A Bayesian Network for Flood Detection Combining SAR Imagery and Ancillary Data

Annarita D'Addabbo, Alberto Refice, Guido Pasquariello, Francesco P. Lovergine, Domenico Capolongo, and Salvatore Manfreda

***Abstract*—Accurate flood mapping is important for both planning activities during emergencies and as a support for the successive assessment of damaged areas. A valuable information source for such a procedure can be remote sensing synthetic aperture radar (SAR) imagery. However, flood scenarios are typical examples of complex situations in which different factors have to be considered to provide accurate and robust interpretation of the situation on the ground. For this reason, a data fusion approach of remote sensing data with ancillary information can be particularly useful. In this paper, a Bayesian network is proposed to integrate remotely sensed data, such as multitemporal SAR intensity images and interferometric-SAR coherence data, with geomorphic and other ground information. The methodology is tested on a case study regarding a flood that occurred in the Basilicata region (Italy) on December 2013, monitored using a time series of COSMO-SkyMed data. It is shown that the synergetic use of different information layers can help to detect more precisely the areas affected by the flood, reducing false alarms and missed identifications which may affect algorithms based on data from a single source. The produced flood maps are compared to data obtained independently from the analysis of optical images; the comparison indicates that the proposed methodology is able to reliably follow the temporal evolution of the phenomenon, assigning high probability to areas most likely to be flooded, in spite of their heterogeneous temporal SAR/InSAR signatures, reaching accuracies of up to 89%.**

***Index Terms*—Bayesian networks (BNs), data fusion, flood mapping, synthetic aperture radar (SAR) change detection, synthetic aperture radar (SAR)/interferometric SAR (InSAR) time series analysis.**

Manuscript received May 6, 2015; revised November 25, 2015; accepted December 21, 2015. Date of publication March 9, 2016; date of current version April 27, 2016. This work was supported in part by the Italian Ministry of Research (MIUR) within the "Apulia Space" PON project PON03PE_00067_6. The work of D. Capolongo was supported by the Research Project PRIN 2011 "Response of morphoclimatic system dynamics to global changes and related geomorphological hazard."

A. D'Addabbo, A. Refice, G. Pasquariello, and F. P. Lovergine are with Istituto di Studi sui Sistemi Intelligenti per l'Automazione-Consiglio Nazionale delle Ricerche (ISSIA-CNR), 70125 Bari, Italy (e-mail: daddabbo@ba.issia.cnr.it; refice@ba.issia.cnr.it; pasquariello@ba.issia.cnr.it; f.lovergine@ba.issia.cnr.it).

D. Capolongo is with the Department of Earth and Environmental Science, University of Bari, 70125 Bari, Italy (e-mail: domenico.capolongo@uniba.it).

S. Manfreda is with Dipartimento delle Culture Europee e del Mediterraneo (DICEM), University of Basilicata, 75100 Matera, Italy (e-mail: salvatore.manfreda@unibas.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2016.2520487

I. Introduction

FLOODING is a recurrent and dramatic natural disaster that affects several areas in the world, both in tropical and temperate regions, often causing great damages to lives and property, industrial settlements and infrastructures, and artistic and historical sites, aside from dramatically affecting local ecosystems. It is expected that the frequency of these phenomena will increase in the next future, due to climate change and the occurrence of several extreme weather events. Accurate knowledge of flood extents is crucial to improve disaster management and to mitigate the impact of flood episodes, during both the emergency phase, when an overall view is needed in order to plan relief efforts, and the aftermath, for the assessment of damaged areas.

Remote sensing data of affected areas, together with effective methods to exploit their information, can be particularly useful in the damage estimation phase: they offer a synoptic overview of wide areas, giving a snapshot of the extents of the flooding event [1]–[3]. Both optical and radar data have been used to infer information about the flooded areas [4]. Optical data have been used, particularly in urban areas [5], [6], where the effects of shadowing, layover, and foreshortening limit the results obtained by synthetic aperture radar (SAR) images, although new insight on these topics has been recently provided, e.g., in [3] and [7]. SAR sensors are particularly useful to produce flood maps, owing to their all-weather and day–night operation capability [8]–[10]. In particular, the latest generation of SAR sensors, such as TerraSAR-X and the COSMO-SkyMed constellation, supplies a large amount of daily acquired radar images, characterized by high spatial resolution, that can be particularly useful in flood detection problems [3], [11], [12].

Nevertheless, flood scenarios are typical examples of complex situations in which different factors have to be considered to provide accurate and robust interpretation of the situation on the ground. The presence of many land cover types, each one with a particular signature in the presence of flood, requires modeling the behavior of different "objects" in the scene (not necessarily associated to land cover categories) in order to associate them to flood or no-flood conditions. In principle, a single SAR image acquired during an event could provide a flood map: its capability to detect inundations is based on the low backscatter response from smooth water covering the terrain, which allows to efficiently separate flooded areas from rougher nonflooded terrain that, instead, scatters the impinging signal in many different directions and produces a higher backscattering coefficient (σ^0). However, many factors potentially complicate this simple situation. For instance, the contrast

between dry soil and flooded one can be significantly reduced in the presence of wind roughening the water surface. Also, vertical structures on the ground, such as trees or man-made objects, could determine a double bounce mechanism of the radar signal, which is enhanced when the ground surface is flooded with respect to dry conditions: this can make the radar return from vegetated or urban flooded areas higher than that under nonflooded conditions [12].

In general, the exploitation of a set of images (as a multitemporal stack) seems to provide more satisfactory results in the presence of these critical situations [8], [13]. Moreover, the interferometric SAR (InSAR) coherence information can be a viable means for detecting flooded areas as an extension of conventional change detection methods [14], [15]. However, the relative difficulty in planning pairs of SAR acquisitions suitable for InSAR use in coincidence with flood events, for legacy sensors, posed some problems in devising operational scenarios for the use of InSAR in flood monitoring. A renewed attention to this issue came with the availability of the aforementioned high-resolution sensors.

Recently, the combined analysis of multi-temporal SAR intensity and coherence trends has been proposed to detect flooded areas, reducing false alarms and missed identifications which affect algorithms based on SAR intensity alone [16], [17]. More generally, the fusion of multitemporal, multisensor, multiresolution, and/or multiplatform Earth observation image data, together with other ancillary information, seems to have a key role in the pursuit of a consistent interpretation of complex scenes: the decisions or actions taken considering data synergy exploitation are better (both qualitatively and quantitatively, in terms of accuracy, robustness, etc.) than would be possible if the sources were used individually [18]. In the case of flooding, distance from the river, terrain elevation, hydrologic information, or some combination thereof can add useful information that leads to a better performance in flood detection. In fact, the integration of remote sensing-derived flood extent maps and hydraulic models is now a promising research field (see, e.g., [19] and the references therein). In [20], a procedure was developed in a geographic information system (GIS) environment in order to estimate the flooded area combining SAR imagery with ancillary information extracted from digital elevation topography. In [21], two different statistical methods have been used to integrate SAR data with high-precision topography and a river flow model. More recently, flood maps have been produced combining multitemporal SAR data and a hydrodynamic model by using a classification algorithm based on fuzzy logic [22].

In this paper, we propose a statistical model able to perform flood detection by using information and data fusion. It is based on Bayesian networks (BNs), one of the most common types of probabilistic graphical models combining probability theory and graph theory, which introduce graph structures into a probabilistic model to represent dependence assumptions among the involved variables [23], [24]. BNs are a statistically well-founded method to combine imagery information with ancillary data, such as distance from the river, digital elevation models (DEMs), hydraulic models, etc. This combination is not restricted to a specific sensor, and it can exploit the information from time series of images. In our case, it is also capable of describing the dynamics of the flooding event.

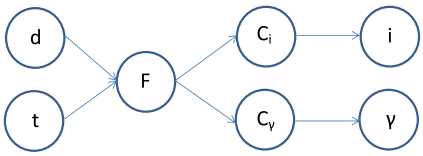

Fig. 1. BN scheme.

In recent years, BNs and, more generally, probabilistic graphical models have been used as a data mining tool in many research fields, but rarely to process remote sensing data. In [25], a BN has been applied to two Landsat TM scenes acquired at different times on the same area, obtaining an overall change detection classification accuracy that overcomes the one obtained by other methods. However, to our knowledge, the capability of BNs to join different kinds of data has been scarcely exploited. In [5], a simple probabilistic graphical model has been applied to estimate the state of roads during flooding, combining a DEM with satellite optical images: the obtained results outperform other methods based only on the exploitation of imagery data or only on DEM information. On the other hand, BNs have been successfully used to construct risk assessment systems for natural disasters, such as rock falls [26], avalanches [27], or wildfires [28]: various kinds of data, such as land cover types extracted from the Corine Land Cover database, road and building densities, have been fused by means of BNs to obtain reliable risk maps, although none included remotely sensed imagery. In this paper, a BN is used in a flood detection problem to integrate multitemporal SAR intensity and InSAR coherence data with geomorphic and other ground information.

This paper is organized as follows. In Section II, the proposed method is discussed in detail. In Section III, some information on the study area and the considered flood event are reported, and the available data, as well as the reference data used for the evaluation, is described. The experimental results are presented in Section IV. Finally, in Section V, some conclusions are drawn.

II. Methods: BNs for Flood Map Production

BNs combine probability theory and graph theory, introducing graph structures into a probabilistic model to represent dependence assumptions among the involved variables. The interactions of the variables of a "system" are encoded through a directed acyclic graph (DAG) (see Fig. 1): the nodes of the DAG represent the random variables, and the edges among them correspond to the direct influence of one node on another, modeled as conditional probabilities [23], [24]. The BN structure thus provides information about the underlying processes and the way that various variables communicate and share "information" as it spreads around the network. Formally speaking, the DAG encodes independence assumptions between variables.

