{<2>}$)	$\frac{A}{A+B+C}$	Adjustment of the PC for the quantity being forecast (Wilks 2011)	(i) (iii) (iv)
"Adjusted" critical success index	$\frac{\sum_i w_{i,\delta} A_i}{\sum_i w_{i,\delta} A_i + \sum_i w_{i,\delta} B_i + \sum_i w_{i,\delta} C_i}$	Adjustment of the CSI that aims to remove the issue (iv). The weight factor $w_{i,\delta}$ depends both on ground features and model performances (Prestininzi et al. 2011)	(i) (vi)
$F^{<3>}$	$\frac{A-C}{A+B+C}$	Designed to penalise underprediction	(i) (iv)
$F^{<4>}$	$\frac{A-B}{A+B+C}$	Designed to penalise overprediction	(i)
Hit rate (H)	$H = \frac{A}{A+C}$	Fraction of the observed flood that is correctly predicted (it detects underprediction)	(i) (iv)
False alarm rate (F)	$F = \frac{B}{B+D}$	Fraction of dry areas that are incorrectly predicted (it detects overprediction)	(i) (iv)
Receiver operating characteristics (ROC)	H versus F	The ROC space depicts the relative trade-offs between benefits (H) and costs (F)	(i) (iv)
Pierce skill score (PSS)	$H - F$	Threshold that aims at maximising difference between H and F	(i) (iii) (iv)

consistency of binary measures was reported by many authors. Pappenberger et al. (2007) interestingly related the reliability of a binary performance measure to the features of the image processing algorithm. For instance, they stated that active contour algorithms, as well as the CSI, favour the same type of flooding pattern (large areas in contrast to a fragmented floodplain), and this combination may hinder the inefficiency of the CSI. Comprehensive analyses by Hunter et al. (2005), Schumann et al. (2009a) and Stephens et al. (2014) identified six possible issues: (i) the sensitivity to the magnitude of the flood; (ii) the bias towards unflooded areas; (iii) the bias towards overprediction or underprediction of the flooded areas; (iv) the sensitivity to the shape of the valley; (v) the sensitivity to the domain size; and (vi) the sensitivity to the resolution of the model. The last column of Table 6 summarises the contributions of the above listed studies.

The use of any binary performance measure is thus subordinated to the preliminary acknowledgement of its limitations (Stephens et al. 2014), while the contemporary use of different formulations has the potential to provide a more complete description of the prediction accuracy (Pappenberger et al. 2007).

Schumann et al. (2014) formulated a statistical measure in the receiver operating characteristics (ROC) space to allow an automated model calibration without the need of a classification of the SAR image. A SAR image was classified using histogram thresholding with threshold values varying from $-\infty$ to $+\infty$. In a Monte Carlo framework, the ROC curve having the largest area pinpoints the optimal parameter set and the optimal SAR classification threshold value. The RS-D flooded area and the optimal parameter set identified by the ROC-based method were consistent with the results of Di Baldassarre et al. (2009b) and Schumann et al. (2009b, 2013).

As an alternative to area based binary performance measures, Di Baldassarre et al. (2009a) and Prestininzi et al. (2011) used the mean average error (MAE) to compare deterministic values of modelled and observed inundation width.

Due to the recent progresses in RS image analysis, some authors tested the possibility of using water level-based performance indices. In particular, Mason et al. (2009) proposed the use of the Student *t* test, while Stephens et al. (2012) used the RMSE to compare the values of modelled and observed spatially distributed values of water level. Both of the proposed water level-based measures were more sensitive to the roughness parameters and less sensitive to clustered errors in the observed data when compared to the traditionally used water extent-based measures. These characteristic resulted in a higher effectiveness of water level-based measures over flood extent-based measures in constraining the parameter space.

6.3 Calibrating Uncertain Models with Uncertain Remote Sensing-Derived Observations

The need for the acknowledgement of the intrinsic uncertainty in RS-D observation of flood extent and level has been claimed by many authors. For instance, Pappenberger et al. (2007) observed that ignoring RS-D data quality issues may mask the real information content of the observations and mislead the conditioning of the models. Di Baldassarre et al. (2009b) and Tarpanelli et al. (2013b) showed that successive, yet independent, deterministic calibration exercises based on different interpretations of the same SAR image led to different "optimal" parameters. In an attempt to reduce the impact of uncertainties of RS-D data, Stephens et al. (2012) used multiple samples of spatially independent water level values derived from the same SAR image. However, the subjective selection of a sample of the observed data unintentionally included and excluded observed data error, and as a result the choice of subset caused variations in the computed "optimal" parameter set. Conversely, Horritt and Bates (2002) showed that accounting for observation errors in an uncertain classification procedure led to more consistent results between data sets with different accuracies (i.e. an ERS-1 image and an ENVISAT-ASAR Wide ScanSAR Mode image). Recent studies have consequently looked at the analysis of the uncertainty in RS-D observations of flood extent and levels as an opportunity to fully exploit the information content of RS imagery (e.g., Pappenberger et al. 2007; Hostache et al. 2009; Schumann et al. 2009b; Stephens et al. 2012).

Methodologies such as the extended GLUE procedure that acknowledge the uncertainty inherent in both the flood modelling process and the observed data have progressively gained favour, urging the need for methods to describe the uncertainties in RS-D observations of flood extent and levels. Bias in the distributions of RS data has hampered so far the use of a probabilistic distribution (Schumann et al. 2008d). However, a number of alternative strategies have been proposed. These strategies are based on fuzzy logic, reliability diagrams and the statistical analysis of the results of the calibration exercise.

6.3.1 Fuzzy Performance Measures

In a fuzzy logic approach, all RS-D data are affected by an interval representing their uncertainty. Performance measures quantifying the degree of belonging (i.e. membership functions) of the model results to these intervals are used for model evaluation. Different shapes of the membership function and different approaches to quantifying the uncertain intervals of water levels, inundation widths and wet/dry boundaries have been proposed. A simple step membership function was used by Schumann et al. (2008a): at each cross section, the maximum uniform weight (that is 1) was attributed to the realisations of the hydraulic model that predicted water levels within the interval of the observed water level data. A zero weight was used otherwise. Slightly more complex trapezoidal membership functions were adopted by Hostache et al. (2009) and Matgen et al. (2004) to quantify the degree of belonging of modelled along-reach water levels and inundation width values to the uncertain RS-D data. Finally, Pappenberger et al. (2007) used the standard similarity function of Hagen (2003) to assess the level of similarity of the observed and modelled fuzzy flooded areas. All these approaches were used for the calibration of a hydraulic model of the 2003 flood event in the Alzette River (Grand Duchy of Luxemburg). However, the lack of model validation inhibited any conclusive recommendation on the most effective approach.

An accurate description of all sources of uncertainty in RS-D data proved to be pivotal. When a step membership function was used to assign a constant weight to intervals of RS-D water level (Schumann et al. 2008a), the constraint of the parameter space was driven by the very restricted number of cross sections characterised by a small height variation leading to the risk of artificially overfitting the model to these specific locations. Conversely, a comprehensive analysis of RS data uncertainty and a trapezoidal membership function (Hostache et al. 2009) included higher information content and allowed the representation of the average hydraulic behaviour of the catchment. Despite being positively correlated with a large number of deterministic binary measures (Pappenberger et al. 2007), each fuzzy performance measures normally yielded a lower constraint of the parameter space. Nevertheless, this increased degree of freedom can be seen as an opportunity to build a hydraulic model able to reasonably reproduce flood events of different magnitude. For instance, Hostache et al. (2009) showed that a one-dimensional (1D) hydraulic model of the Alzette River, calibrated using RS-D water levels observed during a 50-year Average Recurrence Interval (ARI) flood in 2003 and a trapezoidal fuzzy performance measure could reproduce a 2-year ARI flood in 2007.

6.3.2 Reliability Diagrams

Reliability diagrams are graphs of the observed frequency of an event plotted against the forecasted probability of an event (Hartmann et al. 2002). A perfect forecast system will result in forecasts with a probability of X % being consistent with the eventual outcome X % of the time. Hence, when plotting a reliability diagram, comparisons are made against the diagonal.

Exploiting the quite unique opportunity of two coincident satellite images during the 2006 flood in the River Dee (UK), Schumann et al. (2009b) derived a “possibility of inundation map” applying five different processing protocols to images having different resolutions and characteristics. Di Baldassarre et al. (2009b) then constructed a reliability diagram by classifying the “possibility of inundation map” of Schumann et al. (2009b) into regions of similar possibility of inundation and counting the number of simulated wet cells in each region. In a Monte Carlo framework, each model realisation was evaluated by computing the RMSE between the cloud of points and the diagonal. These weights were used to combine the results of these numerical simulations and obtain an uncertain flood inundation map.

6.3.3 "Post"-Analysis of the Calibration Results

Considering the impossibility of accounting for all the sources of uncertainties in RS-D data, some authors suggested performing a statistical analysis on the results of calibration protocols. Stephens et al. (2012) calibrated a two-dimensional (2D) hydraulic model of the 2006 flood in the river Dee (UK) using a deterministic approach and multiple samples of spatially independent water level values extracted from a SAR image. The frequency with which each parameter set appeared as the optimum was used as a weight within a GLUE-style uncertainty framework. The resulting flood uncertain inundation map reflected the confidence in the observed data and showed an overall good agreement with the observed data.

6.4 Validation Strategies

A good model is one that, once calibrated for one flood event, reproduces a different event with the required predictive capability (e.g., Horritt et al. 2007; Schumann et al. 2009a). Defining the "required predictive capability" is subjective. Under the assumption of ideally perfect models, some authors (e.g., Prestininzi et al. 2011) adopted the upper boundary target of replicating an event with the same degree of uncertainty as the observations. However, real uncertain models hardly (if ever) achieve this target during calibration; further, calibrated parameter values are not physically realistic (Sect. 6) and flood events having different magnitude might require a different set of parameter values. In a GLUE framework, a model is more robust if the cloud of the best scoring values in the parameter space does not change when moving from one event to another (e.g., from calibration to validation). In order to allow some degree of freedom in model performance, a threshold to identify the best scoring values should be defined. The lack of a broad consensus on protocols and criteria to evaluate the performance of hydraulic models for flood forecasts has been pointed out by many authors (e.g., Horritt 2006; Yan et al. 2015).

Models calibrated using RS-D data have been validated using field data in both a deterministic (e.g., Tarpanelli et al. 2013b; Domeneghetti et al. 2014) and stochastic (e.g., Hostache et al. 2009) framework. However, the use of lumped field data might not allow a comprehensive validation of the model behaviour at large spatial scales.

Measuring the consistency between uncertain flood inundation maps and observed floods would be conceptually easy if a large number of observed flood events were available allowing use of the reliability diagrams (Horritt 2006). However, since RS observations are currently available for a very small number of flood events, only limited form of validation is possible by estimating how well the uncertain flood inundation map matches the observed extent. Albeit this comparison can only be approximate because the observed flood extent has been used to derive the weights of the parameter space used for the computation of the uncertain flood inundation map (Mason et al. 2009), it allows some sensitive analysis on model performances.

Most studies listed in the literature simply commented on a visual comparison between observed and modelled uncertain flood inundation maps. Such a qualitative evaluation hampers any benchmarking analysis on the effectiveness of the proposed calibration strategies, even when the same flood event and the same hydraulic model were used as case study (e.g., Matgen et al. 2004; Pappenberger et al. 2007; Schumann et al. 2008a; Di Baldassarre et al. 2009b; Stephens et al. 2012). A protocol based on the use of quantitative measures to evaluate uncertain flood inundation maps when one single image is available was suggested by Horritt (2006).

Two desirable properties of an uncertain flood inundation map are accuracy and precision. A precise flood map will contain large areas which are classified as definitely wet or dry. However, such a precise map may not be accurate, as it may not coincide perfectly with the real flood extent. Horritt (2006) suggested the Shannon entropy as a measure of model precision and the reliability diagram as a measure of the accuracy of probabilistic predictions. The use of the Shannon entropy for the comparison of different data sets ability in constraining the parameter space of hydraulic models was previously demonstrated by Hunter et al. (2005). The reliability diagram was constructed by classifying the uncertain flood map into regions of similar predicted possibility of flooding and counting the number of observed wet cells in each region. The protocol was then applied for assessing both the precision and accuracy of uncertain predictions of flood extent of two events that occurred in the river Severn (UK), in 1998 and 2000, respectively. A SAR image was available for each event. When both the measures of precision and accuracy were applied, differences in the optimal parameter sets needed to reproduce two events of different magnitude became evident, exposing the challenge of using models in a predictive mode and the danger of over-fitting the model to the calibration data (Romanowicz and Beven 2003).

The need for testing the model results for both precision and accuracy in order to avoid misleading conclusions was subsequently recognised by Mason et al. (2009) and Di Baldassarre et al. (2009b). Alternatively to the reliability diagram, Mason et al. (2009) used a chi-square goodness-of-fit test to verify the null hypothesis that the observed flood extent matches the predicted uncertain flood map, while Di Baldassarre et al. (2009b) computed the difference between the modelled and the observed possibility of inundation at pixel level.

6.5 Utility of Remote Sensing-Derived Observations of Flood Extent and Level in Contrast to Field Data

It is generally agreed in the scientific community that the information content of RS data is very valuable. However, they do not yet seem capable of being a complete substitute for in situ observations. The fuzziness and poor description of the uncertainty of the SAR-derived values of inundation width (e.g., Matgen et al. 2004), flooded area (e.g., Hunter et al. 2005; Pappenberger et al. 2007) and water level (e.g., Schumann et al. 2008a) and the limited spatial distribution of RALT observations (e.g., Domeneghetti et al. 2014) limit the ability to constrain the parameter space of any hydraulic model.

