Scientific Research Publishing

A Scalable Synthesis of Multiple Models of Geo Big Data Interpretation

Alessia Goffi[1,2], Gloria Bordogna[1], Daniela Stroppiana[1], Mirco Boschetti[1], Pietro Alessandro Brivio[1]

[1]CNR IREA, Milano, Italy
[2]Terraria s.r.l., Milano, Italy
Email: bordogna.g@irea.cnr.it

Abstract

The paper proposes a scalable fuzzy approach for mapping the status of the environment integrating several distinct models exploiting geo big data. The process is structured into two phases: the first one can exploit products yielded by distinct models of remote sensing image interpretation defined in the scientific literature, and knowledge of domain experts, possibly ill-defined, for computing partial evidence of a phenomenon. The second phase integrates the partial evidence maps through a learning mechanism exploiting ground truth to compute a synthetic Environmental Status Indicator (ESI) map. The proposal resembles an ensemble approach with the difference that the aggregation is not necessarily consensual but can model a distinct decision attitude in between pessimistic and optimistic. It is scalable and can be implemented in a distributed processing framework, so as to make feasible ESI mapping in near real time to support land monitoring. It is exemplified to map the presence of standing water areas, indicator of water resources, agro-practices or natural hazard from remote sensing by considering different models.

Keywords

Environmental Status Indicators, Ordered Weighted Averaging Operators, Machine Learning, Decision Attitude

1. Introduction

Nowadays, Information and Communication Technologies are mature to manage and share geo big data on the Web by coping with huge volumes, variable velocity of creation rates, great variety and complexity of both data structures, formats, and semantics.

As far as the storage of huge volumes of geo big data, the horizontal scaling architecture with many small/medium size servers connected through a network has established over the vertical scaling architecture, consisting of a huge centralized server, for its greater and more flexible expandability at lower costs, its great resilience to faults, and its suitability to implement distributed processing frameworks [1]. The distributed processing paradigm, based on distributed file system and map-reduce, allows to process huge volumes of data efficiently, by partitioning independent tasks that can be executed in parallel by distinct mappers installed on independent machines, and by combining their reorganized results by means of reducers, installed on distinct machines [2] [3]. NoSQL databases have demonstrated to effectively manage geospatial data with different data structure, both in the form of semi-structured information, and grid and raster data [4]. Finally, Web geo services of the Open Geospatial Consortium (OGC) provide standards interoperable methods for sharing on the Web distributed and heterogeneous geo big data [5].

Nevertheless, when analyzing multi-source heterogeneous geo big data to capture the status of environmental phenomena, data redundancy might cause inconsistent information about the representation of the phenomenon, which may induce doubts on both data reliability and suitability for taking decisions to benefit territory management [6]. This is particularly true in the case of information derived from remote sensing processing and analysis for territorial monitoring and managing. Different preprocessing procedures, domain experts, models and methods, can yield different resulting products depicting the same phenomenon so that it becomes questionable which one to trust and to communicate to the final stakeholders.

In order to cope with such problems, especially when exploiting remote sensing derived information and *in situ* data, flexible human-interpretable synthesis of geo big data could reconcile results yielded by distinct models and experts, generating a kind of ensemble environmental status maps [7]. In order to be practical for territorial monitoring, the process generating the synthesis should be applicable in near-real time, to help decision makers to understand ongoing phenomena and plan timely mitigation measures.

In this paper, we propose a human interpretable synthesis of environmental status indicator (ESI) maps, obtained from the ensemble of products derived from distinct models of remote sensing data interpretation. Since each product can be defined on a distinct domain, remote sensing knowledge, possibly ill-defined, derived from either the literature or data analysis performed by experts, is exploited to compute an independent map from each product. In our experimental conditions these products can also be regarded as partial evidence of a phenomenon brought by different indicators. This is done in the first phase of the proposed algorithm framework by specifying (soft) constraints, defined by membership functions of fuzzy sets on the domain of the products taking values in [0, 1], so as to achieve both normalization of the products domains and (fuzzy) segmentation of the products. Subsequently, in the second phase, the synthesis of different

maps describing the phenomenon under analysis is performed by aggregating them based on an Ordered Weighted Averaging (OWA) operator [8]. OWA operators are a family of non-linear mean-like fuzzy operator [8], that can be learned by an adaptation mechanism exploiting ground truth data (GTD) available in a region of interest (ROI) [9]. The final ESI map is a more robust and reliable representation of the phenomenon than each single product alone. Besides this advantage, the semantics of the synthesis can also be expressed linguistically to describe a decision attitude, thus achieving human interpretability. The learning mechanism was proposed in [10] exploiting GTD to learn the best OWA operator for a given area by iteratively minimizing error between OWA results with respect to the GTD. The novelty of the current paper with respect to what has been proposed in [10] is the use of crisp segmentation thresholds as defined in the literature to compute binary partial evidence maps in the first phase, and an alternative approach to define the OWA aggregation by choosing a desired decision attitude expressed linguistically.

With this experiment we mimic the basic conditions that can occur in operational decision making for territorial management: 1) use of standard non-customized products and 2) need to produce a synthesis according to a specific attitude towards facing decisions. As far as point 1 GIS technicians can already use available maps such as the Copenicus products or autonomously produced layers from specific indicators/indexes computed on the basis of available models in literature and/or published thresholds. Usually point 2 is performed in a supervised way by a manual interaction of the user.

The proposal is exemplified to map the status of standing water areas by aggregating the results of eight distinct models defined in the literature in three regions of interest in Northern Italy, characterized by distinct environmental conditions.