A BN is characterized by many attractive properties that can be particularly useful in modeling many real-world systems.

$$
\begin{aligned}
&p(F = \text{flood} | i = i^*, \gamma = \gamma^*, t = t^*, d = d^*) \\
&\quad = \frac{p(F = \text{flood}, i = i^*, \gamma = \gamma^*, t = t^*, d = d^*)}{p(i = i^*, \gamma = \gamma^*, t = t^*, d = d^*)} \\
&\quad = \frac{\left[\sum_{C_i} p(i = i^* | C_i) p(C_i | F = \text{flood})\right] \left[\sum_{C_\gamma} p(\gamma = \gamma^* | C_\gamma) p(C_\gamma | F = \text{flood})\right] p(F = \text{flood} | t = t^*, d = d^*)}{\sum_F \left[\sum_{C_i} p(i = i^* | C_i) p(C_i | F)\right] \left[\sum_{C_\gamma} p(\gamma = \gamma^* | C_\gamma) p(C_\gamma | F)\right] p(F | t = t^*, d = d^*)}
\end{aligned} \tag{1}
$$

First of all, it naturally provides a general methodology for handling data fusion, allowing the joint distribution of all variables to be written down tractably, even in cases where its explicit representation is particularly large. Moreover, a BN graphical structure can be based on prior knowledge, allowing to construct models that provide an accurate reflection of the understanding of a domain. Alternatively, the structure may be learned from data, or even a combination of the two approaches can be used: pose a prior BN, and update it based on observations. Finally, it enables inference: as previously mentioned, the BN captures the joint distribution of all variables, and, via inference, any/all conditional distribution(s) of interest, in any direction, can be expressed. These features appear particularly suited for the application domain of natural disaster monitoring and assessment [29].

We consider here a BN designed as represented in Fig. 1: it combines information extracted by a set of images, and some ancillary information about the *a priori* spatial distribution of flood in the analyzed area. In detail, the random variable F is discrete and consists of only two states, flood and no flood, so that $P(F = \text{flood}) = 1 - P(F = \text{no flood})$. This is the variable that we want to estimate by statistical inference. The variable i corresponds to the n-dimensional vector obtained from the SAR intensity imagery. Since, often, the flood state does not exhibit a simple causal relation with the imagery, due to the presence of particular types of land cover, particular backscattering mechanisms, etc., the intermediate variable class C_i is introduced [5]. It consists of K_i possible states that can be determined in two different ways: either by using some *a priori* knowledge on the scene or by extracting them automatically from the imagery. Anyway, the number K_i of possible states and the usefulness of the C_i variable are strongly dependent on the complexity of the analyzed scene. For example, if the area under observation is characterized only by unvegetated/barren areas and a single SAR intensity image is available, the variable C_i is unnecessary, and the i random variable can be directly connected to F: the backscatter values of flooded areas, in this case, will be very low and well separated from nonflooded ones. Otherwise, for more complex scenarios, such as in the presence of vegetated, forested, or urban areas, or if we want to follow the flood dynamics owing to the availability of multitemporal data, the introduction of the intermediate variable C_i allows to classify the various scattering behaviors into a finite number of types and then to focus the attention on the most interesting ones [5], [8]. The variable γ corresponds to the m-dimensional vector obtained from the multitemporal InSAR coherence imagery. Similarly to the previous case, this variable is not directly connected with the F one, and a hidden C_γ variable is introduced. It consists of K_γ possible states, for which the previous considerations hold. Finally, the variables t and d represent some ancillary information, i.e., a geomorphic flooding index (GFI) derived from a light-detection-and-ranging-derived DEM at 3-m posting, and the geometric distance from the river, properly modulated by the intensity of the considered event (the maximum expected distance reached by the river flood is estimated by considering the effective water level measurements along the river, available as readings from a gauge station), respectively. The t variable is computed from the definition in [30] and [31] as follows:

$$t = \ln\left(\frac{h_r}{H}\right)$$

where h_r is a function of the contributing area A_r in the nearest section of the drainage network hydrologically connected to the point under exam and H is the elevation difference between the cell under exam and the final point of the previously identified path. In this index, the parameter h_r is representative of the probable water level in a cross section of the river hydraulically connected to the point, and it is computed as a power law of A_r with the exponent set equal to 0.3. The variable t is introduced to assign correct *a priori* flood probability to zones, even relatively far from the river course, that may be flooded due to the presence of dense channel networks, which become a preferential vector for the flow during a flood.

The second component of the BN is a set of local probability models. As previously mentioned, the relations between variables are modeled as conditional probabilities, and the conditional independence assumptions hold, i.e., if two variables are not directly connected in the graph, they are independent. We want to infer the value of the conditional probability of the variable F, which derives from the BN structure in Fig. 1 and is given in (1) shown at the top of the page, where i^*, γ^*, t^*, and d^* are observed values for the random variables i, γ, t, and d, respectively.

In particular, the conditional probability $p(i|C_i)$ is given as an n-dimensional probability function generated by a mixture of K_i Gaussian distributions, $\mathcal{N}(\mu, \Sigma)$. The parameters $\{\mu_{C_i}, \Sigma_{C_i}\}$ of each multitemporal intensity function have been automatically computed by applying a K-means algorithm [32]. The number K_i of clusters C_i has been determined by a trial-and-error procedure, in order to find the value that provides a good representation of different classes actually present in the analyzed scene. Similarly, the conditional probability $p(\gamma|C_\gamma)$ is given as an m-dimensional probability function generated by a mixture of K_γ Gaussian distributions. Again, the parameters

$\{\mu_{C_\gamma}, \Sigma_{C_\gamma}\}$ of each coherence function have been automatically computed by applying a K-means algorithm, and the number K_γ of clusters C_γ has been determined by a trial-and-error procedure. It is worth noting that SAR and InSAR data distributions are usually modeled through Gamma or Wishart functions [33]–[35]; however, if the number of looks of the filtered SAR images is sufficiently high, both the SAR intensity and InSAR coherence pdfs can be well approximated with Gaussian functions. Thus, the aforementioned use of Gaussian multivariate pdfs is justified in view of the filtering applied to the data in the preprocessing steps (see Section III-B). Some considerations on the clustering step and its influence on the applicability and portability of the whole proposed method are in order. First of all, our structure implies that the InSAR coherence is considered as a channel independent from the SAR intensity. InSAR coherence carries information about the temporal stability of the microscopic arrangement of the individual scatterers on the terrain, so it measures the degree of similarity of two SAR images, at the level of the speckle pattern of single pixels. Therefore, strictly, the coherence computed between two SAR acquisitions is not independent from them. However, our probability distribution in this case satisfies a conditional independence property, i.e., if we know that a given pixel corresponds to a flooded area, the knowledge of the coherence value no longer gives us information about the intensity. More formally, $P(i|F = \text{flood}, \gamma = \gamma^*) = P(i|F = \text{flood})$. Similarly, $P(\gamma|F = \text{flood}, i = i^*) = P(\gamma|F = \text{flood})$. Therefore, the conditional independence holds.

Moreover, the best number of clusters is not generally known in advance. As previously stated, it depends on the analyzed scene, on the number of available images, and on the extension of the flood with respect to the whole scene. This information allows to narrow our search, but in general, a trial-and-error procedure is needed. It can be time consuming if several tests are performed. However, even in this case, the BN framework helps to reduce the computational time and cost because it guarantees good performances even if the search is not refined: it is mandatory to avoid underclustering, but overclustering is not a problem. In fact, if the number of clusters is too low, there could be mixed clusters, composed by objects with different spectral signatures. In this case, the final result could be severely compromised. On the contrary, if the number of clusters is higher than the best one, there could be several clusters associated to the same spectral signature, thus having equal *a priori* probability values $P(F|C_k)$, with $k = i, \gamma$, but this does not compromise the final result.

The probabilities $p(C_i|F)$ and $p(C_\gamma|F)$ have been computed by applying the Bayes rule

$$p(C_k|F) = \frac{p(F|C_k)P(C_k)}{\sum_{C_k} p(F|C_k)P(C_k)} \tag{2}$$

where the $P(C_k)$ terms, with $k \in \{i, \gamma\}$, are computed from the results of the K-means algorithm and the $P(F|C_k)$ are assigned by the user. The conditional probability $p(F|t, d)$ has been modeled as a sigmoid function

$$p(F|t, d) = \frac{1}{1 + \exp\left[-(At - Bd + C)\right]} \tag{3}$$

where the A, B, and C parameters are set by some *a priori* knowledge about the scene and the considered event. In particular, the A and B parameters represent the relative weight given to the two variables t and d, respectively, and regulate the span of the sigmoidal function in (3). The C parameter controls the set of points where $p(F|t, d) = 0.5$ (see the Appendix for further considerations).

Summarizing, the proposed BN structure encodes our intuition about "the way the world works" in a flood detection problem, when the aforementioned variables are considered. However, the proposed method is quite general and can be considered as a particular case of a general "class" of BNs, in which M conditional independent imagery sources and N independent ancillary data can be considered. In particular, it holds also if one source of information is lacking, without having to modify our algorithm structure at all. For example, if one of the two ancillary information sources is not available, it is sufficient to put the corresponding coefficient (A or B) equal to 0 in (3), and the model still holds. If no ancillary information is available, the following condition is verified: $p(F = \text{flood}|t, d) = p(F = \text{no flood}|t, d) = 0.5$, and (1) still holds. Similarly, if one of the two imagery sources x is not available, the following condition is verified: $\sum_{C_x} p(x|C_x)p(C_x|F = \text{flood}) = \sum_{C_x} p(x|C_x)p(C_x|F = \text{no flood}) = 0.5$ and the model is again usable.