The measurements provided by ground-based gauging stations still remain invaluable for constraining flood model predictions. In some cases, adequate model calibration was only possible when RS-D observations were used in combination with field data (e.g., Matgen et al. 2004; Hunter et al. 2006; Hostache et al. 2009). However, currently available SAR-derived uncertain values of inundation width (e.g., Matgen et al. 2004; Di Baldassarre et al. 2009a), flooded area (e.g., Hunter et al. 2005) and water level (e.g., Hostache et al. 2009, 2010; Domeneghetti et al. 2014) can integrate traditional field data to improve the accuracy of the calibration of the hydraulic model. Hostache et al. (2009), Di Baldassarre et al. (2009a) and Dung et al. (2011) further showed that discrete field data representing the bulk response of the catchment might not allow adequate verification of a hydraulic model, while the intrinsically two-dimensional features of remote sensing observations densely distributed over space naturally induce more coherent and more explicative modalities of comparison. For instance, using SAR-derived inundation widths, Di Baldassarre et al. (2009a) demonstrated that a model which is a good predictor of internal hydraulic properties might not be a good predictor of the flooded area. Hostache et al. (2009) further stated that the local nature of field data can lead to an overfitting of the hydraulic model to a specific location or event, thus limiting the overall predictive skills of the model. Conversely, the increased degree of freedom allowed by RS data uncertainty provides an opportunity to build a hydraulic model that can reasonably reproduce flood events having different magnitudes.

Dung et al. (2011) showed that the mutually exclusive use of stage hydrographs and flooded areas can lead to biased and contradictory model behaviour as information on temporal dynamics from in situ gauging stations and spatial dynamics from RS-D inundation maps is complementary. The authors then successfully adopted a multi-objective calibration protocol to exploit the information content derived from a network of 12 river stage hydrographs and a series of inundation maps derived from ENVISAT-ASAR imagery.

A multi-objective calibration exercise completed by Hunter et al. (2005) highlighted the relevance of the accuracy of the RS data. More specifically, a high-resolution optical air photograph was much more effective in discriminating between parameter sets than the coincident ERS-1 SAR image (25-m ground resolution). Furthermore, they showed that the information on flooded area retrieved from a single, accurate air photograph led to a similar degree of reduction in the parameter space uncertainty of 84 field water marks. The effectiveness of water marks in constraining the parameter space was previously demonstrated by Werner et al. (2005). Considering the rare availability of field water marks, the results highlight the value of accurate RS-D observations of flood extent and levels.

Notwithstanding the need for more accurate and temporally continuous field data, many studies showed that calibration performed exclusively using RS-D observations was still able to provide a valuable prediction of the flooded area (e.g., Matgen et al. 2004 used ENVISAT-ASAR- and ERS-2-derived inundation width values; Pappenberger et al. 2007 used ENVISAT-ASAR-derived flooded area) or enhance the modelling of the average streamflow of a large river (e.g., Domeneghetti et al. 2014 used ENVISAT RALT data).

In summary, RS-D data are not yet ready to replace field data for hydraulic model calibration and validation entirely, with improvement in the accuracy and description of their uncertainty required. The temporal resolution of RS data is likely to increase, and investigations into the optimal acquisition time frame are under way. The use of field data and RS-D data in a combined, multi-objective calibration manner is currently the best strategy (e.g., Hall et al. 2011; Bates et al. 2014b; Domeneghetti et al. 2014) in order to fully exploit the complementary information content on temporal dynamics from in situ gauging stations and spatial dynamics from RS-D data. RS-D data provide information that cannot be achieved with field data; as such more research focusing on the use of RS data for the calibration of hydraulic models is required.

6.6 Utility of Low-Resolution Remote Sensing

Higher temporal repeat of acquisition is currently possible through the use of low-resolution sensors. Furthermore, RS data from low-resolution sensors are often available at global scale and at no or low cost. Several studies consequently investigated the possibility of using coarse RS data to improve the accuracy of flood forecasting, especially in ungauged or data-sparse areas (see review from Schumann et al. 2009a; Di Baldassarre et al. 2011; Yan et al. 2015).

Inundation width values and flooded areas derived from ENVISAT-ASAR WSM images superimposed on a 2-m LiDAR DEM were successfully used by Di Baldassarre et al. (2009a), Prestininzi et al. (2011) and Tarpanelli et al. (2013b) to calibrate hydraulic models of the Po and Tiber rivers in Italy. Schumann et al. (2010) showed that when only freely available low-resolution space-borne data sets with global coverage can be used, water profiles derived from the intersection between an ENVISAT-ASAR WSM image and the global 90-m SRTM DEM were able to discriminate between competing model parameterizations. The satisfactory results obtained for medium-sized catchments (in the Po and Tiber river case studies the floodplain has an extension varying from 400 to 4000 and from 50 to 350 m, respectively) highlighted the potential of using coarse RS data to support flood modelling. However, Schumann et al. (2010) underlined that accounting for RS-D observations and DEM uncertainties is essential for the success of the calibration exercise, especially in medium to steep topography. Furthermore, RS-D water level values are likely to be more sensitive to the low accuracy of the original data, and RS-D inundation width, flooded area and river reach water level profiles are recommended when using a coarse data set (Schumann et al. 2008b).

As previously stated, the systematic use of optical data is limited by the inability of the sensor to penetrate the cloud cover. Yet, when cloud-free, coarse optical images proved to be useful. One example is the successful use of an ASTER image of the 2004 flood in Haiti by Brandimarte et al. (2009). Some studies, e.g., Lai et al. (2014) and Ticehurst et al. (2014, 2015), investigated the more appealing opportunity of using the large number of images derived from the "frequent" revisit time of coarse optical sensors for the analysis of floods in large, slow-motion catchments. In particular, Lai et al. (2014) used flood extent maps derived from a series of MODIS images to calibrate the hydraulic model of the 180-km^2-wide Mengwa flood detention area (China). Ticehurst et al. (2013) showed that MODIS spatial consistency and high temporal frequency (1–2 times per day) can be of great value in detecting general changes in water movement in large scale Australian catchments such as the lower Condamine–Balonne flood plain (3800 km^2), the Fitzroy River (32,000 km^2) and the Macquarie Marshes (3000 km^2). The best results were achieved in the wide lower part of the catchments and during the slow flood recession phase.

6.7 Timing Considerations

Model response to the calibration process might be significantly different for high and low flows on the inter- as well as the intra-event scale (e.g., Romanowicz and Beven 2003; Hunter et al. 2005). Stand-alone low-resolution sensors and medium- to high-resolution sensors carried by a constellation of satellites provide the highest acquisition frequency (see Sect. 2.1.3). However, even in the frame of a fast-progressing technology, the discrete-time coverage will remain an intrinsic limitation of RS data. Consequently, the acquisition time may be crucial in determining the success of using RS data in constraining the parameter space for the purpose of flood modelling (e.g., Schumann et al. 2009a).

With flood forecasting being the focus of the modelling study, observations of larger discharge values might be of greater interest in testing model performance against conditions as close as possible to those being predicted. For instance, Horritt (2006) showed that inundation extents acquired at higher discharges were more effective at constraining model predictions of the 1998 and 2000 floods in the river Severn (UK) than satellite overpasses during the falling limb of the hydrograph, when the floodplain was dewatering slowly. However, an exception to this was shown by the case study of the 2003 "valley filling event" in the Alzette River (Grand Duchy of Luxembourg), as investigated by Matgen et al. (2004) and Schumann et al. (2008c). In a "valley filling event", flood extents show little sensitivity to high discharge values, and shoreline positional errors can cause significant uncertainties in stage retrieval. In these cases, acquisitions at a lower flow than peak may provide a better test of model performances. At a preliminary stage of the flood, small changes in the channel roughness coefficient may have a large impact on the simulated extent and thus provide useful information on channel hydraulics. Further, when the river channel is at bankfull, small changes in the discharge tend to induce large changes of the flood extent, and the lumped friction parameters need to be adjusted accordingly. Using multi-temporal airborne radar imagery of the 2000 flood in the river Severn (UK), Horritt et al. (2007) showed that observations made during the receding limb of the hydrograph, when the discharge was approximately at bankfull and then below bankfull were more effective at distinguishing models than imagery captured at peak flow.

In conclusion, care is required to ensure that the observations provide useful information on the hydraulic processes that the model is trying to represent (Schumann et al. 2009a). Inundation patterns at different stages during the flood event may be dominated by the topography of the floodplain, and these effects may limit the ability of observations to constrain the model parameter space. A strong correlation between the observed quantity and flood dynamics is pivotal in determining the success of the calibration/validation protocol, and the advisable acquisition time is a function of the specific case study.

6.8 Selection of the Remote Sensing-Derived Product

As can be inferred from the above, a range of RS-D data can be used for the calibration and validation of hydraulic models: (1) maps of flood extent, (2) spatially distributed inundation width values and (3) spatially distributed water level values. A large variability of the observed values within the flooding process allows a more effective constraint of the parameter space. Flood extent and inundation width data are highly correlated with flood dynamics in floodplains with slowly varying slopes. Conversely, only water level values vary strongly in floodplains constrained by steep slopes or in the occurrence of valley filling events (e.g., Horritt 2006; Mason et al. 2009; Stephens et al. 2012, 2014). In order to capture the capability of each observation data set to evaluate the performance of various features of the model, many studies (e.g., Schumann et al. 2008d; Hall et al. 2011) suggested a multi-objective calibration framework representing a compromise between observed flood area and spatially distributed water stages.

7 Remote sensing-Derived Observations for Real-Time Updating of Hydraulic Flood Forecasting Models

Data assimilation (DA) is an approach to the problem of updating a dynamical system using both current and past observations, together with a model to provide temporal continuity and dynamic coupling among the variables (e.g., Charney et al. 1969; Hunt et al. 2007; Houser et al. 2010). Basically, there are four methods for model updating (Houser et al. 2010): state updating, input updating, error prediction or correction and parameter updating. State updating methods adjust internal model states based on observations; input updating methods involve generating new estimates of input data and then running them through the simulation model; error prediction methods analyse past errors between observations and model predictions to predict future model errors; and parameter updating methods seek to update the parameters of simulation models.

Data assimilation requires a numerical algorithm that, given a (noisy) model of the system dynamics, finds the best estimates of system states from (noisy) observations (Houser et al. 2010). Most current approaches to this problem are derived from either the direct observer or dynamic observer techniques. Direct observer techniques sequentially update the results of the numerical model (i.e. the "forecast" or "a priori estimate") using the difference between observations and model predicted observations. This difference is known as the "innovation", and it is computed whenever new observations are available. The model predicted observations are calculated from the model "background" states. The correction, or "analysis increment", added to the background state vector is the innovation multiplied by a weighting factor or "gain". The resulting estimate of the state vector is known as the "a posteriori estimate". The commonly used direct observer techniques are: (1) direct insertion; (2) statistical correction; (3) successive correction; (4) analysis correction; (5) Newtonian nudging; (6) optimal statistical interpolation; (7) one-, two- or three-dimensional variational assimilation; (8) Kalman filter and variants; and (9) particle filter and variants. The dynamic observer finds the best fit between the forecast model state and the observations by minimising an objective or penalty function over space and time "window", including a background and observation penalty term to account for initial state vector uncertainty and observation uncertainty. Dynamic observer methods are well suited for smoothing problems, but provide information on estimation accuracy only at a considerable computational cost. The 4D (3D in space, 1D in time) "variational" (otherwise known as Gauss-Markov) assimilation and the ensemble Kalman smoother (Dunne and Entekhabi 2005) are examples of dynamic observer techniques. For a more comprehensive analysis of the direct and dynamic observer techniques in land surface data assimilation, we refer to Houser et al. (2010).

Many studies assimilated in situ measurements for hydraulic modelling applications (e.g., Shiiba et al. 2000; Madsen and Skotner 2005; Neal et al. 2007). As pointed out by Schumann et al. (2009a), there is no doubt that a comprehensive remote sensing data assimilation (RS DA) framework has the potential of becoming a critical component in future flood forecasting systems. However, only a few studies have so far attempted to assimilate RS data into hydraulic models (e.g., Matgen et al. 2007b; Giustarini et al. 2011; García-Pintado et al. 2015, please refer to Sect. 7.1 for the complete list). The objective of this section is to analyse the outcomes of the ongoing debate on how to integrate RS data into hydraulic modelling for flood forecasting.

7.1 Overview of the Studies

The assimilation of RS-D observations of water levels, inundation widths or flood extents into hydraulic models for flood forecasting has started to be investigated only in the last decade, taking advantage of the increasing availability of RS data and of the correlated ongoing developments in the interpretation of RS images. The number of historical events for which RS data are available is still limited (see Sect. 5). However, the amount of applications is expected to increase with the launch of new satellites scheduled for the next years. In this context, many authors opted for the analysis of synthetic scenarios in which synthetic RS-D data were generated either from a so-called truth realisation of the numerical model or using a simulator algorithm (such as the Jet Propulsion Laboratory Instrument Simulator, Rodriguez and Moller 2004). Synthetic RS-D data were generated to emulate the characteristics of products potentially provided by specific sensors and the analysis of synthetic scenarios aimed at proving their potential utility for flood forecasting. Furthermore, these studies focused on the problem of scheduling satellite-based acquisitions for sequential assimilation into operational flood modelling. Conversely, studies based on real case scenarios aimed at proving the utility of the assimilation of currently available RS-D data.

For both synthetic and real case scenarios, preliminary choices of the data assimilation algorithm, the updating method, the RS-D observation (i.e. water levels, inundation width or flooded area) have to be made. These aspects are discussed in Sects. 7.2 and 7.6, respectively. Section 7.3 lists the performance measures used to evaluate the impact of RS DA on the accuracy of flood forecast. A few benchmarking analyses have contrasted the effectiveness of assimilating RS and field data into hydraulic models for flood forecasting; the main outcomes are presented in Sect. 7.4. The accuracy of the observations, the first visit time and the frequency of acquisition affect the result of any DA strategy. Section 7.5 summarises the findings and recommendations of the published studies. As DA techniques are not meant to correct systematic errors (biases), Sect. 7.7 discusses the impact of a poor implementation and calibration of the hydraulic model on the accuracy of the flood forecasts.

Table 7 summarises the studies focussing on the assimilation of RS-D observations of flood extent and levels for improving the accuracy of hydraulic models for flood forecasting.

7.2 Data Assimilation Algorithm

The large majority of the approaches applied in the literature for the assimilation of RS data into hydraulic models for flood forecasting were derived from the direct observer technique. In particular, Matgen et al. (2007b) applied a direct insertion method while Andreadis et al. (2007), Andreadis and Schumann (2014), Neal et al. (2009), Matgen et al. (2010), Giustarini et al. (2011, 2012) and García-Pintado et al. (2013) applied filtering methods. Per contra, Lai and Monnier (2009) peculiarly provided an example of the application of a dynamic observer technique. In an artificial test case based on the 1997 flood event in the Mosel River (France), they used a 4D-variational technique to assimilate synthetic RS-D water levels into a 2D hydraulic model for the updating of the input discharge hydrograph.