III. Study Area and Data Set

A. Bradano River Flood Event in December 2013

Several flood events occurred in the last years in the Basilicata region (Southern Italy) and, in particular, in its coastal area on the Jonian Sea, due to its intrinsic lithological and geomorphological properties [36]. In this area, several rivers are flowing (see Fig. 2). The surrounding land cover is mainly agricultural, characterized by cereal and vegetable crops or fruit shrubs, while an extended area near the river mouths and along the coast is occupied by pine forest stands [37]. The frequency of extreme events has increased, in the last five years [38], causing several flooding episodes, even two–three in the same year, with great damage for agriculture and industries. This process is enhanced by the agricultural land use of the alluvial plains, which has reduced the extent of the natural river beds, increasing flood risk. We focus on one of such events involving the Bradano and Basento river basins. It occurred on November 30–December 7, 2013, and was due to rainfall starting from November 30 and continuing until December 5. In particular, we consider an area around the Bradano river, for which the peak flow (measured at the gauging station "SS106" close to the outlet; see Fig. 3) was recorded in the evening of December 1, with the discharge reaching about 800 m^3/s, and caused an inundation that propagated in the surrounding areas, producing floods during the following four days. Another less relevant peak of about 300 m^3/s was registered on December 2. Thereafter, the rainfall event was exhausted, and the dams released some of the flood volume stored. The considered event was triggered by a relevant amount of rainfall that reached the value of about 130 mm in one day, while the total amount of

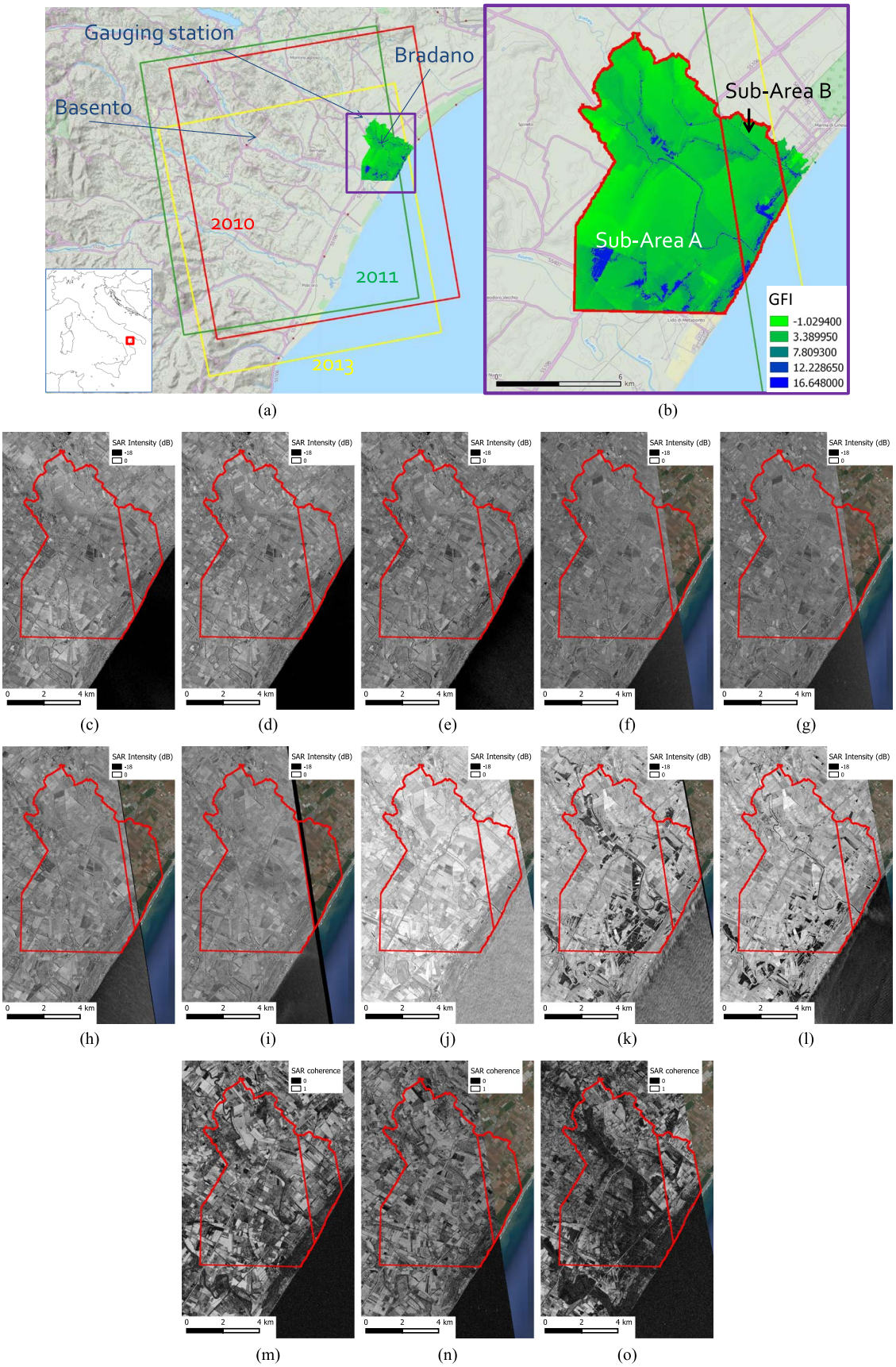

Fig. 2. (a) Location map of the data set. The frame outlines the masters of the three interferometric combinations for 2010, 2011, and 2013, respectively, and the Basento and Bradano rivers, as well as the location of the gauging station whose data are reported in Fig. 3. (b) Zoom map showing the GFI data set and highlighting the two subareas (A and B), delimited by the different data availability (green and yellow lines). (c)–(l) SAR intensity maps ordered in time as in Table I. (m)–(o) SAR coherence maps ordered in time as in Table I.

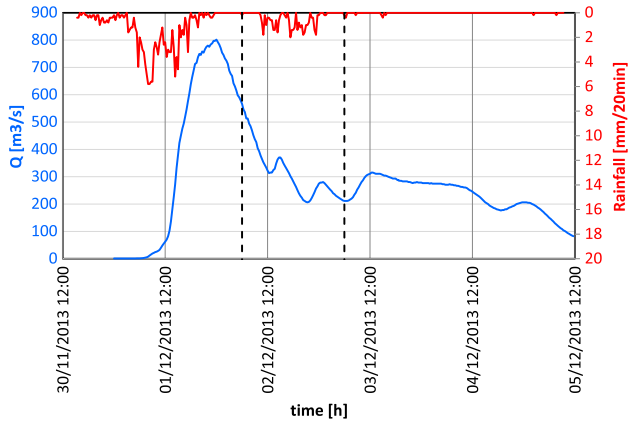

Fig. 3. Temporal evolution of the water discharge (blue curve, left y-axis scale) and rainfall (red curve, right y-axis scale) recorded at the SS106 gauging station on the Bradano river. The vertical dashed lines indicate the time of the two SAR acquisitions during the flood event. Data courtesy of Basilicata Civil Protection.

rainfall reached the value of about 170 mm. In particular, as officially reported in the data of Basilicata Civil Protection, on December 2, the Bradano river provoked inundations that affected large areas concentrated along the river banks. In the following days, the phenomenon was in a decreasing phase, and the flood effects were propagating toward the coast, interesting smaller areas progressively farther away from the river. Nevertheless, this flood should be considered an ordinary event that unfortunately produced significant inundation for a levee failure. Part of the considered test area lies between the Bradano and Basento rivers and could thus be potentially interested by flood due to the cumulative action of both rivers.

B. SAR/InSAR Data Set

A series of three COSMO-SkyMed stripmap SAR images, acquired close to the considered event, was available over the analyzed area, with a ground pixel size of approximately 3×3 m^2, acquired in the same geometry, polarization (VV), and incidence angle (38°), so that InSAR processing could be performed. The acquisition dates are November 16 and December 2 and 3, 2013. As can be noted, the two SAR acquisitions of December 2 and 3 offer a useful observation data set to follow the temporal evolution of the flood wave phenomenon, which was in a crucial phase on December 2 and in a decreasing one on the following day. The InSAR coherence computed between these two images was also considered.

Other seven COSMO-SkyMed stripmap SAR images were considered over the same area, acquired in the absence of flood events on October 2, 3, and 10, 2010, January 17, 18, and 25, 2011, and February 18, 2011, respectively. The InSAR coherence between October 2 and 3, 2010, and January 17 and 18, 2011, was considered as well. These additional SAR data act here as a reference data set that allows to capture multitemporal backscattering and coherence trends, aiding in the detection of flood. The image frames acquired on 2011 do not cover the whole study area; consequently, this has been divided into two subareas (as depicted in the zoom map on the right of Fig. 2), indicated as subarea A and subarea B, respectively: in the first one, the whole set of ten SAR intensity and three InSAR coherence images has been considered, and in the second one, only six intensity and two coherence images have been used.

TABLE I
DATES OF THE AVAILABLE SAR IMAGES AND PERPENDICULAR BASELINES FOR THE USED INSAR PAIRS

Date	$B_\perp$ (m)
02/10/2010	→ 106.9
03/10/2010	
10/10/2010	
17/01/2011	→ 6.2
18/01/2011	
25/01/2011	
18/02/2011	
16/11/2013	
02/12/2013	→ 326.2
03/12/2013	

As can be seen from Table I, the InSAR data set has relative spatial baselines well within the critical value (which, for COSMO-SkyMed data, is of a few kilometers), which ensures that the geometrical decorrelation is negligible [39]. Preprocessing included standard SAR intensity calibration, InSAR coregistration, and coherence and phase estimation. Interferometric processing was thus performed through the DORIS open source software [40]. Images were precisely coregistered using height information from SRTM data [41]. On the complex interferograms, a filtering step was performed using a nonlocal approach [42], which computes averages of homogeneous targets over large windows, with the number of looks that adaptively varies between 1 and the maximum size of the search window used, which, in our case, was set to 25 × 25 pixels. Generally, a prefiltering step is considered necessary when using SAR images in mapping applications, to reduce the spatial heterogeneity of scattering intensity induced by the coherent nature of the sensor. In our case, the adopted nonlocal filtering paradigm increases the effective number of looks of spatially homogeneous image areas while preserving high resolution, which is a peculiar characteristic of new-generation SAR sensors such as COSMO-SkyMed. Since the analyzed area is mainly covered by agricultural fields, having a number of looks that is considerably high, the Gaussian assumption in the BN is justified, as shown in the intensity and coherence histograms reported in Fig. 4. Final filtered intensity and coherence products were geocoded to a common reference frame, thus obtaining regular-grid UTM maps with a posting of 3 m in both dimensions.

C. Reference Data

Two different reference maps, independently obtained and partially overlapped with the analyzed area, are available. The first one, named GT1, has been obtained by the Copernicus Emergency Mapping Service [43]. In this map, delivered at a ground resolution of 5 m, a unique thematic layer, assessing the delineation of the flooded areas, has been derived by

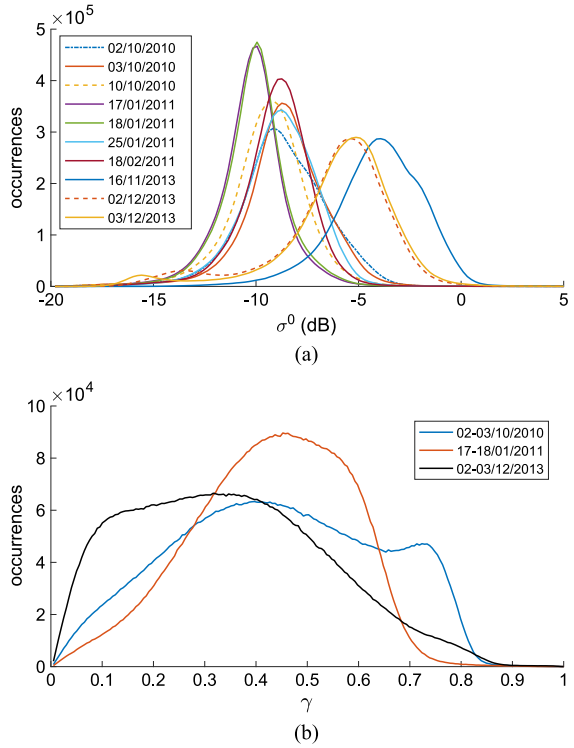

Fig. 4. Histograms of (a) SAR intensity and (b) InSAR coherence.

visual interpretation from a GSD Spot 6 image, acquired on December 4, 2013. The second one, indicated in the following as GT2, has been extracted from a Pleiades-1B multispectral image, at a ground resolution of 2 m, acquired on December 5, 2013. In this case, different informative layers have been extracted by visual inspection: the flooded areas on December 5 and the ones that have been probably inundated before the image acquisition, characterized by the presence of mud or signs of water stagnation. To each area, a score has been assigned, ranging from 1 to 3, corresponding to the confidence level (assigned by the expert analysis) that it has been actually interested by inundation. A brief description of each score level is reported hereafter:

1) areas flooded at the image acquisition time;
2) areas not flooded at the acquisition time, but in which there are evident signs of water presence in the days before and its stagnation;
3) areas not flooded at the acquisition time, but in which there are only slight signs of previous water presence and its stagnation.

It is worth noting that the last two levels in GT2 are particularly interesting in our case because they allow us to deduce information temporally coincident with our SAR/InSAR data set, composed by images acquired on December 2 and 3, 2013. For the purpose of this application, we ignore their distinction and lump them together. For these reasons, a new reference map has been created by suitably merging information extracted by the two aforementioned maps. It covers the area shared by the two primary maps and contains the following information layers:

1) areas flooded before December 4;
2) areas flooded on December 4 and which could be flooded also in earlier days;
3) areas flooded on both December 4 and 5 and which could be flooded also in earlier days;
4) areas flooded from December 5 onward.

TABLE II
REFERENCE DATA DECISION TABLE

	Reference (GT) data layers			
Original ref. data layers	1	2	3	4
Flooded on Spot 6 (Dec. 4) GT1	N	Y	Y	N
Not flooded on Spot 6 (Dec. 4) GT1	Y	N	N	Y
Flooded on Pleiades (Dec. 5) Level 1 GT2	N	N	Y	Y
Not flooded on Pleiades (Dec. 5) Level 2 GT2	Y	Y	N	N
Not flooded on Pleiades (Dec. 5) Level 3 GT2	Y	Y	N	N

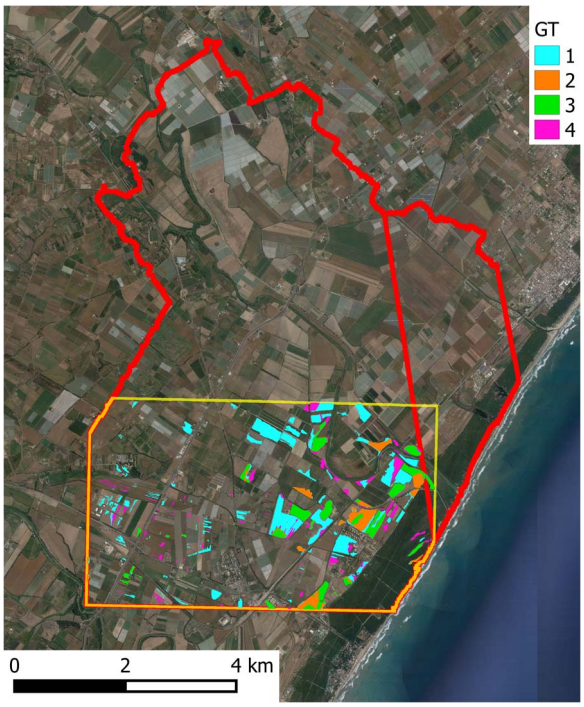

Fig. 5. Reference map GT. The extension of the considered GT area is delimited by the yellow line. The red lines show the two subareas as in Fig. 2. See text for label explanation.

These layers have been obtained by the ones of GT1 and GT2 as summarized in Table II. The final GT map, resampled at the resolution of 3×3 m^2, is reported in Fig. 5. It is worth noting that, in the reference data, some areas are labeled as flooded from December 5 onward (class 4): these seem to be not in agreement with the temporal evolution of the event, which was in a waning phase on this day, as reported in Section III-A. Such apparent inconsistencies may be due to the presence of approximations in the photointerpretation process: first of all, there is a field, very close to the river bank, that has been wrongly labeled as "not flooded" in the GT1 map (see [43]). Moreover, there are many small fields in GT2 that are not visible in GT1 due to their different resolution.

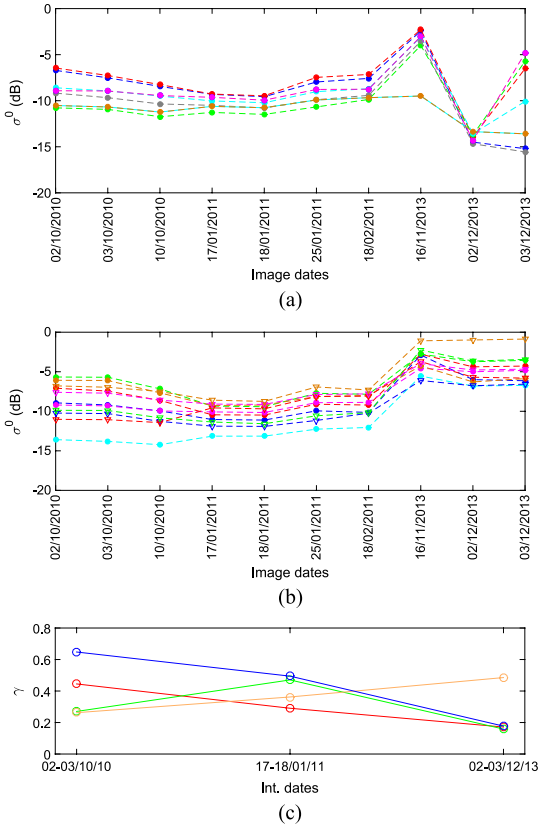

Fig. 6. Centroid values of (a and b) SAR intensity, and (c) InSAR coherence, computed by the K-means algorithm, as a function of time for the most significant clusters.