The next paragraphs present studies that used a filtering approach. The ensemble Kalman filter and the particle filter and their variants were used for the assimilation of RS-D data into hydraulic models. The selection of the filtering algorithm was based on:

1. the ability of the filter to model the uncertainties in the RS-D data; and
2. the physical dimension of the model domain.

Early studies completed by Andreadis et al. (2007) and Neal et al. (2009) were based on the Ensemble Kalman filter (EnKF) analysis scheme proposed by Evensen (1994, 2003, 2004). The EnKF is a Monte Carlo implementation of the Bayesian update problem. The EnKF is computationally efficient; however, it makes the assumption that all probability distributions involved (including the distribution of the measurement error) are Gaussian. As detailed in Sect. 4.2, studies investigating the full empirical distribution and the

Table 7 Summary of the main studies featuring the assimilation of RS derived data into flood forecasting hydraulic models

<table>
<tr><th>References</th><th>Topics (section)</th><th>Study site (Table 3)</th><th>DA algorithm and updating method</th><th>Description of RS data uncertainties</th><th>Hydraulic model</th><th>Performance measure and main results
Analysis step: Dist (FM,AM); SDEns
Forecast: AE (WL), RMSE (t, WL)</th></tr>
<tr><td>Matgen et al. (2007a, b)</td><td>7.2.2</td><td>I
ERS-2 and ENVISAT-ASAR derived WLs</td><td>Direct insertion
State updating</td><td>Intervals of likely water levels</td><td>HEC-RAS 1D (calibrated)</td><td>Forecast
<table>
<tr><td></td><td colspan="2">$\lvert WL_{obs} - WL_{mod} \rvert$ [cm]</td></tr>
<tr><td>Time</td><td>Open loop</td><td>Filter</td></tr>
<tr><td>t_{ass} (ERS-2)</td><td>30</td><td>7</td></tr>
<tr><td>t_{ass} (ERS-2) +4.5h</td><td>14</td><td>14</td></tr>
<tr><td>t_{ass} (ENVISAT–ASAR)</td><td>3</td><td>6</td></tr>
<tr><td>t_{ass} (ENVISAT–ASAR) + 20h</td><td>17</td><td>17</td></tr>
<tr><td>Peak</td><td>25</td><td>25</td></tr>
</table></td></tr>
<tr><td>Andreadis et al. (2007)</td><td>7.2.2
7.5.2</td><td>XI [a]
Synthetic WLs (SWOT)</td><td>EnKF
State updating
Input updating (autoregressive error model)</td><td>Gaussian distribution:
$\mu = 0$
$\sigma = 5$ cm</td><td>LISFLOOD-FP
SRTM DEM; river bathymetry
Spatial discretisation: 270 m</td><td>Forecast
<table>
<tr><td colspan="4">RMSE (Q) [m^3/s]</td></tr>
<tr><td rowspan="3">Open loop</td><td colspan="3">Filter</td></tr>
<tr><td colspan="3">Assimilation frequency</td></tr>
<tr><td>8 days</td><td>16 days</td><td>32 days</td></tr>
<tr><td>23.2 %</td><td>10 %</td><td>12.1%</td><td>16.9%</td></tr>
<tr><td colspan="4">RMSE (WL) [cm]</td></tr>
<tr><td>56.0</td><td>21.6</td><td>24.9</td><td>33.3</td></tr>
</table></td></tr>
<tr><td>Neal et al. (2009)</td><td>7.4
7.7
7.5.1</td><td>I
ENVISAT-ASAR derived WLs</td><td>EnKF</td><td>WL observations that meet the quality requirements</td><td>HEC-RAS 1D (not calibrated)
Measured and simplified river bathymetry</td><td>Analysis step:
<table>
<tr><td></td><td colspan="2">SD(Ens) at t_{ass}</td></tr>
<tr><td></td><td>Open loop</td><td>Filter</td></tr>
<tr><td>WL</td><td>0.65</td><td>0.09</td></tr>
<tr><td>Q</td><td>21.1</td><td>4.4</td></tr>
</table>
<table>
<tr><td></td><td colspan="2">Dist(FM,AM) at t_{ass}</td></tr>
<tr><td></td><td>Open loop</td><td>Filter</td></tr>
<tr><td>WL, RMSE</td><td>0.7</td><td>0.4</td></tr>
<tr><td>Q</td><td>-14</td><td>+10</td></tr>
</table></td></tr>
<tr><td>Matgen et al. (2010)</td><td>7.2.2
7.5.1
7.5.2</td><td>I
Synthetic WLs</td><td>PF
State updating
Input updating (constant error model)</td><td>Gaussian distribution:
$\mu = 0$
$\sigma = 10, 30, 50, 70, 100, 200, 500, 1000$ cm</td><td>HEC-RAS 1D (calibrated)</td><td>Forecast
<table>
<tr><td colspan="4">RMSE (WL) [cm]</td></tr>
<tr><td rowspan="3">Open loop</td><td colspan="3">Filter;σ (RS-D WLs): 30cm</td></tr>
<tr><td colspan="3">Assimilation frequncy</td></tr>
<tr><td>12 h</td><td>24 h</td><td>48 h</td></tr>
<tr><td>34</td><td>18</td><td>19</td><td>25</td></tr>
</table></td></tr>
</table>

Table 7 continued

<table>
<tr><th>References</th><th>Topics (section)</th><th>Study site (Table 3)</th><th>DA algorithm and updating method</th><th>Description of RS data uncertainties</th><th>Hydraulic model</th><th>Performance measure and main results
Analysis step: Dist (FM,AM); SDEns
Forecast: AE (WL), RMSE (t, WL)</th></tr>
<tr><td>Giustarini et al. (2011)</td><td>4.2
7.2.1
7.4
7.7</td><td>I
ERS-2 and ENVISAT-ASAR derived WLs</td><td>PF, global and local formulations
State updating Input updating (constant error model)</td><td>RS data:
Uniform distribution- average spread: 54 cm; accuracy: 30 cm (Hostache et al. 2009)
Field data:
Gaussian distribution:
$\mu = 0$
$\sigma = 1$ cm</td><td>HEC-RAS 1D (calibrated)</td><td>Analysis step
<table>
<tr><td colspan="4">Dist(FM)/ Dist(AM) at many locations;If |Dist(FM) / Dist(AM) |<1 the DA did not improve the analysis</td></tr>
<tr><td colspan="2">Global filter</td><td colspan="2">Local filter</td></tr>
<tr><td>RS WL</td><td>Field data</td><td>RS WL</td><td>Field data</td></tr>
<tr><td>-18÷ 0.7</td><td>-1.4÷0.7</td><td>-1.1÷0.7</td><td>0.8÷1</td></tr>
</table>
Forecast: RMSE(t,WL)
<table>
<tr><td></td><td colspan="3">RMSE, WL [cm]</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">Open loop</td><td colspan="2">Filter</td></tr>
<tr><td>GW</td><td>LW</td></tr>
<tr><td>20h (t_{ass} (ERS-2))</td><td>17</td><td>22</td><td>10</td></tr>
<tr><td>30h</td><td>40</td><td>36</td><td>43</td></tr>
<tr><td>33h - t_{ass} (ENVISAT-ASAR)</td><td>22</td><td>0</td><td>5</td></tr>
<tr><td>52h</td><td>30</td><td>70</td><td>90</td></tr>
</table></td></tr>
<tr><td>Giustarini et al. (2012)</td><td>4.2
7.5.1
7.7</td><td>A) XII [c]
ENVISAT-ASAR [90 SRTM DEM] derived WLs
B) XV
RADARSAT-2 [Lidar DEM 2 m] derived WLs</td><td>PF
State updating
Input updating
(constant error model)</td><td>Cross section-specific empirical histograms</td><td>HEC-RAS 1D (not calibrated)</td><td>Forecast - Po river: AE(WL)
<table>
<tr><td></td><td colspan="2">$\lvert WL_{obs} - WL_{mod}\rvert$ [cm]</td></tr>
<tr><td>time</td><td>Open loop</td><td>Filter</td></tr>
<tr><td>ASAR-WSM</td><td>100</td><td>-30</td></tr>
<tr><td>ASAR-WSM + 110 h</td><td>100</td><td>+30</td></tr>
</table>
Forecast - Sure river: AE(WL)
<table>
<tr><td></td><td colspan="2">$\lvert WL_{obs} - WL_{mod}\rvert$ [cm]</td></tr>
<tr><td>time</td><td>Open loop</td><td>Filter</td></tr>
<tr><td>t_{ass} (RADARSAT-2)</td><td>50</td><td>0</td></tr>
<tr><td>t_{ass} (RADARSAT-2) + 300 h</td><td>47</td><td>15</td></tr>
</table></td></tr>
<tr><td>García-Pintado et al. (2013)</td><td>7.5.2</td><td>XIII
Synthetic WLs (CosmoSkyMed)</td><td>LEnKF
State updating
Input updating (inflow correction: constant bias)</td><td>Gaussian distribution:
$\mu = 0$
$\sigma = 25$ cm</td><td>LISFLOOD-FP 2D (calibrated)</td><td>Forecast:
$\frac{\text{RMSE(Q, early first visit)}}{\text{RMSE(Q,late first visit,)}} = 0.35$
$\frac{\text{RMSE(WL, early first visit)}}{\text{RMSE(WL, late first visit)}} = 0.25$</td></tr>
</table>

Table 7 continued

<table>
<tr><th>References</th><th>Topics (section)</th><th>Study site (Table 3)</th><th>DA algorithm and updating method</th><th>Description of RS data uncertainties</th><th>Hydraulic model</th><th>Performance measure and main results
Analysis step: Dist (FM,AM); SDEns
Forecast: AE (WL), RMSE (t, WL)</th></tr>
<tr><td>Andreadis and Schumann (2014)</td><td>7.3
7.6
7.5.2
7.5.3
7.7</td><td>XI [b]
Synthetic WLs (SWOT)</td><td>LEnKF
State updating</td><td>Gaussian distribution:
$\mu = 0$
σ,WLs = 10 cm
σ (TW, FA) = 20 %</td><td>LISFLOOD-FP, 2D
(calibrated)
Spatial resolution: 600 m</td><td>Cost function: time and spatially-averaged error reduction due to the assimilation when compared to an open loop simulation.
<table>
<tr><td colspan="6">Assimilation of WLs</td></tr>
<tr><td></td><td colspan="5">Lead time [days]</td></tr>
<tr><td></td><td>1</td><td>3</td><td>5</td><td>7</td><td>11</td></tr>
<tr><td>WL [m/km]</td><td>0.59</td><td>0.45</td><td>0.32</td><td>0.23</td><td>0.16</td></tr>
<tr><td>Q [m^3/s/km]</td><td>57.1</td><td>20.2</td><td>-12</td><td>-30</td><td>-30</td></tr>
<tr><td colspan="6">Assimilation of TOP WIDTH</td></tr>
<tr><td></td><td colspan="5">Lead time [days]</td></tr>
<tr><td></td><td>1</td><td>3</td><td>5</td><td>7</td><td>11</td></tr>
<tr><td>Q [m^3/s/km]</td><td>-5.3</td><td>-8.4</td><td>-9.7</td><td>-11</td><td>-14</td></tr>
<tr><td colspan="6">Assimilation of FLOODED AREA</td></tr>
<tr><td></td><td colspan="5">Lead time [days]</td></tr>
<tr><td></td><td>1</td><td>3</td><td>5</td><td>7</td><td>11</td></tr>
<tr><td>FA [km^2]</td><td>14.4</td><td>13.6</td><td>12.2</td><td>10.4</td><td>8.4</td></tr>
</table></td></tr>
<tr><td>García-Pintado et al. (2015)</td><td>7.2.1
7.2.2</td><td>XIII [b]
Real data:
WLs derived from 7 COSMO-SkyMed images</td><td>EnKF, LEnKF
State updating
Input updating (inflow correction: constant bias)
Parameter updating</td><td>Gaussian distribution:
$\mu = 0$
$\sigma = 25$ cm</td><td>LISFLOOD-FP 2D
(calibrated)</td><td>*Forecast:* RMSE(t,Q, across 7 gauge stations)
<table>
<tr><td colspan="2">Local EnKF, state and input updating</td></tr>
<tr><td>Traditional metric</td><td>Original metric</td></tr>
<tr><td>56</td><td>68 - higher time correlation</td></tr>
<tr><td colspan="2">Global EnKF, state and input updating</td></tr>
<tr><td colspan="2">100</td></tr>
</table></td></tr>
</table>

AE average error, *Dist (FM,AM)* distance between the mean value of the forecast step and the mean value of the analysis step, *FA* flooded area, *Q* discharge, *RMSE* root mean square error, *SD (Ens)* standard deviation of the ensamble, *TW* top width, t_{ass} assimilation time, *WLs* water levels, μ mean, σ standard deviation uncertainty description of RS-D data are still limited. In one of the rare studies, Schumann et al. (2008c) showed that a Gaussian error assumption may not be adequate for most of the RS-D ensembles of water levels. Consequently, Matgen et al. (2010) and Giustarini et al. (2011, 2012) proposed an assimilation scheme based on the particle filter (PF) as a possibility to relax the Gaussian assumption of the EnKF. The PF can indeed easily manage the propagation of a non-Gaussian distribution through nonlinear hydrologic and hydraulic models (Moradkhani 2008). The DA protocol proposed by Matgen et al. (2010) and Giustarini et al. (2011, 2012) can thus be adapted to any empirical or theoretical distribution function. However, since theoretical distribution functions alternative to the Gaussian distribution have yet to be specified, following Neal et al. (2009), Matgen et al. (2010) suggested the possibility of applying the PF to assimilate a subsample of measurements that did not fail a normality test. Giustarini et al. (2011, 2012) assumed that the RS-D water levels were uniformly distributed within an interval derived using the protocol proposed by Hostache et al. (2009). In a follow-up study, Giustarini et al. (2012) formulated an approach that allows exploiting the full empirical spread of RS-D water levels, without requiring any assumption on their distribution function (Sect. 4.2).