IV. Experimental Results

A. Multitemporal Intensity Imagery Segmentation and Probability Assignment

First of all, the conditional probabilities $p(i|C_i)$ have been computed as 10-D (or 6-D) probability functions generated by a mixture of Gaussian distributions, as explained in Section II. The parameters $\{\mu_{C_i}, \Sigma_{C_i}\}$ of each function have been automatically computed by the K-means algorithm. We have performed various tests finding that $K_i = 32$ is a number of clusters sufficient to identify univocally all the different types of multitemporal scattering classes present in the data. Of these, a subset represents objects of interest for the specific flood monitoring experiment. In Fig. 6, the temporal trends of the mean value of σ^0 in some clusters are reported. In particular, mean values that exhibit significant variations in correspondence with the flood event are plotted. In Fig. 6(a), seven different trends are shown. Of these, four are characterized by very small σ^0 values on December 2, while before and after this date, the radar return is higher: regions showing this multitemporal signature were considered as flooded only on December 2. Instead, the other three trends exhibit low σ^0 values both on December 2 and 3: it can be deduced that such signatures represent areas flooded in both days. It is interesting to note that there is no cluster showing low values of backscattering signal only on December 3: this occurrence fully agrees with the information concerning the rainfall phenomenon, which was almost finished at this date. Also, most of the σ^0 trends plotted in Fig. 6 show an increase of several decibels (dB) on November 16, 2013. This can be ascribed to the heavy rain which was falling on the area at the time of this acquisition, recorded by meteo stations and reported by several news websites (e.g., [44]), which is likely to have increased the average backscatter level of the fields in the area due to increased soil moisture. As a confirmation of this interpretation, a whitish blob, most likely the signature of heavy precipitation, not uncommon on X-band SAR images (see, e.g., [45]), is also visible on a region (not shown) of the same image acquired on November 16, some tens of kilometers SW of the test site. In the other dates, the backscatter signals are, instead, quite steady because no other changes are present in these clusters. In fact, the whole data set is composed by images acquired in winter and mainly covering agricultural areas, not subject to agricultural practices in this period of the year, and exhibiting no changes due to the phenological cycle of their cultures.

Fig. 6(b) plots 11 different trends of σ^0 value, all characterized by high backscattering signal values on December 2 and 3, similar to those on November 16, and 4–5 dB higher than the preceding days: this could be explained by considering the double bounce effect [12] or could be ascribed to increased soil moisture. The absolute σ^0 levels on the flood dates are scattered on a relatively broad interval of more than 5 dB. This may reflect the complexity of the phenomena and, in case of double bounce, the variations in radar response due to plant height, water depth, and their relative values. Moreover, it is worth noting that these clusters contain almost 43% of the total analyzed pixels and can be found scattered across the whole scene. We did not have *a priori* detailed information concerning land use. For example, we know that, in the area, some agricultural fields covered by vineyards are present, which could be affected, for their structure, by the double bounce effect, but we did not know their location and extension at the event time. The same situation occurs for other land use classes or for the case of increased soil moisture. Fortunately, in such cases, the BN helps us to manage our uncertainty and lack of information, which is expressed in the $P(F|C_i)$ terms assigned. In particular, a value $P(F|C_i) = 0.5$ is given to controversial signatures [e.g., the intensity signatures in Fig. 6(b)]. Conversely, high values of flood probability, $P(F|C_i) = 0.9$, are assigned to clusters with a more straightforward interpretation as flooded fields, such as the ones reported in Fig. 6(a). Finally, very low *a priori* probabilities, $P(F|C_i) = 0.1$, are given to the other clusters showing negligible intensity variations on the flood dates with respect to the preceding days. Subsequently, the probabilities $P(C_i|F)$ are computed by applying the Bayes rule (2), where the $P(C_i)$ terms are assumed to be proportional to the relative populations of the clusters coming from the K-means algorithm.

B. Multitemporal Coherence Imagery Segmentation and Probability Assignment

Independently from the intensity imagery segmentation, the conditional probabilities $p(\gamma|C_\gamma)$ were computed as 3-D (or 2-D) probability functions generated by a mixture of Gaussian distributions, as explained in Section II. On most

types of natural terrain, coherence is expected to be low on areas which have been inundated for some time during the interval between the two InSAR acquisitions, but this signature is also characteristic of other land cover classes such as permanent water and forests, also present in the analyzed scene. The availability of multiple InSAR coherence images is particularly useful in the interpretation of flooded and nonflooded areas because they allow to separate areas exhibiting permanently low coherence values (permanent water or forests) and areas subjected to a decrease of coherence value during the inundation period (flooded areas). On the other hand, it should be noted that the coherence value on flooded areas is low even if they have been covered by water at any time between the first and the second sensor pass, but they are not covered at the exact time of any of the two acquisitions. This aspect could be a drawback when following the evolution of the flood dynamics, but it can be overcome when the coherence information is suitably merged with other kinds of data. Also, in this case, the parameters $\{\mu_{C_\gamma}, \Sigma_{C_\gamma}\}$ of the multivariate Gaussian approximating each representative data type have been automatically assigned by the K-means algorithm. Various tests led to an optimal number of clusters $K_\gamma = 8$, which gives a reliable representation of the analyzed scene, showing several distinguishable behaviors. Different temporal trends have been investigated. In particular, we are interested in clusters showing either a significant reduction or an increase of coherence values during the flood event [see Fig. 6(c)]. The former [blue, green, and red coherence time series in Fig. 6(c)] are the most common ones because they are potentially associated with the flood event, so we assign to them a high value of $P(F|C_\gamma) = 0.9$. The latter (orange series) can be found in the presence of double bounce effects: this may occur when water lying below vertical structures is found in both the InSAR sensor passes, as in the case, e.g., of mangroves on wetlands [46], in urban areas [47], or in vegetated fields when the incoming e.m. waves have a sufficient level of penetration beneath the canopy. We found some evidence of the presence of such phenomena in our data and thus assigned $P(F|C_\gamma) = 0.5$ to such cluster. Conversely, very low probabilities $P(F|C_\gamma) = 0.1$ are given to the other clusters characterized by either constantly high or constantly low coherence values.

We underline that the InSAR coherence information seems to be particularly useful to complement intensity data, particularly in the analysis of areas showing an increased value of backscattering signal during the flood. As previously mentioned, we have found that almost 43% of the total analyzed pixels show an increase in their backscatter intensity value during the flood event, so they could be potentially interested by a double bounce effect. By considering also the InSAR coherence information, we found that only about one-third of these points (corresponding to about 13% of the total analyzed pixels) exhibit also a significant change in the coherence value. In this way, a useful reduction of false alarms can be obtained.

C. Flood Map Production and Evaluation

Two different flood maps have been produced by applying (1), reporting the probability values of inundation of each pixel in the analyzed area on December 2 and 3, respectively. As previously mentioned, the conditional probability $p(F|t, d)$ has been modeled as a sigmoid function [see (3)], where the parameters have been set respectively equal to $A = 0.9$, $B = 1.1$, and $C = -0.4$ for December 2 and to $A = 0.2$, $B = 0.3$, and $C = 2$ for December 3. These values were obtained through a trial-and-error procedure, taking into account some information about the event evolution (see the Appendix).

The flood maps, depicted in Fig. 7, have been obtained by applying a threshold of 0.5 to each final $p(F = \text{flood}|i = i^*, \gamma = \gamma^*, t = t^*, d = d^*)$ map. Overall, the probability map depicting the flood extension on December 2 reports large areas having high values of probability, mainly localized along the Bradano river, and in good agreement with the information about the flood event, confirming that it was in a crucial phase during

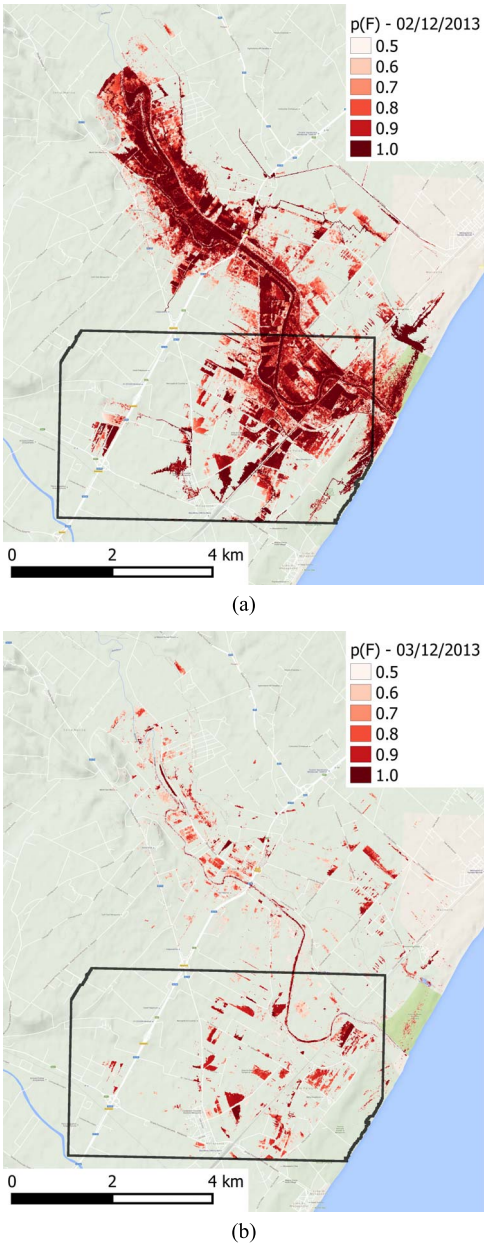

Fig. 7. Flood maps produced by the BN method for (a) the December 2 and (b) the December 3 event, respectively, with the GT area overlapped.

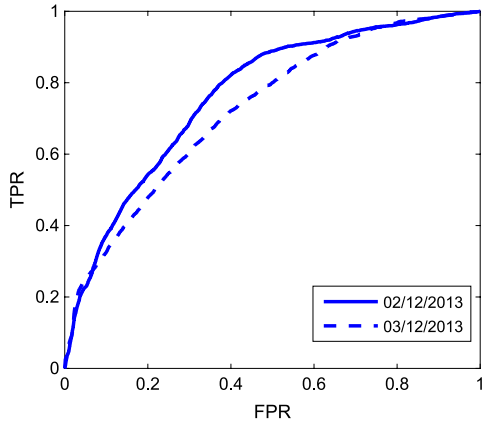

Fig. 8. ROC curves obtained by comparing the probability maps with GT for December 2 and 3. AUC values are 0.77 and 0.73, respectively.

this day. In particular, the BN framework allows to reduce false alarms and missed identifications by performing data fusion. In fact, considering again the pixels exhibiting high values of backscattering intensity (43% of pixels in the analyzed scene), we observed that, by adding both coherence information and ancillary information, they are reduced to only 5%. On the other hand, some areas that we know were flooded but that exhibited high values of backscattering intensity show high values of posterior probability, indicating that the data fusion performed through the BN framework allows to overcome the missed identification which would be obtained by using only one data source. The probability map concerning the flood extension on the following day shows a remarkable reduction of inundated areas, again in agreement with the information about the considered event that was almost concluded on December 3.

The two probabilistic flood maps have been compared with the available reference data, described in Section III-C. The overall accuracies, obtained in correspondence of a probability threshold value equal to 0.5, are respectively equal to 74% and 88% for December 2 and 3. The flood and no-flood instances are very unbalanced on the maps (the extension of the flooded area is relatively small with respect to the analyzed area), so the true positive rate (TPR) and false positive rate (FPR) values are also considered to give more objective evaluations. They are defined as

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

where TP is the number of true positives, i.e., the actual flood data that are correctly classified, FP is the number of false positives, i.e., negative data classified as positive, TN is the number of true negatives, i.e., the actual no-flood data that are correctly classified, and FN is the number of false negatives, i.e., positive data classified as negative. In Fig. 8, the corresponding receiver operator characteristic (ROC) curves are reported: a relative agreement with the reference maps has been obtained, with area under the curve (AUC) values for the December 2 map and for the December 3 map equal to 0.77 and 0.73, respectively. The best overall accuracies, obtained by varying the probability threshold values, are respectively equal to 85% and 89% for December 2 and 3. However, the available reference data are not temporally coincident with the available images and, thus, with the probability maps, so other qualitative and quantitative considerations can be put forward through comparisons with the GT map. In particular, we focus on the individual polygons delimiting fields with uniform labels, assigned during the photointerpretation activity leading to the GT map. Considering the GT fields labeled as "1," i.e., the fields that were not flooded on December 4 and 5 but could be flooded before those dates, we report, in the first row of Table III, the number of pixels, included within each of these fields, showing a posterior probability value greater than 0.5 in each of the two flood maps produced. We know that these areas have been interested by flood before December 4, and from this comparison, we can deduce that the phenomenon concerns principally December 2, while in the following day, the water in these fields has likely been absorbed. This is again consistent with the known temporal evolution of the considered event.

TABLE III
COMPARISON BETWEEN PROBABILITY MAPS AND GT

GT Label	# pxl	# pxl with $P \geq 0.5$ on Dec. 2	# pxl with $P \geq 0.5$ on Dec. 3
1	140,929	64,325 (45.6%)	13,488 (9.6%)
2	40,554	31,599 (77.9%)	13,628 (33.6%)
3	68,382	50,922 (74.5%)	35,019 (51.2%)

In Fig. 9, the GT polygonal fields labeled as "1"in (a), and labeled as "2" and "3" in (b) (see also Fig. 5), are shown: the color of each field represents the percentage of pixels, belonging to that field, with posterior probability values greater than 0.5, computed for December 2. It is interesting to note that there are many areas having more than 80% of pixels with a probability value greater than 0.5 (reddish colors). In particular, these areas are located close to the river, while areas far from the river have lower values (orange and yellow fields). This is true for both the areas labeled as "1" (in panel a) and those labeled as "2" and "3" (in panel b) in the GT map, indicating that, close to the river banks, the inundation had already begun on December 2. In Fig. 10, the percentage of pixels with posterior probability greater than 0.5, belonging to fields labeled as "1" (a) and labeled as "2" and "3" (b) in the GT map, computed for December 3, are reported. In this case, only a few pixels are above the threshold in fields labeled as "1," as shown by their yellow color, while the water is also persistent in these areas, which appear as flooded on December 4 and 5. Finally, we have evaluated the probability map on December 3 also in the following manner: we have considered only the points in the fields labeled as "2" and "3" that exhibit $p(F) > 0.9$ in the map computed for December 2. These pixels are considered as reference data for the flood instance. We have observed that 71% of them have a posterior probability value $p \geq 0.5$ in the map computed for December 3.

V. CONCLUSION

We have shown an application of a BN performing a data fusion approach of SAR intensity imagery, InSAR coherence,

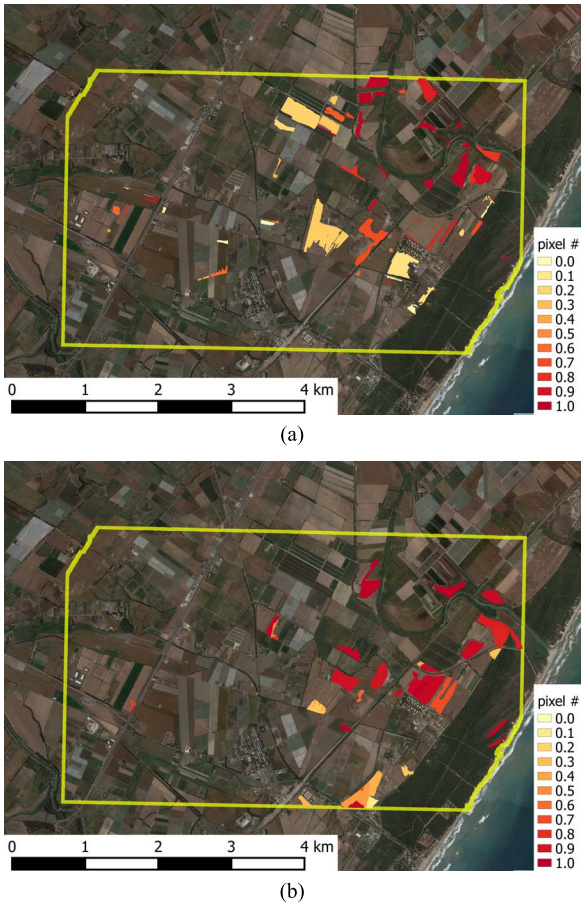

Fig. 9. Percentage of pixels in each field labeled as 1 in (a) and as 2 or 3 in (b), having $p >= 0.5$ in the map computed on December 2.

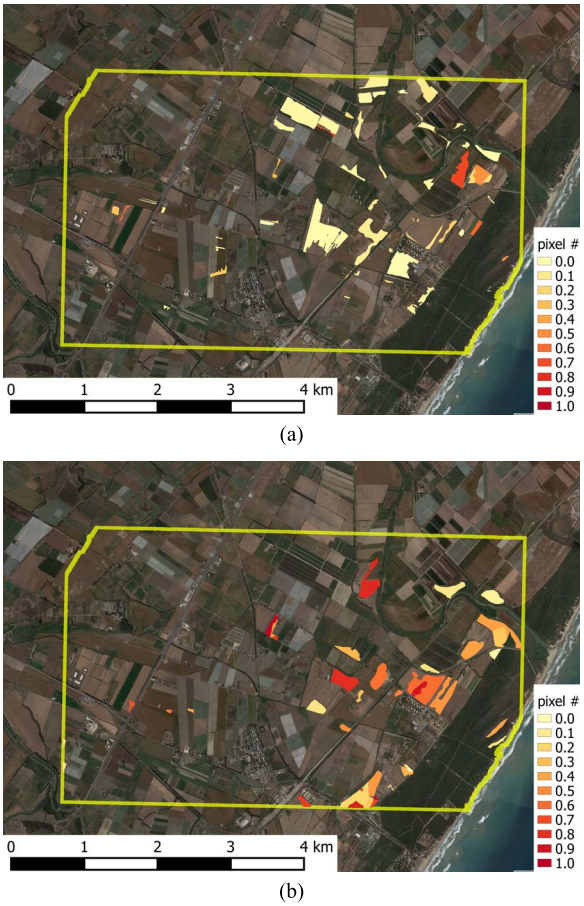

Fig. 10. Percentage of pixels in each field labeled as 1 in (a) and as 2 or 3 in (b), having $p >= 0.5$ in the map computed on December 3.

and ancillary data to detect flooded areas. Multitemporal SAR intensity and InSAR coherence data have been independently segmented, identifying relevant pixel clusters exhibiting homogeneous scattering/coherence classes. Subsequently, data have been combined within the Bayesian framework by assigning *a priori* flood probabilities to individual cluster centroids, and adding some ancillary information, such as the distance from the river course and a GFI.

Experimental results show good capabilities of identification of a large area interested by the flood phenomenon, partially overcoming the obstacle constituted by the presence of scattering/coherence classes corresponding to different land cover types, which respond differently to the presence of water and to inundation evolution. In particular, our BN-based data fusion approach allows to both mitigate the false alarms and to correctly identify flooded areas in events characterized by complex land cover ground conditions and time evolution.

In conclusion, BNs appear to be a powerful tool to perform data fusion in the analysis of complex real-world systems, such as natural hazard detection. Further studies will be conducted by considering the introduction of other variables, such as imagery by optical sensors or other ancillary information. In this respect, another advantage of the BN representation could be exploited, i.e., its modularity. In fact, depending on the independence relations between new and old variables, some local probability models could be reused, and only some conditional probability terms would need to be inserted in the computation, without rewriting the whole joint probability. Furthermore, the proposed methodology exhibits the great advantage that the analysis of conditionally independent variables (which is, at present, the most time-consuming step) can be performed in parallel, thus improving its potential applicability to near-real-time postevent mapping purposes.