García-Pintado et al. (2013) noted the relatively small size and the mono-dimension of the model setups investigated by Matgen et al. (2010) and Giustarini et al. (2011). As discussed by Snyder et al. (2008) and Matgen et al. (2010), despite the potential to improve the PF efficiency for rivers with complex geometry and large dimensional problems, the research is still in progress. Conversely, the feasibility of the EnKF with ensemble sizes (i.e. the number of model realisations) much smaller than the state dimension has been demonstrated in operational numerical weather prediction (Houtekamer and Mitchell 2005) and is theoretically justified (Furrer and Bengtsson 2007). Based on these considerations, the EnKF was recently implemented by García-Pintado et al. (2013, 2015) and Andreadis and Schumann (2014).

7.2.1 Global or Local Data Assimilation Strategy

Global and local formulations for both the PF and the EnKF are possible. The two different strategies were tested and compared, respectively, by Giustarini et al. (2011) and García-Pintado et al. (2015).

Generally speaking, in a global filter the likelihood of a model realisation is based on its ability to correctly predict the state variables (i.e. water levels, inundation width, flooded area and discharge) along the entire river domain. A global PF was adopted in the synthetic experiment implemented by Matgen et al. (2010). They used the same model to generate and assimilate the synthetic satellite observations. Consequently, for a given forcing (inflow hydrograph), the model generally performed equally well (or poorly) at all cross sections along the river, and the choice of a global weighting procedure in which a single particle contained a state vector the water levels at all cross sections was appropriate. However, as observed by Giustarini et al. (2011) in a real case study, model structural errors and parameter uncertainties cause local bias that need to be taken into consideration. As the global weighting procedure favours compromising solutions that provide acceptable results at all model cross sections, it might happen that one model performs well over the entire river reach, but at the same time has a very poor performance at many local levels. This raises difficulties in the selection of a good model application. Moreover, as García-Pintado et al. (2015) remarked in their analysis of the global formulation of the EnKF, spurious updates of the state at locations physically disconnected from the observation are possible. Localisation techniques were then suggested to correct the problem.

Giustarini et al. (2011) proposed a local variant of the PF. The local PF procedure attributes a separate particle set to each cross section (i.e. a single particle contains the water level from one cross section as state vector) and thus associates likelihoods to each model realisation according to its ability to correctly predict the water stage at a given cross section. The global and local formulations were applied to assimilate (1) field data and (2) ERS2- and ENVISAT-ASAR-derived water levels into a hydraulic model of the 2003 flood in the Alzette River (Grand Duchy of Luxembourg). The authors demonstrated that a local weighting procedure is the preferred solution when assimilating unbiased and/or very precise observations, while in ungauged basins where highly uncertain RS-D observations are the only available data source, a global weighting procedure is recommended. Based on these recommendations, Giustarini et al. (2012) used the global weighting formulation of the PF to assimilate water levels derived from a coarse ENVISAT-ASAR Wide Swath Mode image into an uncalibrated 1D hydraulic model of the 2008 flood in the Po River (Italy).

The local formulation of the EnKF assimilation protocol has been traditionally based on an observation localisation method (e.g., Hunt et al. 2007; Nerger and Gregg 2007; Kirchgessner et al. 2014) in which the assimilation is applied independently to a series of disjoint local domains, and only observations within some defined cut-off radius are considered (Kirchgessner et al. 2014). García-Pintado et al. (2015) formulated a novel distance metric based on an along channel network distance that is originally able to account for the physical connectivity of flows. They subsequently tested the efficiency of the global formulation, the traditional local formulation and their own novel local formulation of the EnKF to improve the forecast of the 2012 flood in the Severn and Avon rivers (UK). Water levels derived from 7 COSMO-SkyMed images had a Gaussian distribution with 25 cm standard deviation and were assimilated into a 2D hydraulic model. The authors showed that their novel local formulation of the EnKF was able to filter out the spurious correlations and unphysical relationships dominating the global filter. When the local filter was applied, more observations shared a fair contribution to the updating, making the filter more robust to outliers in the observations.

A local EnKF was subsequently applied by Andreadis and Schumann (2014) to assimilate synthetic values of RS-D water levels, inundation width and flood extent into a 2D hydraulic model for the prediction of the flow regime of the Ohio River over 1 year. The river was partitioned into reaches of equal lengths, and the assimilation was performed exclusively using the forecasts and observations of that reach. In the case of the Ohio River (516 km), the localisation reach length was set at 5 km; however, the authors stated that very similar results were obtained for lengths up to 50 km.

7.2.2 State, Input and Parameter Updating

The most straightforward use of RS-D data to improve the forecasting skills of a hydraulic model is state updating. Each time a RS observation is available, the model state is updated and the new values are then used as the new ‘initial condition” of the hydraulic model. Following the analysis step, the model is propagated in time and its result becomes the forecast in the next analysis cycle (i.e. when new observations become available). For instance, Matgen et al. (2007b) used a direct insertion method that forced water level data simulated by a hydraulic model to fall within the confidence interval of ERS-2-derived water levels. Matgen et al. (2010) subsequently used the PF to update the forecasted water levels using synthetic water levels with an “acquisition” frequency down to 12 h and an accuracy of 0.30 m. Andreadis et al. (2007) and García-Pintado et al. (2013) used the

EnKF to update the modelled water level and discharge values using synthetic observations of water level. Andreadis and Schumann (2014) used the local EnKF to update water levels, discharge or flooded extent values using synthetic observations of water levels, top width and flooded areas. Using different algorithms for the assimilation of different states in real and synthetic scenarios, all the studies listed above concluded that, although the models results were more accurate immediately after the analysis step, the improvement in forecasting skills due to state updating had a short time span. Depending on the specific case study, a few hours or even a few minutes after the state updating, the model forecast returned towards the open-loop model realisation. The positive impact of a mere re-initialization of hydraulic models is indeed limited by the dominating effect of the upstream boundary condition (i.e. inflows).

To tackle the problem of non-persistent model improvements, many authors (e.g., Andreadis et al. 2007; Matgen et al. 2010; Giustarini et al. 2011) suggested a shift from a simple state updating method to a combination of state updating and input updating. Different protocols to correct input errors were used. For instance, Andreadis et al. (2007) computed a relative error term based on the difference between the a priori and the a posteriori discharge values at each analysis step. Then, from the next step until the next assimilation time, every member of the ensemble of input discharge hydrographs was corrected applying the autoregressive error forecast model proposed by Madsen and Skotner (2005). The autoregressive error forecast model essentially regresses the current value of a time series against the value at the previous time step. Over an 84-day simulation of the hydraulic behaviour of the Ohio River, the filter was able to recover water levels and discharge even when a 32-day satellite acquisition frequency was used.

Alternatively, Matgen et al. (2010) adopted a constant error forecast model: the relative error computed at one analysis step was applied to correct every member of the ensemble of discharge hydrographs from the next step until the next assimilation time. In this approach, the underlying assumption is that current model errors are due to an over- or underestimation of water stored in the basin. In their analysis of the 2003 flood event in the Alzette River, Matgen et al. (2010) showed that the time-averaged RMSE of water level was 0.34 m for an open-loop simulation and 0.33 m when a simple state updating method was applied to assimilate synthetic RS-D water levels having a Gaussian distribution with zero mean and standard deviation of 0.30 m and a frequency of 48 h. The time-averaged RMSE reduced to 0.25 m when the same RS-D water levels were used in a combined state and input updating approach.

Giustarini et al. (2011) applied the scheme for state updating and input correction based on the constant inflow error forecast model proposed by Matgen et al. (2010) to a real case scenario. Water levels derived from an ERS-2 and an ENVISAT-ASAR image, respectively, acquired during the rising limb and immediately after the flood peak of the 2003 flood event in the Alzette River were assimilated into a 1D hydraulic model. The proposed inflow correction model led to improved predictions for more than 5 h after the assimilation. Nevertheless, during the receding limb the forecasts with filtering performed worse than the open-loop predictions. Giustarini et al. (2011) observed that the two satellite observations were acquired when model errors are known to be only weakly correlated in time. During the rising limb of the event, precipitation errors continuously add to model parameter and model structural errors and the assumption of constant relative errors throughout the flood event might not be valid. However, Giustarini et al. (2012) obtained more persistent improvements in the predicted water levels when the PF with the same constant inflow error forecast model was applied to two different real cases. First, a 2-year Average Recurrence Interval (ARI) flood in the Po River observed by a coarse-resolution

(75 m) ENVISAT-ASAR Wide Swath Mode image recorded immediately before the flood peak, and second, a 5-year ARI flood event in the Surat River observed by a high-resolution (5 m) RADARSAT-2 image recorded immediately after the flood peak. Both test cases obtained improvements in the predicted water levels over several hours. The reasons for this might be the different approach used to represent RS-D data uncertainties (see Sect. 4.2) and/or the different flooding characteristics of the study sites. In fact, Giustarini et al. (2011) hypothesised that in larger river systems the dominating effect of the boundary condition is reduced and this favours more persistent model improvements.

García-Pintado et al. (2013, 2015) argued that, for storm–flood event durations and in the absence of discharge data, the low current satellite revisit frequency hampers the operational implementation of more complex approaches than a constant error forecast model. Consequently, the deviations of the error forecast model from the real error dynamics will diminish the improvement achieved using the data assimilation protocol. In particular, García-Pintado et al. (2013) used a reverse approach in a synthetic scenario to show that the speed at which the updated model predictions drift away from the truth after the analysis step is driven by the lack of match between the used inflow error forecast model and the true inflow error evolution.

It is important to underline that Andreadis et al. (2007), Matgen et al. (2010), Giustarini et al. (2011, 2012) and García-Pintado et al. (2013) all assumed the hydraulic model to be perfect in its structure, implementation and parameters, while the boundary inflows were considered as the only source of error. An attempt towards parameter updating was performed by García-Pintado et al. (2015). Using the real case scenario of the 2007 flood event in the Severn and Avon rivers (UK), they investigated whether, with an imminent flood situation, it is more effective to focus on state updating; joint state and inflow updating; or joint state, input and parameter updating. In the latter hypothesis, state values (water levels), inputs and uncertain friction and bathymetry are estimated at the same time. In the studied case, despite a benefit in the development of forecast error covariances and a consistent convergence of the parameters, the simultaneous parameter updating did not improve the flood forecast skill. However, the authors suggested that an accurate tuning of the localisation parameters used for bathymetry estimation in the frame of the local EnKF is likely to lead to an improved feedback on the flood forecasts.

7.3 Performance Measures

The capability of the proposed RS DA strategies in improving the accuracy of hydraulic models for flood forecasting has been evaluated by comparing the results of the analysis and of the forecast steps against "true" (i.e. observed or synthetic) values of water levels/discharge at individual gauge stations. In particular, the effectiveness of the analysis step in reducing the discrepancy between the forecast and the "truth" is shown by plotting the histogram of the a priori and a posteriori water level/discharge value ensemble at specific locations along the river reach and quantified by the mean error value. The latter is the change in distance between the mean of the a priori histogram and the "truth" compared to the distance between the mean of the a posteriori histogram and the truth. The effectiveness of the analysis step in reducing the uncertainty of the predicted ensemble of water level and/or discharge values is measured by computing the standard deviation of the a priori and a posteriori histograms.

The effectiveness of the forecast step in improving the predictions of a hydraulic model is generally quantified by the time-averaged RMSE between the modelled and the "true" time series of water level/discharge values at discrete points along the river reach. Some studies (e.g., Giustarini et al. 2012) limited the assessment to a visual comparison between predicted and "true" time series of water levels/discharge values. García-Pintado et al. (2013) used the RMSE, the Brier Skill Score (BSS) and rank histograms. The BSS was used to evaluate the accuracy of forecast when compared to the open-loop simulation. The rank histograms were used for determining the reliability of ensemble forecasts and for the diagnosis of errors in its mean and spread. Andreadis and Schumann (2014) defined a time and spatially averaged cost function to measure the reduction in the forecast error due to the assimilation when compared to an open-loop simulation.

The errors in the timing and value of the flood peak have so far not been included in the assessment of the effectiveness of RS DA in enhancing the forecast skills of a hydraulic model. Furthermore, as explained in Sect. 6.5, checking the skills of a hydraulic model for flood forecasting exclusively at local level might not be representative for the model behaviour at catchment level.

7.4 Utility of Remote Sensing-Derived Observations of Flood Extent and Level in Contrast to Field Data

A few studies contrasted the use of RS-D and field observations in a data assimilation exercise. Neal et al. (2009) showed that, over a 10-km domain of the Alzette River, water levels derived from a 25-m resolution ENVISAT-ASAR image taken four hours after the 2003 flood peak were able to provide similar reductions in water level uncertainty as two wrack marks having a standard deviation below 0.04 m. This result demonstrated that, even when a dense and accurate network of ground measurements is available (it is worth noting here that gauging stations in Europe are usually constructed every 10–60 km), the use of a RS image is appealing.

Giustarini et al. (2011) performed a benchmark test to contrast the improvements in flood forecasting accuracy obtained when assimilating, respectively, very precise but poorly distributed ground-surveyed information or spatially distributed but highly uncertain RS-D water level data. For the 2003 flood event in the Alzette River, hydrometric data with a standard deviation of 0.1 m were recorded by six gauge stations; RS-D water levels having a mean uncertainty of ±0.54 cm (Hostache et al. 2009) were retrieved from two subsequent 25-m resolution ERS2 and ENVISAT-ASAR images. In situ data were assimilated only at the time steps of the satellite overpasses. When a global weighting approach was used, the two data sets showed similar performances in reducing the a priori uncertainty, and the a posteriori distribution of the water levels generally encompassed the truth. However, when a local weighting approach was used, the ground data outperformed the RS-D data due to their higher accuracy. Nevertheless, Giustarini et al. (2011) highlighted that a local weighting approach in combination with precise but poorly spatially distributed field data can potentially lead to an over-correction of models as the performance of a model at a local level might not be truly representative for its behaviour at a regional level. In this case, the assimilation can even lead to a deterioration in model performance.

In summary, RS-D data are not yet a substitute for a dense in situ gauge network, as they are only available infrequently and with a lesser accuracy than in situ gauge data. However, RS-D data provide added value through their spatially distributed information. The combination of both RS-D and field data sets is therefore likely to yield the best assimilation results.