APPENDIX
ANCILLARY DATA PARAMETERS

The conditional probability function $p(F|t,d)$ has been modeled as a sigmoid, which is generally written as a function of two free parameters, μ and σ

$$f(y) = \frac{1}{1 + \exp\left[-\sigma \cdot (y - \mu)\right]} \tag{4}$$

where μ is the y value corresponding to $f(y) = 0.5$ and σ is the steepness of the curve. In our case, $p(F|t,d)$ is a function of two independent variables [see (3)], and three parameters, denoted by A, B, and C, have been considered to properly take into account each contribution. This particular form of the *a priori* probability function, as well as its parametrization, is not mandatory for the success of the proposed methodology,

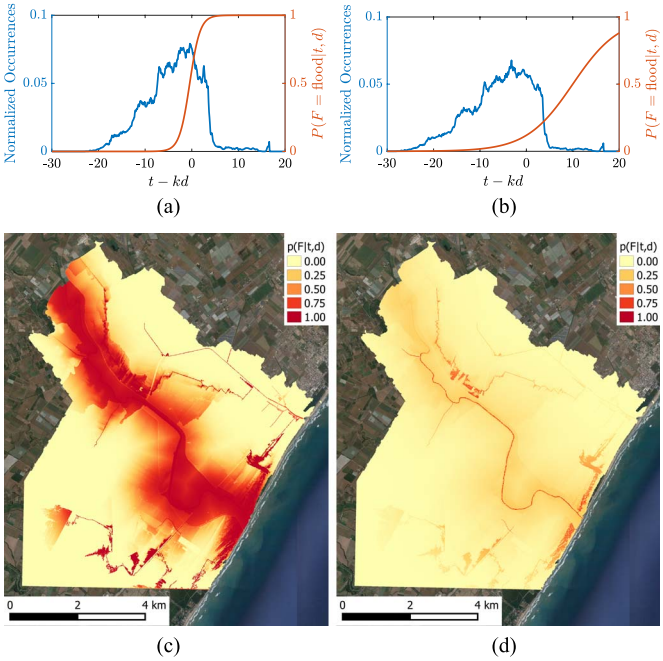

Fig. 11. (a) Ancillary data sigmoid function and data histogram for December 2, 2013. (b) Same as (a) for December 3, 2013. (c) Ancillary data map for December 2, 2013. (d) Same as (c) for December 3, 2013.

and other equivalent forms could be used. In this appendix, we illustrate the reasoning behind the choice of the parameters used in the modeling.

First of all, in order to simplify the parameter setting, the two variables, t, i.e., the GFI, and the distance from the river course, d, have been combined into a single variable $\overline{y} = t - k \cdot d$, which can be interpreted as a "modified" Euclidean distance, modulated by the index t, through a weight factor k which also serves as a normalization constant to bring both variables within a common interval. Therefore, comparing (4) to (3), the following relations hold:

$$At - Bd + C = A\left(\overline{y} + \frac{C}{A}\right) = \overline{\sigma}(\overline{y} - \overline{\mu}) \tag{5}$$

with $A = \overline{\sigma}$, $B = kA$, and $C = \overline{\mu}A$. The k, $\overline{\sigma}$, and $\overline{\mu}$ have been set by considering information about the considered event, and the A, B, and C values are then computed by using the previous relations. In particular, for December 2, we know that the flood was in a crucial phase and was concentrated near the river (about 500 m around the banks), while farther than 5 km, only areas with high values of t could be inundated for rain accumulation (i.e., these areas had not been reached by floodwaters yet). For these reasons, $k = 1.2$ has been first set in order to give a slightly greater weight to the d variable with respect to t. Subsequently, the $\overline{y}$ variable has been considered [see Fig. 11(a)], and the information about the event has been conveyed into the other variables in the following way:

1) $\overline{\mu} = -0.35$, by considering that, in this way, values $p(F|t,d) > 0.5$ are assigned to a terrain belt around the river, wide about 500 m;
2) $\overline{\sigma} = 0.9$, which corresponds to setting $p(F|t,d) \simeq 0$ on areas farther than about 5 km from the river.

These positions, together with $k = 1.2$, give the A, B, and C values reported in Section IV and the ancillary data map depicted in Fig. 11(c). Conversely, on December 3, the flood event was in a decreasing phase, and the water was flowing away from the river, mostly concentrated in areas where it persisted for accumulation. For this reason, $k = 1.5$ has been set in order to give a slightly greater weight to areas corresponding to greater values of the d variable. Successively, by considering the $\overline{y}$ variable distribution [see Fig. 11(b)], the following values have been set:

1) $\overline{\sigma} = 0.2$, in order to expand the function steepness and assign to the ancillary information a lower weight with respect to the imagery information;
2) $\overline{\mu} = 10$, which corresponds to setting $p(F|t,d) = 0.5$ on areas with great accumulation.

These positions give the A, B, and C values reported in Section IV and the ancillary data map depicted in Fig. 11(d).

ACKNOWLEDGMENT

COSMO-SkyMed images are courtesy of the Italian Space Agency (ASI). The authors would like to thank Ing. L. Candela, ASI, for the support in the data acquisition and Dr. N. De Musso for the help in the optical image analysis and interpretation. InSAR processing was performed by Dr. D. O. Nitti of GAP s.r.l.

REFERENCES

[1] S. Serpico *et al.*, "Information extraction from remote sensing images for flood monitoring and damage evaluation," *Proc. IEEE*, vol. 100, no. 10, pp. 2946–2970, Oct. 2012.
[2] S. Martinis, A. Twele, and S. Voigt, "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.*, vol. 9, no. 2, pp. 303–314, Mar. 2009.
[3] D. Mason, I. Davenport, J. Neal, G. Schumann, and P. D. Bates, "Near real-time flood detection in urban and rural areas using high resolution synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3041–3052, Aug. 2012.
[4] J. Sanyal and X. Lu, "Application of remote sensing in flood management with special reference to monsoon Asia: A review," *Nat. Hazards*, vol. 33, no. 2, pp. 283–301, Oct. 2004.
[5] D. Frey, M. Butenuth, and D. Straub, "Probabilistic graphical models for flood state detection of roads combining imagery and DEM," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 6, pp. 1051–1055, Nov. 2012.
[6] Y. Wang, J. Colby, and K. Mulcahy, "An efficient method for mapping flood extent in a coastal floodplain using Landsat TM and DEM data," *Int. J. Remote Sens.*, vol. 23, no. 18, pp. 3681–3696, Jan. 2002.
[7] D. Mason *et al.*, "Flood detection in urban areas using TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 882–894, Feb. 2010.
[8] L. Pulvirenti, M. Chini, N. Pierdicca, L. Guerriero, and P. Ferrazzoli, "Flood monitoring using multi-temporal COSMO-SkyMed data: Image segmentation and signature interpretation," *Remote Sens. Environ.*, vol. 115, no. 4, pp. 990–1002, Apr. 2011.
[9] S. Dellepiane and E. Angiati, "A new method for cross-normalization and multitemporal visualization of SAR images for the detection of flooded areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2765–2779, Jul. 2012.
[10] P. Matgen *et al.*, "Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies," *Phys. Chem. Earth*, vol. 36, no. 7/8, pp. 241–252, 2011.
[11] F. Covello *et al.*, "COSMO-SkyMed—An existing opportunity for observing the Earth," *J. Geodyn.*, vol. 49, no. 3/4, pp. 171–180, Apr. 2010.
[12] N. Pierdicca, L. Pulvirenti, M. Chini, L. Guerriero, and L. Candela, "Observing floods from space: Experience gained from COSMO-SkyMed observations," *Acta Astron.*, vol. 84, pp. 122–133, Mar./Apr. 2013.