7.5 Importance of Remote Sensing Observations Resolution, Acquisition Time and Coverage

7.5.1 Resolution

The need for a rapid dissemination of information is probably of greater importance than the production of a high-resolution product (e.g., Blyth 1997). Given the strong inverse relationship between spatial resolution and revisit time for satellites (e.g., Di Baldassarre et al. 2011), many authors investigated the possibility of assimilating data from coarse satellite imagery into hydraulic models. Synthetic experiments completed by Matgen et al. (2010) and the analysis of real case studies performed by Neal et al. (2009) and Giustarini et al. (2011, 2012) showed that forecasting improvements can be achieved using the existing coarse, low accuracy SAR data. In particular, Matgen et al. (2010) hypothesised standard deviations of observed water levels from 0.1 m up to 10 m in order to investigate the performance of the assimilation scheme as a function of observation accuracy. The rather satisfactory results obtained with standard deviations up to 5 m indicated that the assimilation of data having errors typical of coarse-resolution imagery (Di Baldassarre et al. 2011) can lead to forecast improvements.

Giustarini et al. (2012) used the full empirical histogram distribution of the water levels derived from the overlapping of a 150-m resolution ENVISAT-ASAR Wide Swath Mode image with the global 90-m resolution SRTM DEM. Given the coarse resolution of both the data set and the overestimation of the actual river bathymetry, the RS-D water levels displayed a wide spread of values for each cross section and even included values lower than the bed level. Moreover, due to the underestimation of the actual shape of the flow area, for a given input discharge, the modelled water levels overestimated the true water levels. Nevertheless, after the assimilation step, the reduction in the forecast error was relevant.

7.5.2 First Visit Time and Frequency of Acquisition

Knowledge of the sensitivity of the forecasting accuracy to the first visit time and to the subsequent acquisition frequency is an important design parameter in a time limited (and often resource-limited) emergency scenario. These questions were investigated by ad hoc synthetic experiments. Andreadis et al. (2007) analysed the impact of the assimilation frequency on the modelling of both low-flow and high-flow periods. The synthetic RS-D water levels were representative of the observations that will be provided by the SWOT altimeter and the 8-day planned overpass interval of SWOT for the Ohio River (USA) was used as a benchmark. Two additional experiments with overpass intervals of 16 and 32 days were conducted. All the filter simulations performed better than the open-loop simulation; however, as expected, the assimilation system performance degraded as the observation frequency became sparser. The time-averaged discharge and channel water level RMSE of the open-loop simulation were 23.2 % and 56 cm. These numbers were reduced to 10, 12.1 and 16.9 % and to 21.6, 24.9 and 33.3 cm, when RS-D data are assimilated with a frequency of 8, 16 and 32 days.

The timing of acquisition, rather than its frequency, is pivotal in improving flood events forecast. Matgen et al. (2010) showed that the water level RMSE averaged over the total length of a flood event decreased from the open-loop simulation value of 0.34 m to 0.18, 0.19 and 0.25 m when RS-D water level data were acquired with a frequency of 12, 24 and

48 h, respectively, and the first visit time was before the flood peak at the upstream cross section. In fact, a higher frequency was effective in improving the forecast skills only when acquisitions were performed during the rising limb and until shortly after the flood peak. On the contrary, no significant positive effect was obtained by an increased frequency during the receding limb. Analogously, the synthetic experiments performed by García-Pintado et al. (2013) showed that imagery obtained early in the flood had a large influence on forecast statistics, and revisit interval was most influential for early observations. The time-averaged discharge RMSE when the first acquisition happened during the rising limb was roughly equal to the 35 % of the time-averaged discharge RMSE when the first acquisition happened during the receding limb, while the frequency of acquisition appeared to be a secondary factor. Per contra, in a real case scenario, Giustarini et al. (2011) denounced the low persistence of the improvements of the flood forecast skills achieved using two RS images acquired before and immediately after the flood peak, respectively. Nevertheless, in this case study, not even the use of highly accurate field data assimilated at the same acquisition time of the RS images could enhance the performance of the forecast model and a poor calibration was likely the cause of such a lack of improvements (Sect. 7.7).

In summary, the required first visit and revisit times depend on the temporal correlation of model errors. During the hydrograph's rising limb errors are difficult to predict as precipitation errors continuously add to model parameter and model structural errors (e.g., Giustarini et al. 2011). The operational scheduling of satellite SAR acquisitions should try to capture the early stages of the rising limb, possibly with the highest available observation frequency. After the flood peak, it becomes economically and computationally convenient to distribute the observations in time. This kind of strategy should enable the forecast to be kept on track for a longer time at minimum cost (García-Pintado et al. 2013).

7.5.3 Coverage

In the analyses described so far, the RS observation provided a full coverage of the model domain. However, depending on the specific orbit, the spatial coverage of satellite observations might be partial, and examining the impact of observations that are not spatially coincident with the forecast reach becomes of interest. Biancamaria et al. (2011), García-Pintado et al. (2013) and Andreadis and Schumann (2014) showed that the benefit of assimilating upstream observations at an early overpass is propagated downstream and can be highly influential. The experiment completed by Andreadis and Schumann (2014) suggested that the sequential assimilation of synthetic RS-D observations of flood extents having limited coverage can be useful to investigate errors and uncertainties in the model implementation and structure. Moving from upstream to downstream of the Ohio River (USA), they assimilated observations covering gradually increasing areas of the model domain. By extending the spatial coverage of their synthetic observations downstream, they were able to point out the locations where the assimilated observations actually degraded the forecast. A further investigation revealed large errors in the DEM (3.5 and 4.7 m) at those locations. It was also proven that, because of the negative impact of the assimilation at these locations, the assimilation of observations covering a much larger area is required to recover the same level of forecast accuracy that could be achieved assimilating more accurate observations over a smaller area.

7.6 Selection of the Remote Sensing-Derived Product

All studies listed above assimilated RS-D water levels. Andreadis and Schumann (2014) compared the impact of assimilating RS-D observations of water level, top width and inundated area on the forecasting skills of a hydraulic model of the Ohio River (USA). In a synthetic scenario, a simple state updating approach was applied during both low- and high-flow periods and for lead times from 1 to 11 days. The assimilation of RS-D water levels improved the prediction of water levels for all lead times. The impacts of the assimilation of RS-D water level and top width values on discharge prediction were then evaluated. The imperfect knowledge of channel bathymetry and the fact that discharge is mostly governed by the boundary inflows forcing (i.e. the low temporal autocorrelation of discharge) render discharge prediction highly challenging. Despite those difficulties, the impact of the RS-D water level observations on the discharge prediction skills was positive for lead times up to 7 days over the entire study period. In contrast, the assimilation of top width observations actually degraded the forecast skills on average (both spatially and temporally). The impact of the DA exercise on the forecast skill is partly dependent on the covariance between the forecast variable (i.e. discharge) and the observations (i.e. top width). For channel cross sections that are close to rectangular, top width does not vary significantly despite large variations in discharge (and water depth) and the information content of the width observations is limited. Finally, the impact of the assimilation of RS-D flooded area was a function of the acquisition time (Sect. 7.5.2) and showed a short persistence due to the quick variation of the inflow conditions during a flood event.

7.7 Importance of Hydraulic Models Implementation and Calibration

The analysis of real case scenarios performed by Neal et al. (2009), Giustarini et al. (2011, 2012) and Andreadis and Schumann (2014) showed that, despite the fact that the RS DA improved the accuracy of the flood forecasts, the impact of uncertainties in the implementation, calibration or model structure was still significant, especially at the local level. In particular, Neal et al. (2009), Giustarini et al. (2012) and Andreadis and Schumann (2014) tested the potential value of satellite observations for the reduction in flood forecasting uncertainty when knowledge of topography, and river bathymetry is highly uncertain as only coarse data are available. Neal et al. (2009) obtained rather satisfactory results comparing the accuracy of a model of the 2003 flood in the Alzette River when either LiDAR data were used to model the cross sections, or simplified channel geometry was used. In fact, the same real case study was also analysed by Giustarini et al. (2011) using a different calibration, a different inflow ensemble and a different filter. In the latter study, even according to the best case scenario, i.e. when precise in situ measurements were assimilated into a hydraulic model, difficulties raised from the fact that model accuracy varies in space, and the poor quality of the forecast results at some cross sections could be explained by a poorly calibrated model. When only a global coarse data set for model implementation were available, Giustarini et al. (2012) and Andreadis and Schumann (2014) showed that, despite its effectiveness in correcting input errors and modelled water levels, the RS DA protocol could not impede the underestimation of river discharge due to the discrepancy between the actual river bathymetry and the bathymetry derived from the global database. Based on these analyses, it can be concluded that the set-up and calibration of the hydraulic model are paramount prior to any data assimilation exercise. Nevertheless, RS DA can provide a valuable tool to identify weaknesses in the model structure or local errors in the model implementation/parameterisation. For instance, based on the results shown by Andreadis and Schumann (2014), ad hoc experiments consisting of the assimilation of (synthetic) remote sensing data having a partial coverage can allow the identification of the most "valuable" (i.e. improving forecast skill) observation locations as well as locations where the assimilated observations may lead to a persistent degradation of the forecast skill. In a resource-limited scenario, this analysis is a valuable tool as it points out the locations where assimilation is pivotal (as they benefit the flood forecasts) as well as the areas/locations requiring refined implementation data and parameter assessment.

8 Discussion and Conclusions

Based on this literature review, there is no doubt that the use of remote sensing data has the potential to become a critical component in flood forecasting systems, more specifically for the calibration, validation and real-time constraint of hydraulic models.

The number of space missions which are able to provide coarse- to high-resolution RS imagery is currently increasing, and the observation of flood events from a remote location is likely to shift from being opportunistic to being strategic. RS techniques provide a large amount of data, which can solve the well know issue of scarcity of spatially distributed data for flood inundation modelling (e.g., Bates et al. 1997, 2004; Horritt 2000; Werner et al. 2005). Such data availability requires a better understanding on how to effectively exploit the information content to improve the accuracy of hydraulic models for flood forecasting. In fact, more detailed information might not directly lead to improved model performance as different limitations and uncertainty sources affect flood inundation modelling. These can be categorised as: (1) model structural uncertainty, which is originated by the inability of models to perfectly schematize the (often incompletely understood) physical processes involved; (2) inflow uncertainty; (3) scarcity and uncertainty in the observations used for calibration, validation and real-time constraint; (4) inadequacy of calibration, validation and data assimilation protocols; and (5) time-changing physical characteristics of the flooded area (e.g., Di Baldassarre and Uhlenbrook 2012; Dottori et al. 2013; Bates et al. 2014b).

The published literature demonstrated that the availability of RS-D observations of flood extent and water levels have the potential to reduce some of the above listed sources of uncertainty. The main conclusions and research needs pointed out during the last two decades of scientific efforts are listed here with the aim of paving the way for the development of improved protocols for flood inundation prediction.

- RS-D observations of flooded area and water levels are currently a complement, not an alternative to field data for hydraulic model calibration/validation and real-time constraint. The main drawbacks of RS-D data are the low precision, accuracy and acquisition frequency.
- Further research on RS image processing algorithms is required in order to improve accuracy and precision and reduce the delivery time of RS-D data.
- The quantification of RS-D data uncertainty is pivotal for any calibration, validation and DA exercise.
- The choice of the most appropriate RS-D observation (i.e. inundation extent or level) and of the advisable acquisition first visit and revisit time is tied to the morphological and flooding characteristics of the catchment. A large variability of the observed quantities within the flooding process and a strong correlation between the observed

quantity and flood dynamics is pivotal in determining the success of the calibration, validation, and data assimilation protocol. Early acquisitions are recommended as during the hydrograph's rising limb small changes in the channel roughness coefficient may have a large impact on the simulated extent and water levels. Observations of flood extent at flood peak conditions are advisable for model calibration and validation in flat, large floodplains. Conversely, in V-shaped valleys affected by "valley filling" events, flood extents show little sensitivity to high discharge values. In these cases, observations of water levels at a lower flow than peak may provide a better test of model performances. In real-time forecast, precipitation errors continuously add to model parameter and model structural errors; first visit time and increased acquisition frequency during the rising limb until the flood peak proved to be more effective in improving the models' forecast skill.

- The implementation of multi-objective calibration and validation frameworks is advised and the definition of consistent performance measures is required.
- The selection of the appropriate data assimilation algorithm is an open question. It is worth noting that no study so far has applied different DA algorithms (for instance the EnKF and the PF) to the same case study. Simple state updating methods improve the forecast for a short time. Input updating has been based so far on very simple models and more appropriate approaches have been advocated. Future research should look at the combination of state updating, input updating and parameter updating.
- Uncertainties in the upstream boundary conditions have been identified has one of the main sources of errors in flood forecasting (e.g., Bates et al. 2014b). In all the studies analysed, these uncertainties were propagated through the system using ensemble input forecasts retrieved from hydrologic models or uncertainty in the rating curve of gauged hydrographs. The propagation of errors and uncertainties through coupled hydrologic and hydraulic models has to be investigated.

In summary, further research in image processing, description of uncertainty, calibration–validation and data assimilation protocols, performance measures, error propagation are required to make full use of the new RS data sets. Nevertheless, limitations and uncertainty sources will always affect flood inundation modelling; the models are a representation of the real word, and the real word might undergo sensitive changes over the years or during the flood event itself. Consequently, possible uncertainty sources should be always identified and evaluated, for instance, through multiple scenario analyses.

Acknowledgments This study is financially supported by the Bushfires and Natural Hazards CRC project—Improving flood forecast skill using remote sensing data. Valentijn Pauwels is funded by ARC Future Fellow grant FT130100545. The authors would like to acknowledge the Australian Bureau of Meteorology and Geoscience Australia for their valuable comments and support.

References

Adhikari P, Hong Y, Douglas K, Kirschbaum D, Gourley J, Adler R, Robert Brakenridge G (2010) A digitized global flood inventory (1998–2008): compilation and preliminary results. Nat Hazards 55:405–422. doi:10.1007/s11069-010-9537-2

Alsdorf DE (2002) Interferometric SAR observations of water level changes: potential targets for future repeat-pass AIRSAR missions. In: Paper presented at the AIRSAR earth science and application workshop, Pasadena, California, March 4–6 2002

Alsdorf DE, Melack JM, Dunne T, Mertes LAK, Hess LL, Smith LC (2000) Interferometric radar measurements of water level changes on the Amazon flood plain. Nature 404:174–177

Alsdorf DE, Smith LC, Melack JM (2001) Amazon floodplain water level changes measured with interferometric SIR-C radar. Geosci Remote Sens IEEE Trans 39:423–431. doi:10.1109/36.905250

Alsdorf D, Dunne T, Melack J, Smith L, Hess L (2005) Diffusion modeling of recessional flow on central Amazonian floodplains. Geophys Res Lett 32:L21405. doi:10.1029/2005GL024412

Alsdorf D, Bates P, Melack J, Wilson M, Dunne T (2007) Spatial and temporal complexity of the Amazon flood measured from space. Geophys Res Lett 34:L08402. doi:10.1029/2007GL029447

Andreadis KM, Schumann GJP (2014) Estimating the impact of satellite observations on the predictability of large-scale hydraulic models. Adv Water Resour 73:44–54. doi:10.1016/j.advwatres.2014.06.006

Andreadis KM, Clark EA, Lettenmaier DP, Alsdorf DE (2007) Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. Geophys Res Lett 34:L10403. doi:10.1029/2007GL029721

Aronica G, Hankin B, Beven K (1998) Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Adv Water Resour 22:349–365. doi:10.1016/S0309-1708(98)00017-7

Aronica G, Bates PD, Horritt MS (2002) Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. Hydrol Process 16:2001–2016. doi:10.1002/hyp.398

Bamler R, Hartl P (1998) Synthetic aperture radar interferometry. Inverse Prob 14:R1

Bates PD (2004) Remote sensing and flood inundation modelling. Hydrol Process 18:2593–2597. doi:10.1002/hyp.5649

Bates PD, Horritt MS, Smith CN, Mason D (1997) Integrating remote sensing observations of flood hydrology and hydraulic modelling. Hydrol Process 11:1777–1795. doi:10.1002/(SICI)1099-1085(199711)11:14<1777:AID-HYP543>3.0.CO;2-E

Bates PD, Stewart MD, Siggers GB, Smith AM, Hervouet JM, Sellin RHJ (1998) Internal and external validation of a two-dimensional finite element code for river flood simulations. Proc Inst Civ Eng Water Marit Energy 130:127–141. doi:10.1680/iwtme.1998.30972

Bates PD, Horritt MS, Aronica G, Beven K (2004) Bayesian updating of flood inundation likelihoods conditioned on flood extent data. Hydrol Process 18:3347–3370. doi:10.1002/hyp.1499

Bates PD, Wilson MD, Horritt MS, Mason DC, Holden N, Currie A (2006) Reach scale floodplain inundation dynamics observed using airborne synthetic aperture radar imagery: data analysis and modelling. J Hydrol 328:306–318. doi:10.1016/j.jhydrol.2005.12.028

Bates P, Neal J, Alsdorf D, Schumann GP (2014a) Observing global surface water flood dynamics. Surv Geophys 35:839–852. doi:10.1007/s10712-013-9269-4

Bates PD, Pappenberger F, Romanowicz RJ (2014b) Uncertainty in flood inundation modelling. In: Applied uncertainty analysis for flood risk management. Imperial College Press, London (UK), pp 232–269. doi:10.1142/9781848162716_0010

Bazi Y, Bruzzone L, Melgani F (2005) An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images. Geosci Remote Sens IEEE Trans 43:874–887. doi:10.1109/TGRS.2004.842441

Bercher N, Kosuth P (2012) Monitoring river water levels from space: quality assessment of 20 years of satellite altimetry data. In: Paper presented at the 20 years of progress in radar altimetry, Venice (Italy), 24–29 September 2012

Beven K (2006) A manifesto for the equifinality thesis. J Hydrol 320:18–36. doi:10.1016/j.jhydrol.2005.07.007

Beven K, Binley A (1992) The future of distributed models: model calibration and uncertainty prediction. Hydrol Process 6:279–298. doi:10.1002/hyp.3360060305

Biancamaria S, Bates PD, Boone A, Mognard NM (2009) Large-scale coupled hydrologic and hydraulic modelling of the Ob river in Siberia. J Hydrol 379:136–150. doi:10.1016/j.jhydrol.2009.09.054

Biancamaria S et al (2011) Assimilation of virtual wide swath altimetry to improve Arctic river modeling. Remote Sens Environ 115:373–381. doi:10.1016/j.rse.2010.09.008

Biancamaria S, Lettenmaier DP, Pavelsky TM (2015) The SWOT mission and its capabilities for land hydrology. Surv Geophys. doi:10.1007/s10712-015-9346-y

Birkett CM, Mertes LAK, Dunne T, Costa MH, Jasinski MJ (2002) Surface water dynamics in the Amazon Basin: application of satellite radar altimetry. J Geophys Res Atmos 107:LBA 26-21. doi:10.1029/2001JD000609

Blyth KEN (1997) Floodnet: a telenetwork for acquisition, processing and dissemination of earth observation data for monitoring and emergency management of floods. Hydrol Process 11:1359–1375. doi:10.1002/(SICI)1099-1085(199708)11:10<1359:AID-HYP529>3.0.CO;2-6

Brandimarte L, Brath A, Castellarin A, Baldassarre GD (2009) Isla Hispaniola: a trans-boundary flood risk mitigation plan. Phys Chem Earth A/B/C 34:209–218. doi:10.1016/j.pce.2008.03.002

Brisco B, Kapfer M, Hirose T, Tedford B, Liu J (2011) Evaluation of C-band polarization diversity and polarimetry for wetland mapping. Can J Remote Sens 37:82–92. doi:10.5589/m11-017

Brivio PA, Colombo R, Maggi M, Tomasoni R (2002) Integration of remote sensing data and GIS for accurate mapping of flooded areas. Int J Remote Sens 23:429–441. doi:10.1080/01431160010014729

Calabresi G (1995) The use of ERS SAR for flood monitoring: an overall assessment. In: Paper presented at the 2nd ERS applications workshop, London, UK, 6–8 December

Charney J, Halem M, Jastrow R (1969) Use of incomplete historical data to infer the present state of the atmosphere. J Atmos Sci 26:1160–1163. doi:10.1175/1520-0469(1969)026<1160:UOIHDT>2.0.CO;2

Choudhury BJ (1989) Monitoring global land surface using Nimbus-7 37 GHz data theory and examples. Int J Remote Sens 10:1579–1605. doi:10.1080/01431168908903993

Coxon G, Freer J, Westerberg IK, Wagener T, Woods R, Smith PJ (2015) A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. Water Resour Res 51:5531–5546. doi:10.1002/2014WR016532

De Groeve T (2010) Flood monitoring and mapping using passive microwave remote sensing in Namibia. Geomat Nat Hazards Risk 1:19–35. doi:10.1080/19475701003648085

De Roo A, Van Der Knijff J, Horritt MS, Schmuck G, De Jong S (1999) Assessing flood damages of the 1997 Oder flood and the 1995 Meuse flood. In: Paper presented at the 2nd international ITC symposium on operationalization of remote sensing, Enschede, The Netherlands

Delmeire S (1997) Use of ERS-1 data for the extraction of flooded areas. Hydrol Process 11:1393–1396. doi:10.1002/(SICI)1099-1085(199708)11:10<1393:AID-HYP528>3.0.CO;2-N

Di Baldassarre G, Montanari A (2009) Uncertainty in river discharge observations: a quantitative analysis. Hydrol Earth Syst Sci 13:913–921. doi:10.5194/hess-13-913-2009

Di Baldassarre G, Uhlenbrook S (2012) Is the current flood of data enough? A treatise on research needs for the improvement of flood modelling. Hydrol Process 26:153–158. doi:10.1002/hyp.8226

Di Baldassarre G, Schumann G, Bates P (2009a) Near real time satellite imagery to support and verify timely flood modelling. Hydrol Process 23:799–803. doi:10.1002/hyp.7229

Di Baldassarre G, Schumann G, Bates PD (2009b) A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent. J Hydrol 367:276–282. doi:10.1016/j.jhydrol.2009.01.020

Di Baldassarre G, Schumann G, Brandimarte L, Bates P (2011) Timely low resolution SAR imagery to support floodplain modelling: a case study review. Surv Geophys 32:255–269. doi:10.1007/s10712-011-9111-9

Domeneghetti A, Castellarin A, Brath A (2012) Assessing rating-curve uncertainty and its effects on hydraulic model calibration. Hydrol Earth Syst Sci 16:1191–1202. doi:10.5194/hess-16-1191-2012

Domeneghetti A, Tarpanelli A, Brocca L, Barbetta S, Moramarco T, Castellarin A, Brath A (2014) The use of remote sensing-derived water surface data for hydraulic model calibration. Remote Sens Environ 149:130–141. doi:10.1016/j.rse.2014.04.007

Domeneghetti A, Castellarin A, Tarpanelli A, Moramarco T (2015) Investigating the uncertainty of satellite altimetry products for hydrodynamic modelling. Hydrol Process 29:4908–4918. doi:10.1002/hyp.10507

Dottori F, Di Baldassarre G, Todini E (2013) Detailed data is welcome, but with a pinch of salt: accuracy, precision, and uncertainty in flood inundation modeling. Water Resour Res 49:6079–6085. doi:10.1002/wrcr.20406

Dung NV, Merz B, Bárdossy A, Thang TD, Apel H (2011) Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data. Hydrol Earth Syst Sci 15:1339–1354. doi:10.5194/hess-15-1339-2011

Dunne S, Entekhabi D (2005) An ensemble-based reanalysis approach to land data assimilation. Water Resour Res 41:W02013. doi:10.1029/2004WR003449

Durand JM, Gimonet BJ, Perbos JR (1987) SAR data filtering for classification. Geosci Remote Sens IEEE Trans 25:629–637. doi:10.1109/TGRS.1987.289842

Evans TL, Costa M, Telmer K, Silva TSF (2010) Using ALOS/PALSAR and RADARSAT-2 to map land cover and seasonal inundation in the Brazilian Pantanal. Sel Top Appl Earth Obs Remote Sens IEEE J 3:560–575. doi:10.1109/JSTARS.2010.2089042

Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res Oceans 99:10143–10162. doi:10.1029/94JC00572

Evensen G (2003) The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn 53:343–367. doi:10.1007/s10236-003-0036-9

Evensen G (2004) Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn 54:539–560. doi:10.1007/s10236-004-0099-2

Faruolo M, Coviello I, Lacava T, Pergola N, Tramutoli V (2009) Real time monitoring of flooded areas by a multi-temporal analysis of optical satellite data. In: Geoscience and remote sensing symposium, 2009 IEEE international, IGARSS 2009, 12–17 July 2009. pp IV-192–IV-195. doi:10.1109/IGARSS.2009.5417339

Franceschetti G, Iodice A, Riccio D (2002) A canonical problem in electromagnetic backscattering from buildings. IEEE Trans Geosci Remote Sens 40:1787–1801. doi:10.1109/TGRS.2002.802459

Frappart F, Calmant S, Cauhopé M, Seyler F, Cazenave A (2006) Preliminary results of ENVISAT RA-2-derived water levels validation over the Amazon basin. Remote Sens Environ 100:252–264. doi:10.1016/j.rse.2005.10.027

Frost VS, Stiles JA, Shanmugam KS, Holtzman JC, Smith SA (1981) An adaptive filter for smoothing noisy radar images. Proc IEEE 69:133–135. doi:10.1109/PROC.1981.11935

Frost VS, Stiles JA, Shanmugan KS, Holtzman J (1982) A model for radar images and its application to adaptive digital filtering of multiplicative noise. Pattern Anal Mach Intell IEEE Trans 4:157–166. doi:10.1109/TPAMI.1982.4767223

Fu L-L (2001) Chapter 3.3 Ocean circulation and variability from satellite altimetry. In: Gerold Siedler JC, John G (eds) International geophysics, vol 77. Academic Press, Cambridge, Massachusetts (USA), pp 141-XXVIII. doi:10.1016/S0074-6142(01)80116-9

Furrer R, Bengtsson T (2007) Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. J Multivar Anal 98:227–255. doi:10.1016/j.jmva.2006.08.003

García-Pintado J, Neal JC, Mason DC, Dance SL, Bates PD (2013) Scheduling satellite-based SAR acquisition for sequential assimilation of water level observations into flood modelling. J Hydrol 495:252–266. doi:10.1016/j.jhydrol.2013.03.050

García-Pintado J, Mason DC, Dance SL, Cloke HL, Neal JC, Freer J, Bates PD (2015) Satellite-supported flood forecasting in river networks: a real case study. J Hydrol 523:706–724. doi:10.1016/j.jhydrol.2015.01.084

Giustarini L et al (2011) Assimilating SAR-derived water level data into a hydraulic model: a case study. Hydrol Earth Syst Sci Discuss 8:2103–2144. doi:10.5194/hessd-8-2103-2011

Giustarini L, Matgen P, Hostache R, Dostert J (2012) From SAR-derived flood mapping to water level data assimilation into hydraulic models. In: Neale CMU, Maltese A (eds) Remote sensing for agriculture, ecosystems, and hydrology XIV, Edinburgh (UK), 24–26 September. Proceedings of SPIE, pp 85310U–85310U–85312. doi:10.1117/12.974655

Giustarini L, Hostache R, Matgen P, Schumann GJP, Bates PD, Mason DC (2013) A change detection approach to flood mapping in urban areas using TerraSAR-X. IEEE Trans Geosci Remote Sens 51:2417–2430. doi:10.1109/tgrs.2012.2210901

Giustarini L et al (2015) Accounting for image uncertainty in SAR-based flood mapping. Int J Appl Earth Obs Geoinf 34:70–77. doi:10.1016/j.jag.2014.06.017

Grayson R, Blöschl G (2001) Spatial patterns in catchment hydrology: observations and modelling. Cambridge University Press, Cambridge

Gupta HV, Sorooshian S, Yapo PO (1998) Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. Water Resour Res 34:751–763. doi:10.1029/97WR03495

Hagen A (2003) Fuzzy set approach to assessing similarity of categorical maps. Int J Geogr Inf Sci 17:235–249. doi:10.1080/13658810210157822