[13] A. Refice *et al.*, "Inundation monitoring through high-resolution SAR/InSAR data and 2D hydraulic simulations," in *Proc. EARSel*, 2013, pp. 15–22.
[14] G. Nico, M. Pappalepore, G. Pasquariello, A. Refice, and S. Samarelli, "Comparison of SAR amplitude vs. coherence flood detection methods—A GIS application," *Int. J. Remote Sens.*, vol. 21, no. 8, pp. 1619–31, Jan. 2000.
[15] S. Dellepiane, G. Bo, S. Monni, and C. Buck, "SAR images and interferometric coherence for flood monitoring," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2000, vol. 6, pp. 2608–2610.
[16] A. Refice *et al.*, "SAR and InSAR for flood monitoring: Examples with COSMO-SkyMed data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 7, pp. 2711–2722, Jul. 2014.
[17] N. Pierdicca *et al.*, "Flood mapping by SAR: Possible approaches to mitigate errors due to ambiguous radar signatures," in *Proc. IEEE IGARSS*, 2014, pp. 3850–3853.
[18] B. Dasarathy, "Information fusion: What, where, why, when and how?" *Inf. Fusion*, vol. 2, no. 2, pp. 75–76, Jun. 2001.
[19] G. Schumann, P. Bates, M. Horritt, P. Matgen, and F. Pappenberger, "Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models," *Rev. Geophys.*, vol. 47, no. 4, Dec. 2009, Art. no. RG4001.
[20] P. A. Brivio, R. Colombo, M. Maggi, and R. Tomasoni, "Integration of remote sensing data and GIS for accurate mapping of flooded areas," *Int. J. Remote Sens.*, vol. 23, no. 3, pp. 429–441, Jan. 2002.
[21] P. Matgen, G. Schumann, J. Henry, L. Hoffmann, and L. Pfister, "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 Observ. Geoinf.*, vol. 9, no. 3, pp. 247–263, Aug. 2007.
[22] L. Pulvirenti, N. Pierdicca, G. Boni, M. Fiorini, and R. Rudari, "Flood damage assessment through multitemporal COSMO-SkyMed data and hydrodynamic models: The Albania 2010 case study," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 7, pp. 2848–2855, Jul. 2014.
[23] D. Barber, *Bayesian Reasoning and Machine Learning*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[24] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA, USA: MIT Press, 2009.
[25] D. Qin, M. Jianwen, and O. Y. Yun, "Remote sensing data change detection based on the CI test of Bayesian networks," *Comput. Geosci.*, vol. 32, no. 2, pp. 195–202, Mar. 2006.
[26] D. Straub, "Natural hazards risk assessment using Bayesian networks," in *Proc. Proc. ICOSSAR*, A. *et al.*, Ed., 2005, pp. 2509–2516.
[27] A. Gret-Regamey and D. Straub, "Spatially explicit avalanche risk assessment linking Bayesian networks to a GIS," *Nat. Hazards Earth Syst. Sci.*, vol. 6, no. 6, pp. 911–926, Oct. 2006.
[28] P. Papakosta and D. Straub, "A Bayesian network approach to assessing wildfire consequences," in *Proc. ICOSSAR*, A. *et al.*, Ed., 2013, pp. 1–8.
[29] K. Vogel, C. Riggelsen, O. Korup, and F. Scherbaum, "Bayesian network learning for natural hazard analyses," *Nat. Hazards Earth Syst. Sci.*, vol. 14, no. 9, pp. 2605–2626, Sep. 2014.
[30] S. Manfreda, C. Samela, A. Sole, and M. Fiorentino, "Flood-prone areas assessment using linear binary classifiers based on morphological indices," in *Vulnerability, Uncertainty, and Risk*. Reston, VA, USA: ASCE, 2014, pp. 2002–2011.
[31] S. Manfreda *et al.*, "Flood-prone areas assessment using linear binary classifiers based on flood maps obtained from 1D and 2D hydraulic models," *Natural Hazards*, vol. 79, no. 2, pp. 735–754, Nov. 2015.
[32] C. Bishop, *Neural Networks for Pattern Recognition*. London, U.K.: Oxford Univ. Press, 1995.
[33] E. Rignot and J. van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 4, pp. 896–906, Jul. 1993.
[34] J. Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1017–1028, Sep. 1994.
[35] R. Touzi and A. Lopes, "Statistics of the Stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 519–531, Mar. 1996.
[36] F. Boenzi *et al.*, "Late pleistocene-holocene landscape evolution in Fossa Bradanica, Basilicata (Southern Italy)," *Geomorphology*, vol. 102, no. 3/4, pp. 297–306, Dec. 2008.
[37] M. Fiorentino, S. Manfreda, and V. Iacobellis, "Peak runoff contributing area as hydrological signature of the probability distribution of floods," *Adv. Water Resources*, vol. 30, no. 10, pp. 2123–2134, Oct. 2007.
[38] M. Piccarreta, A. Pasini, D. Capolongo, and M. Lazzari, "Changes in daily precipitation extremes in the Mediterranean from 1951 to 2010: The Basilicata region, southern Italy," *Int. J. Climatol.*, vol. 33, no. 15, pp. 3229–3248, Dec. 2013.
[39] H. A. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE. Trans. Geosci. Remote Sens.*, vol. 30, no. 5, pp. 950–959, Sep. 1992.
[40] B. Kampes, R. Hanssen, and Z. Perski, "Radar interferometry with public domain tools," in *Proc. FRINGE*, 2003, pp. 1–10.
[41] D. O. Nitti, R. Hanssen, A. Refice, F. Bovenga, and R. Nutricato, "Impact of DEM-assisted coregistration on high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 112–743, Mar. 2011.
[42] C. Deledalle, L. Denis, and F. Tupin, "NL-InSAR: Nonlocal interferogram estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1441–1452, Apr. 2011.
[43] Copernicus Emergency Management Service. [Online]. Available: http://emergency.copernicus.eu/.
[44] "Meteoblog di Meteomarconia.it (in Italian), 2014. [Online]. Available: www.meteomarconia.it/meteoblog/index.php/2013/11/16/situazione-delle-precipitazioni-in-basilicata-dalle-ore-000-del-16-novembre-2013.
[45] L. Pulvirenti, F. Marzano, N. Pierdicca, S. Mori, and M. Chini, "Discrimination of water surfaces, heavy rainfall, and wet snow using COSMO-SkyMed observations of severe weather events," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 858–869, Feb. 2014.
[46] S.-H. Hong, S. Wdowinski, S.-W. Kim, and J.-S. Won, "Multi-temporal monitoring of wetland water levels in the Florida Everglades using interferometric synthetic aperture radar (InSAR)," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2436–2447, Nov. 2010.
[47] D. Mason, L. Giustarini, J. Garcia-Pintado, and H. Cloke, "Detection of flooded urban areas in high resolution synthetic aperture radar images using double scattering," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 28, pp. 150–159, May 2014.

Annarita D'Addabbo received the Laurea degree *(cum laude)* and the Ph.D. degree in physics from the University of Bari, Bari, Italy, in 1998 and 2010, respectively.

Since 1999, she has been with the Italian Research Council, where she is currently a Researcher with the Istituto di Studi sui Sistemi Intelligenti per l'Automazione, Bari. Her main research interests are in the area of pattern recognition and machine learning (classification, regression and estimation, and data fusion) with applications both to remote-sensing images and to biological data.

Alberto Refice received the Laurea degree *(cum laude)* and the Ph.D. degree in physics from the University of Bari, Bari, Italy, in 1994 and 1998, respectively.

Since 2001, he has been with Consiglio Nazionale delle Ricerche, and since 2003, he has been with the Istituto di Studi sui Sistemi Intelligenti per l'Automazione, Bari, as a Researcher. He has been and is involved in many research projects devoted to the development of algorithms for multitemporal point and distributed target processing for radar interferometry, synthetic aperture radar data analysis, multisource data fusion, and geographical information system data management and integration. His research interests concern advanced processing techniques for remote sensing data and their applications to the retrieval of geophysical parameters for environmental hazard monitoring. He is also active on geomorphological terrain analysis and landscape evolution modeling. He has authored and coauthored more than 25 among refereed journal papers and book chapters. In 2010 and 2011, he held the Chair of Remote Sensing principles for the Geology Degree MsC Course, at the University of Bari.

Guido Pasquariello received the Laurea degree in physics from the University of Bari, Bari, Italy, in 1975.

In 1976, he joined the CSATA, where he worked in the field of statistical data analysis. From 1977 to 1978, he was a Scientific Fellow of the Commission of European Communities at the Central Bureau for Nuclear Measurements, Geel, Belgium. From 1978 to 1980, he joined the National Laboratory of Frascati, Rome, of the Italian Institute of Nuclear Physics. From 1980 to 1985, he joined the TECNOPOLIS Science Park of Bari, where he was the leader of projects related to the classification of satellite remote sensing images. Since 1985, he has worked as a Senior Scientist at the Institute for Signal and Image Processing of the Italian National Research Council (CNR), Bari. His research interests include applications of artificial intelligence tools and neural networks to digital image processing in the field of computer vision, remote sensing, and medical data. He is the author or coauthor of more than 150 scientific publications, including journals and conference proceedings. In 1994, he was a member of the International Users Committee of Radiometer Vegetation on SPOT-4 satellite. In 1997, he was a member of the Working Group for the definition of the Italian Space Plan, 1998–2002. From 2006 to 2008, he was the President of the scientific and technological park Tecnopolis (Valenzano, Bari, Italy). From 2009 to 2013, he was a member of the board of directors of InnovaPuglia spa, the Regione Puglia in-house society for innovation in the field of ICT. From 2010 to 2013, he was a member of the Consiglio Tecnico-Scientifico of the Italian Space Agency (ASI).

Francesco P. Lovergine received a degree in computer science from the University of Bari, Bari, Italy, in 1991. Since his thesis activity, he worked on artificial vision methods, with specific interest for visual reconstruction of surfaces and application of approximation theory to vision systems and data fusion.

Since 1997, he has been a Researcher at the Italian National Research Council (CNR)-Istituto di Studi sui Sistemi Intelligenti per l'Automazione, Bari, and he gradually moved his area of expertise to distributed computing in the remote sensing context and the use and implementation of geographical information system (GIS) applications. His current interests are about the representation and understanding of geographical information, open GIS standards and protocols, and free software for GIS. He is a member of ACM, FSFE, and OSGEO. Since 1999, he works in his spare time for the Debian Project and is a cofounder and the current team leader of the DebianGIS group. Most of his contributions are currently among the main free and open-source software products in the GIS area.

Domenico Capolongo received the Ph.D. degree in geomorphology and environmental dynamics from the University of Bari, Bari, Italy, in 1999 and the Master's degree in environmental engineering from the Polytechnic of Turin, Turin, Italy, in 1994.

Since 2002, he has worked at the Earth and Environmental Science Department, University of Bari. His research interests are in the field of geomorphology, particularly on the spatial and temporal distribution of erosion processes at different scales in high-rate evolving landscapes. He uses geographical information system, remote sensing, and computer simulation as principal tools for qualitative and quantitative insight into complex systems and as a virtual laboratory to explore theory by simulation. His studies also focus on the geomorphological hazards and risks. He is a member of the Italian Association of Geomorphology (AIGEO).

Salvatore Manfreda received the Master's degree in civil engineering and the Ph.D. degree in "methods and techniques for environmental monitoring" from the University of Basilicata, Matera, Italy, in 2001 and 2004, respectively.

He started his career as a Researcher Associate at Princeton University, Princeton, NJ, USA (2004–2005). He became Assistant Professor in 2008 and Associate Professor in Hydraulic Constructions and Hydrology in 2014 at the University of Basilicata. He has a broad interest in hydrology and ecohydrology, with particular emphasis on flood prediction, delineation of flood prone areas, stochastic processes in hydrology, space-time rainfall dynamics, soil moisture process, low flow mathematical filters, plant physiology, and vegetation patterns. His research activity has produced more than 90 publications in peer review journals, book chapters, and international and national proceedings.