Hall J, Tarantola S, Bates P, Horritt M (2005) Distributed sensitivity analysis of flood inundation model calibration. J Hydraul Eng 131:117–126. doi:10.1061/(ASCE)0733-9429(2005)131:2(117)

Hall JW, Manning LJ, Hankin RKS (2011) Bayesian calibration of a flood inundation model using spatial data. Water Resour Res 47:W05529. doi:10.1029/2009WR008541

Hall AC, Schumann GJP, Bamber JL, Bates PD, Trigg MA (2012) Geodetic corrections to Amazon River water level gauges using ICESat altimetry. Water Resour Res 48:W06602. doi:10.1029/2011WR010895

Hamilton SK, Sippel SJ, Melack JM (1996) Inundation patterns in the Pantanal wetland of South America determined from passive microwave remote sensing. Arch Hydrobiol 137:1–23

Hartmann HC, Pagano TC, Sorooshian S, Bales R (2002) Confidence builders: evaluating seasonal climate forecasts from user perspectives. Bull Am Meteorol Soc 83:683–698. doi:10.1175/1520-0477(2002)083<0683:CBESCF>2.3.CO;2

Henderson FM, Lewis AJ (2008) Radar detection of wetland ecosystems: a review. Int J Remote Sens 29:5809–5835. doi:10.1080/01431160801958405

Henry JB, Chastanet P, Fellah K, Desnos YL (2006) Envisat multi-polarized ASAR data for flood mapping. Int J Remote Sens 27:1921–1929. doi:10.1080/01431160500486724

Hess LL, Melack JM, Simonett DS (1990) Radar detection of flooding beneath the forest canopy: a review. Int J Remote Sens 11:1313–1325. doi:10.1080/01431169008955095

Hong S-H, Wdowinski S, Sang-Wan K (2010) Evaluation of TerraSAR-X observations for wetland InSAR application. Geosci Remote Sens IEEE Trans 48:864–873. doi:10.1109/TGRS.2009.2026895

Horritt M (1999) A statistical active contour model for SAR image segmentation. Image Vis Comput 17:213–224. doi:10.1016/S0262-8856(98)00101-2

Horritt MS (2000) Calibration of a two-dimensional finite element flood flow model using satellite radar imagery. Water Resour Res 36:3279–3291. doi:10.1029/2000WR900206

Horritt MS (2006) A methodology for the validation of uncertain flood inundation models. J Hydrol 326:153–165. doi:10.1016/j.jhydrol.2005.10.027

Horritt MS, Bates PD (2002) Evaluation of 1D and 2D numerical models for predicting river flood inundation. J Hydrol 268:87–99. doi:10.1016/S0022-1694(02)00121-X

Horritt MS, Mason DC, Cobby DM, Davenport IJ, Bates PD (2003) Waterline mapping in flooded vegetation from airborne SAR imagery. Remote Sens Environ 85:271–281. doi:10.1016/S0034-4257(03)00006-3

Horritt MS, Di Baldassarre G, Bates PD, Brath A (2007) Comparing the performance of a 2-D finite element and a 2-D finite volume model of floodplain inundation using airborne SAR imagery. Hydrol Process 21:2745–2759. doi:10.1002/hyp.6486

Hostache R, Matgen P, Schumann G, Puech C, Hoffmann L, Pfister L (2009) Water level estimation and reduction of hydraulic model calibration uncertainties using satellite SAR images of floods. Geosci Remote Sens IEEE Trans 47:431–441. doi:10.1109/TGRS.2008.2008718

Hostache R, Lai X, Monnier J, Puech C (2010) Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: use of a remote sensing image of Mosel River. J Hydrol 390:257–268. doi:10.1016/j.jhydrol.2010.07.003

Houser P, De Lannoy GM, Walker J (2010) Land surface data assimilation. In: Lahoz W, Khattatov B, Menard R (eds) Data assimilation. Springer, Berlin, pp 549–597. doi:10.1007/978-3-540-74703-1_21

Houtekamer PL, Mitchell HL (2005) Ensemble Kalman filtering. Q J R Meteorol Soc 131:3269–3289. doi:10.1256/qj.05.135

Hunt BR, Kostelich EJ, Szunyogh I (2007) Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. Phys D 230:112–126. doi:10.1016/j.physd.2006.11.008

Hunter NM, Bates PD, Horritt MS, De Roo APJ, Werner MGF (2005) Utility of different data types for calibrating flood inundation models within a GLUE framework. Hydrol Earth Syst Sci 9:412–430. doi:10.5194/hess-9-412-2005

Hunter NM, Bates PD, Horritt MS, Wilson MD (2006) Improved simulation of flood flows using storage cell models. Proc Inst Civil Eng Water Manag 159:9–18. doi:10.1680/wama.2006.159.1.9

Jin YQ (1999) A flooding index and its regional threshold value for monitoring floods in China from SSM/I data. Int J Remote Sens 20:1025–1030. doi:10.1080/014311699213064

Jung HC et al (2010) Characterization of complex fluvial systems using remote sensing of spatial and temporal water level variations in the Amazon, Congo, and Brahmaputra rivers. Earth Surf Proc Land 35:294–304. doi:10.1002/esp.1914

Jung HC et al (2012) Calibration of two-dimensional floodplain modeling in the central Atchafalaya Basin Floodway System using SAR interferometry. Water Resour Res 48:W07511. doi:10.1029/2012WR011951

Karim F, Petheram C, Marvanek C, Ticehurst C, Wallace J, Gouweleeuw B (2011) The use of hydrodynamic modelling and remote sensing to estimate floodplain inundation and flood discharge in a large tropical catchment. In: Paper presented at the 19th international congress on modelling and simulation, 12–16 December, Perth (Australia), pp 3796–3802

Kirchgessner P, Nerger L, Bunse-Gerstner A (2014) On the choice of an optimal localization radius in ensemble Kalman filter methods. Mon Weather Rev 142:2165–2175. doi:10.1175/MWR-D-13-00246.1

Kouraev AV, Zakharova EA, Samain O, Mognard NM, Cazenave A (2004) Ob'river discharge from TOPEX/Poseidon satellite altimetry (1992–2002). Remote Sens Environ 93:238–245. doi:10.1016/j.rse.2004.07.007

Kugler Z, De Groeve T (2007) The global flood detection system. JRC scientific and technical reports, pp 45. EUR 23303 EN, ISSN 1018-5593

Lacomme P, Hardange JP, Marchais JC, Normant E (2001) Air and spaceborne radar systems—an introduction, vol Spie Press Monograph (Book 108). William Andrew Publishing/Noyes, NY (USA), pp 524

Lai X, Monnier J (2009) Assimilation of spatially distributed water levels into a shallow-water flood model. Part I: mathematical method and test case. J Hydrol 377:1–11. doi:10.1016/j.jhydrol.2009.07.058

Lai X, Liang Q, Yesou H, Daillet S (2014) Variational assimilation of remotely sensed flood extents using a 2-D flood model. Hydrol Earth Syst Sci 18:4325–4339. doi:10.5194/hess-18-4325-2014

Lang MW, Kasischke ES (2008) Using C-band synthetic aperture radar data to monitor forested wetland hydrology in Maryland's Coastal Plain, USA. Geosci Remote Sens IEEE Trans 46:535–546. doi:10.1109/TGRS.2007.909950

Lee J (1983) A simple speckle smoothing algorithm for synthetic aperture radar images. Syst Man Cybern IEEE Trans 13:85–89. doi:10.1109/TSMC.1983.6313036

Lee J, Jen-Hung W, Ainsworth TL, Kun-Shan C, Chen AJ (2009) Improved sigma filter for speckle filtering of SAR imagery. Geosci Remote Sens IEEE Trans 47:202–213. doi:10.1109/TGRS.2008.2002881

Lee H et al (2011) Characterization of terrestrial water dynamics in the Congo Basin using GRACE and satellite radar altimetry. Remote Sens Environ 115:3530–3538. doi:10.1016/j.rse.2011.08.015

Li Y, Grimaldi S, Walker J, Pauwels V (2016) Application of remote sensing data to constrain operational rainfall-driven flood forecasting: a review. Remote Sens 8:456. doi:10.3390/rs8060456

Liu Y et al (2012) Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. Hydrol Earth Syst Sci 16:3863–3887. doi:10.5194/hess-16-3863-2012

MacIntosh H, Profeti G (1995) The use of ERS SAR data to manage flood emergencies at the smaller scale. In: Paper presented at the 2nd ERS Applications Workshop, London, UK

Madsen H, Skotner C (2005) Adaptive state updating in real-time river flow forecasting—a combined filtering and error forecasting procedure. J Hydrol 308:302–312. doi:10.1016/j.jhydrol.2004.10.030

Marcus WA, Fonstad MA (2008) Optical remote mapping of rivers at sub-meter resolutions and watershed extents. Earth Surf Proc Land 33:4–24. doi:10.1002/esp.1637

Martinis S, Twele A, Voigt S (2009) Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. Nat Hazards Earth Syst Sci 9:303–314

Martinis S, Twele A, Voigt S (2011) Unsupervised extraction of flood-induced backscatter changes in SAR data using markov image modeling on irregular graphs. IEEE Trans Geosci Remote Sens 49:251–263. doi:10.1109/TGRS.2010.2052816

Martinis S, Kersten J, Twele A (2015) A fully automated TerraSAR-X based flood service. ISPRS J Photogramm Remote Sens 104:203–212. doi:10.1016/j.isprsjprs.2014.07.014

Mason DC, Horritt MS, Dall'Amico JT, Scott TR, Bates PD (2007) Improving river flood extent delineation from synthetic aperture radar using airborne laser altimetry. Geosci Remote Sens IEEE Trans 45:3932–3943. doi:10.1109/TGRS.2007.901032

Mason DC, Bates PD, Dall'Amico JT (2009) Calibration of uncertain flood inundation models using remotely sensed water levels. J Hydrol 368:224–236. doi:10.1016/j.jhydrol.2009.02.034

Mason DC, J-p. Schumann G, Bates PD (2010) Data utilization in flood inundation modelling. In: Flood risk science and management. Wiley, New York, pp 209–233. doi:10.1002/9781444324846.ch11

Mason DC, Davenport IJ, Neal JC, Schumann GJP, Bates PD (2012a) Near real-time flood detection in urban and rural areas using high-resolution synthetic aperture radar images. IEEE Trans Geosci Remote Sens 50:3041–3052. doi:10.1109/tgrs.2011.2178030

Mason DC, Schumann GJP, Neal JC, Garcia-Pintado J, Bates PD (2012b) Automatic near real-time selection of flood water levels from high resolution Synthetic Aperture Radar images for assimilation into hydraulic models: a case study. Remote Sens Environ 124:705–716. doi:10.1016/j.rse.2012.06.017

Mason DC, Giustarini L, Garcia-Pintado J, Cloke HL (2014) Detection of flooded urban areas in high resolution synthetic aperture radar images using double scattering. Int J Appl Earth Obs Geoinf 28:150–159. doi:10.1016/j.jag.2013.12.002

Massonnet D, Rossi M, Carmona C, Adragna F, Peltzer G, Feigl K, Rabaute T (1993) The displacement field of the Landers earthquake mapped by radar interferometry. Nature 364:138–142

Matgen P, Henry JB, Pappenberger F, Fraipont PD, Hoffmann L, Pfister L (2004) Uncertainty in calibrating flood propagation models with flood boundaries derived from synthetic aperture radar imagery. In: Proceedings of the 20th congress of the International Society of Photogrammetry and Remote Sensing, 12–23 July, Instanbul, Turkey, pp 352–358

Matgen P, Schumann G, Henry JB, Hoffmann L, Pfister L (2007a) Integration of SAR-derived river inundation areas, high-precision topographic data and a river flow model toward near real-time flood management. Int J Appl Earth Obs Geoinf 9:247–263. doi:10.1016/j.jag.2006.03.003

Matgen P, Schumann G, Pappenberger F, Pfister L (2007b) Sequential assimilation of remotely sensed water stages in flood inundation models. In: Paper presented at the IAHS symposium on remote sensing for environmental monitoring and change detection—24th General Assembly of the International Union of Geodesy and Geophysics (IUGG), Perugia, Italy, 2–13 July

Matgen P et al (2010) Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept. Hydrol Earth Syst Sci 14:1773–1785. doi:10.5194/hess-14-1773-2010

Matgen P, Hostache R, Schumann G, Pfister L, Hoffmann L, Savenije HHG (2011) Towards an automated SAR-based flood monitoring system: lessons learned from two case studies. Phys Chem Earth A/B/C 36:241–252. doi:10.1016/j.pce.2010.12.009

Mitchell AL, Milne AK, Tapley I (2015) Towards an operational SAR monitoring system for monitoring environmental flows in the Macquarie Marshes. Wetlands Ecol Manag 23:61–77. doi:10.1007/s11273-014-9358-2

Moradkhani H (2008) Hydrologic remote sensing and land surface data assimilation. Sensors (Basel, Switzerland) 8:2986–3004. doi:10.3390/s8052986

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE 50:885–900

Mueller N et al (2016) Water observations from space: mapping surface water from 25 years of Landsat imagery across Australia. Remote Sens Environ 174:341–352. doi:10.1016/j.rse.2015.11.003

Musa ZN, Popescu I, Mynett A (2015) A review of applications of satellite SAR, optical, altimetry and DEM data for surface water modelling, mapping and parameter estimation. Hydrol Earth Syst Sci 19:3755–3769. doi:10.5194/hess-19-3755-2015

Neal JC, Atkinson PM, Hutton CW (2007) Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. J Hydrol 336:401–415. doi:10.1016/j.jhydrol.2007.01.012

Neal J, Schumann G, Bates P, Buytaert W, Matgen P, Pappenberger F (2009) A data assimilation approach to discharge estimation from space. Hydrol Process 23:3641–3649. doi:10.1002/hyp.7518

Neal J, Schumann G, Bates P (2012) A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. Water Resour Res 48:W11506. doi:10.1029/2012WR012514

Nerger L, Gregg WW (2007) Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. J Mar Syst 68:237–254. doi:10.1016/j.jmarsys.2006.11.009

Oberstadler R, HÖNsch H, Huth D (1997) Assessment of the mapping capabilities of ERS-1 SAR data for flood mapping: a case study in Germany. Hydrol Process 11:1415–1425. doi:10.1002/(SICI)1099-1085(199708)11:10<1415:AID-HYP532>3.0.CO;2-2

O'Grady D, Leblanc M (2014) Radar mapping of broad-scale inundation: challenges and opportunities in Australia. Stoch Environ Res Risk Assess 28:29–38. doi:10.1007/s00477-013-0712-3

Oliver C, Quegan S (2004) Understanding Synthetic Aperture Radar Images. SciTech Publishing, London (UK), pp 479

O'Loughlin F, Trigg MA, Schumann GJP, Bates PD (2013) Hydraulic characterization of the middle reach of the Congo River. Water Resour Res 49:5059–5070. doi:10.1002/wrcr.20398

Papa F, Prigent C, Rossow WB, Legresy B, Remy F (2006) Inundated wetland dynamics over boreal regions from remote sensing: the use of Topex-Poseidon dual-frequency radar altimeter observations. Int J Remote Sens 27:4847–4866. doi:10.1080/01431160600675887

Pappenberger F, Beven K, Horritt M, Blazkova S (2005) Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. J Hydrol 302:46–69. doi:10.1016/j.jhydrol.2004.06.036

Pappenberger F, Frodsham K, Beven K, Romanowicz R, Matgen P (2007) Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. Hydrol Earth Syst Sci 11:739–752. doi:10.5194/hess-11-739-2007

Parker WV (2012) Discover the benefits of radar imaging. Earth Imaging J Remote Sens Satell Images Satell Imag. http://eijournal.com/2012/discover-the-benefits-of-radar-imaging

Pavelsky TM, Smith LC (2008) RivWidth: a software tool for the calculation of river widths from remotely sensed imagery. Geosci Remote Sens Lett IEEE 5:70–73. doi:10.1109/LGRS.2007.908305

Prestininzi P, Di Baldassarre G, Schumann G, Bates PD (2011) Selecting the appropriate hydraulic model structure using low-resolution satellite imagery. Adv Water Resour 34:38–46. doi:10.1016/j.advwatres.2010.09.016

Proud SR, Fensholt R, Rasmussen LV, Sandholt I (2011) Rapid response flood detection using the MSG geostationary satellite. Int J Appl Earth Obs Geoinf 13:536–544. doi:10.1016/j.jag.2011.02.002

Pulvirenti L, Chini M, Pierdicca N, Guerriero L, Ferrazzoli P (2011a) Flood monitoring using multi-temporal COSMO-SkyMed data: image segmentation and signature interpretation. Remote Sens Environ 115:990–1002. doi:10.1016/j.rse.2010.12.002

Pulvirenti L, Pierdicca N, Chini M, Guerriero L (2011b) An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic. Nat Hazards Earth Syst Sci 11:529

Raclot D (2006) Remote sensing of water levels on floodplains: a spatial approach guided by hydraulic functioning. Int J Remote Sens 27:2553–2574. doi:10.1080/01431160600554397

Raclot D, Puech C (2003) What does ai contribute to hydrology? aerial photos and flood levels. Appl Artif Intell 17:71–86. doi:10.1080/713827055

Rees WG (2012) Physical principles of remote sensing. Cambridge University Press, Cambridge

Revilla-Romero B, Hirpa F, Pozo J, Salamon P, Brakenridge R, Pappenberger F, De Groeve T (2015) On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions. Remote Sens 7:15702

Rodriguez E (2015) Surface water and ocean topography mission (SWOT), Science Requirements Document, JPL document D-61923. https://swot.jpl.nasa.gov/files/swot/SRD_021215.pdf

Rodriguez E, Moller D (2004) Measuring surface water from space. In: Paper presented at the AGU San Francisco (USA)

Romanowicz R, Beven K (1998) Dynamic real-time prediction of flood inundation probabilities. Hydrol Sci J 43:181–196. doi:10.1080/02626669809492117

Romanowicz R, Beven K (2003) Estimation of flood inundation probabilities as conditioned on event inundation maps. Water Resour Res 39:1073. doi:10.1029/2001WR001056

Romanowicz RJ, Beven KJ, Tawn J (1996) Bayesian calibration of flood inundation models. Floodplain processes. Wiley, Chichester, pp 333–360

Schlaffer S, Matgen P, Hollaus M, Wagner W (2015) Flood detection from multi-temporal SAR data using harmonic analysis and change detection. Int J Appl Earth Obs Geoinf 38:15–24. doi:10.1016/j.jag.2014.12.001

Schmugge T (1987) Remote sensing applications in hydrology. Rev Geophys 25:148–152. doi:10.1029/RG025i002p00148

Schumann GJP, Moller DK (2015) Microwave remote sensing of flood inundation. Phys Chem Earth A/B/C 83–84:84–95. doi:10.1016/j.pce.2015.05.002

Schumann G, Henry J-B, Hoffmann L, Pfister L, Pappenberger F, Matgen P (2005) Demonstrating the high potential of remote sensing in hydraulic modelling and flood risk management. In: Paper presented at the annual conference of the remote sensing and photogrammetry society with the NERC earth observation conference, Portsmouth, UK, 6–9 September 2005

Schumann G, Hostache R, Puech C, Hoffmann L, Matgen P, Pappenberger F, Pfister L (2007a) High-resolution 3-D flood information from radar imagery for flood hazard management. Geosci Remote Sens IEEE Trans 45:1715–1725. doi:10.1109/TGRS.2006.888103

Schumann G, Matgen P, Hoffmann L, Hostache R, Pappenberger F, Pfister L (2007b) Deriving distributed roughness values from satellite radar data for flood inundation modelling. J Hydrol 344:96–111. doi:10.1016/j.jhydrol.2007.06.024

Schumann G, Cutler M, Black A, Matgen P, Pfister L, Hoffmann L, Pappenberger F (2008a) Evaluating uncertain flood inundation predictions with uncertain remotely sensed water stages. Int J River Basin Manag 6:187–199. doi:10.1080/15715124.2008.9635347

Schumann G, Matgen P, Cutler MEJ, Black A, Hoffmann L, Pfister L (2008b) Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM. ISPRS J Photogramm Remote Sens 63:283–296. doi:10.1016/j.isprsjprs.2007.09.004

Schumann G, Matgen P, Pappenberger F (2008c) Conditioning water stages from satellite imagery on uncertain data points. Geosci Remote Sens Lett IEEE 5:810–813. doi:10.1109/LGRS.2008.2005646

Schumann G, Pappenberger F, Matgen P (2008d) Estimating uncertainty associated with water stages from a single SAR image. Adv Water Resour 31:1038–1047. doi:10.1016/j.advwatres.2008.04.008

Schumann G, Bates PD, Horritt MS, Matgen P, Pappenberger F (2009a) Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models. Rev Geophys. doi:10.1029/2008RG000274

Schumann G, Di Baldassarre G, Bates PD (2009b) The utility of spaceborne radar to render flood inundation maps based on multialgorithm ensembles. Geosci Remote Sens IEEE Trans 47:2801–2807. doi:10.1109/TGRS.2009.2017937

Schumann G, Di Baldassarre G, Alsdorf D, Bates PD (2010) Near real-time flood wave approximation on large rivers from space: Application to the River Po, Italy. Water Resour Res 46:W05601. doi:10.1029/2008WR007672

Schumann GJP, Neal JC, Mason DC, Bates PD (2011) The accuracy of sequential aerial photography and SAR data for observing urban flood dynamics, a case study of the UK summer 2007 floods. Remote Sens Environ 115:2536–2546. doi:10.1016/j.rse.2011.04.039

Schumann GJP, Bates PD, Di Baldassarre G, Mason DC (2012) The use of radar imagery in riverine flood inundation studies. In: Fluvial remote sensing for science and management. Wiley, pp 115–140. doi:10.1002/9781119940791.ch6

Schumann GJP et al (2013) A first large-scale flood inundation forecasting model. Water Resour Res 49:6248–6257. doi:10.1002/wrcr.20521

Schumann GJP, Vernieuwe H, De Baets B, Verhoest NEC (2014) ROC-based calibration of flood inundation models. Hydrol Process 28:5495–5502. doi:10.1002/hyp.10019

Schumann GJP, Bates PD, Neal JC, Andreadis KM (2015) Chapter 2—Measuring and mapping flood processes. In: Baldassarre JFSPD (ed) Hydro-meteorological hazards, risks and disasters. Elsevier, Boston, pp 35–64. doi:http://dx.doi.org/10.1016/B978-0-12-394846-5.00002-3

Shiiba M, Laurenson X, Tachikawa Y (2000) Real-time stage and discharge estimation by a stochastic-dynamic flood routing model. Hydrol Process 14:481–495. doi:10.1002/(SICI)1099-1085(20000228)14:3<481:AID-HYP950>3.0.CO;2-F

Siddique-E-Akbor AHM, Hossain F, Lee H, Shum CK (2011) Inter-comparison study of water level estimates derived from hydrodynamic–hydrologic model and satellite altimetry for a complex deltaic environment. Remote Sens Environ 115:1522–1531. doi:10.1016/j.rse.2011.02.011

Sippel SJ, Hamilton SK, Melack JM, Novo EMM (1998) Passive microwave observations of inundation area and the area/stage relation in the Amazon River floodplain. Int J Remote Sens 19:3055–3074

Sivapalan M et al (2003) IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. Hydrol Sci J 48:857–880. doi:10.1623/hysj.48.6.857.51421

Smith LC (1997) Satellite remote sensing of river inundation area, stage, and discharge: a review. Hydrol Process 11:1427–1439. doi:10.1002/(SICI)1099-1085(199708)11:10<1427:AID-HYP473>3.0.CO;2-S

Snyder C, Bengtsson T, Bickel P, Anderson J (2008) Obstacles to high-dimensional particle filtering. Mon Weather Rev 136:4629–4640. doi:10.1175/2008MWR2529.1

Speck R, Turchi P, Süß H (2007) An end-to-end simulator for high-resolution spaceborne SAR systems. In: Algorithms for synthetic aperture radar imagery XIV, Orlando, Florida, USA, 7 May. Proceedings of SPIE. doi:10.1117/12.717222

Stephens EM, Bates PD, Freer JE, Mason DC (2012) The impact of uncertainty in satellite data on the assessment of flood inundation models. J Hydrol 414–415:162–173. doi:10.1016/j.jhydrol.2011.10.040

Stephens E, Schumann G, Bates P (2014) Problems with binary pattern measures for flood model evaluation. Hydrol Process 28:4928–4937. doi:10.1002/hyp.9979

Stokstad E (1999) Scarcity of rain, stream gages threatens forecasts. Science 285:1199–1200. doi:10.1126/science.285.5431.1199

Tarpanelli A, Barbetta S, Brocca L, Moramarco T (2013a) River discharge estimation by using altimetry data and simplified flood routing modeling. Remote Sens 5:4145

Tarpanelli A, Brocca L, Melone F, Moramarco T (2013b) Hydraulic modelling calibration in small rivers by using coarse resolution synthetic aperture radar imagery. Hydrol Process 27:1321–1330. doi:10.1002/hyp.9550

Ticehurst C, Chen Y, Karim F, Dushmanta D (2013) Using MODIS for mapping flood events for use in hydrological and hydrodynamic models: experiences so far. In: Paper presented at the 20th international congress on modelling and simulation—MODSIM 2013, Adelaide (Australia)

Ticehurst C, Guerschman J, Chen Y (2014) The strengths and limitations in using the daily MODIS open water likelihood algorithm for identifying flood events. Remote Sens 6:11791

Ticehurst C, Dutta D, Karim F, Petheram C, Guerschman J (2015) Improving the accuracy of daily MODIS OWL flood inundation mapping using hydrodynamic modelling. Nat Hazards 78:803–820. doi:10.1007/s11069-015-1743-5

Tomkins KM (2014) Uncertainty in streamflow rating curves: methods, controls and consequences. Hydrol Process 28:464–481. doi:10.1002/hyp.9567

Townsend PA (2002) Estimating forest structure in wetlands using multitemporal SAR. Remote Sens Environ 79:288–304. doi:10.1016/S0034-4257(01)00280-2

Wang Y (2004) Using Landsat 7 TM data acquired days after a flood event to delineate the maximum flood extent on a coastal floodplain. Int J Remote Sens 25:959–974. doi:10.1080/0143116031000150022

Werner M, Blazkova S, Petr J (2005) Spatially distributed observations in constraining inundation modelling uncertainties. Hydrol Process 19:3081–3096. doi:10.1002/hyp.5833

Westerhoff RS, Kleuskens MPH, Winsemius HC, Huizinga HJ, Brakenridge GR, Bishop C (2013) Automated global water mapping based on wide-swath orbital synthetic-aperture radar. Hydrol Earth Syst Sci 17:651–663. doi:10.5194/hess-17-651-2013

Whitcomb J, Moghaddam M, McDonald K, Kellndorfer J, Podest E (2009) Mapping vegetated wetlands of Alaska using L-band radar satellite imagery. Can J Remote Sens 35:54–72. doi:10.5589/m08-080

Wilson M et al (2007) Modeling large-scale inundation of Amazonian seasonally flooded wetlands. Geophys Res Lett 34:n/a–n/a. doi:10.1029/2007GL030156

Woodhouse IH (2005) Introduction to microwave remote sensing. CRC Press, Florida, pp 400

Wright N, Villanueva I, Bates P, Mason DC, Wilson M, Pender G, Neelz S (2008) Case study of the use of remotely sensed data for modeling flood inundation on the river Severn, UK. J Hydraul Eng 134:533–540. doi:10.1061/(ASCE)0733-9429(2008)134:5(533)

Yan K, Di Baldassarre G, Solomatine DP, Schumann GJP (2015) A review of low-cost space-borne data for flood modelling: topography, flood extent and water level. Hydrol Process. doi:10.1002/hyp.10449

Yu D, Lane SN (2006) Urban fluvial flood modelling using a two-dimensional diffusion-wave treatment, part 1: mesh resolution effects. Hydrol Process 20:1541–1565. doi:10.1002/hyp.5935

Zwally HJ et al (2002) ICESat's laser measurements of polar ice, atmosphere, ocean, and land. J Geodyn 34:405–445. doi:10.1016/S0264-3707(02)00042-X

Zwenzner H, Voigt S (2009) Improved estimation of flood parameters by combining space based SAR data with very high resolution digital elevation data. Hydrol Earth Syst Sci 13:567–576. doi:10.5194/hess-13-567-2009