In the following section the related literature is outlined. In Section 3 and outline of the proposal is described. Section 4 illustrates the case study and the obtained results. Finally the conclusion summarizes the main achievements.

2. Related Literature

Currently, machine learning techniques, namely deep and convolutional neural networks [11], are the most up-to-date approaches to achieve geo big data dimensionality reduction to the purposes of scene classification, object and anomaly detection [12] [13] [14]. They are data-driven approaches requiring a preliminary training phase in which, given a set of ground truth data, they learn the classifier which can be subsequently applied on the entire ROI. Although these approaches demonstrated to be very successful in several contexts, they are opaque mechanisms, not explicating the classification rules they automatically learn during the training phase. Moreover, in order to train properly the algorithms, the training data set must be large enough and representative: this is necessary in order to avoid overfitting, that is the situation in which the learned model is unable to generalize to new examples that were not in the training set [15]. Never-

theless, GTD are generally scarce and sparse and in many rural areas remote sensing data are the only observations available. This poses a major impediment to the application of deep learning approaches for ESI mapping from remote sensing. Finally, one main problem is that by changing the ROI or the sensors one generally needs to repeat the training phase with GDT of the new ROI, since transfer of a pre-trained network greatly depends on the choice of a proper network architecture for the target purpose [16]. This is the reason why more traditional approaches defined based on expert's knowledge are still appealing. These approaches for environmental status assessment rely on models of the physical interaction of the electromagnetic radiation with the environmental phenomenon of interest. In the literature different models were proposed which defined distinct Spectral Indexes (SI) on a real domain to reduce a large multispectral image to a more compact form that captures a specific "shape" of the spectral signature of the environmental variable of interest: each model defines and applies a distinct function combining the band signals, to generate SI maps which are finally properly segmented to identify vegetation vigor, bare soil areas, water bodies, and so on [17]. Although these methods have the advantage of being human explainable, they are often ineffective to describe complex phenomena, such as flooded areas and burned areas, for several reasons. First of all, they need an accurate calibration phase to determine the best segmentation threshold that may vary depending on ROI. In facts, many environmental phenomena have a different appearance when changing the geographic context and observation conditions (presence of clouds, shadows, specific land covers, etc.). A single model may be not sufficient to capture all aspects of a given phenomenon. For example, to identify all types of water such as shallow water, deep water, wetlands, rivers and inland water bodies, rice flooded fields many distinct SIs have been defined [18]-[23]. Using distinct models to map a given phenomenon may result in redundant or conflicting maps. Furthermore, not all models yield results defined on the same domain, so one needs to normalize the values to compare them. These are the main reasons that motivated our proposal of a scalable synthesis of distinct models based on a kind of flexible data-driven ensemble approach.

3. Outline of the Scalable Synthesis of Multiple Models for ESI Mapping

The starting point of our proposal is exploiting multiple models defined in the literature for mapping a given environmental variable describing a phenomenon of interest. These models may compute products, *i.e.*, images, in which pixel values are defined on distinct domains to represent the status of the phenomenon. In some cases the models also indicate the proper thresholds on the pixel domains that should be applied to segment the phenomenon footprint with greatest accuracy.

In the following we regard these products as distinct contributing factors which can concur or complement one another [24] to determine the ESI map describing the spatial evidence of the studied phenomenon to distinct extent.

The objective is to integrate the products to generate an ESI map that is both synthetic and robust.

Synthetic means that we want to generate an ESI map that offers a manageable and comprehensive representation of the phenomenon. Robust means that the ESI accuracy should be stable by changing ROI. We first exploit domain knowledge in order to compute from each product a partial evidence map of the phenomenon. Then, we fuse the partial evidence maps by means of fuzzy aggregation operators, with behavior that can be flexibly tuned in between that of the intersection and union, *i.e.*, and Ordered Weighted Averaging operator [8]. OWA operators allow to define fusion strategies with distinct mean like semantics ranging from the minimum to the maximum of the values they aggregate. The mean like nature of fusion strategies has been outlined by many authors and is recognized as particularly useful in the context of spatial decision making [25]-[30]. In our original proposal [24] applied for many distinct purposes [31] [32] [33] ill-defined knowledge obtained from remote sensing experts analyzing GTD was used to maps partial evidence of a phenomenon. The experts were also asked to choose an aggregation operator to generate the synthetic ESI map. In the evolution of this approach described in [10] ill-defined experts' knowledge was still used in combination with an adaptive mechanism that exploits reliable georeferenced observations of the phenomenon in a ROI to learn the most suitable OWA operator to apply in the ROI for computing the ESI map. In this approach the georeferenced observations are used as GTD. As discussed in [10] a remarkable aspect of this data-driven phase is that the algorithm does not require a huge amount of GTD to converge.

In the present paper, instead of relying on experts' knowledge we propose the use of distinct models defined in the literature to assess a phenomenon and the synthesis of their results based on the adaptive mechanism defined in [10].

Furthermore, we offer an alternative approach to select the OWA operator, that consists in suggesting OWA operators that model desired decision makers' attitudes towards facing the occurrence of the phenomena. Representative OWA operators are defined and associated with distinct linguistic labels, each one expressing a decision attitude towards facing risks. Namely, decision makers can be pessimistic or optimistic on the possible occurrence of the worst scenario, or can have an attitude in between these two extremes. The case of pessimism the aggregation must yield an ESI map that depicts the worst possible scenario that might occur, while in the case of optimism the ESI map should depict the best possible scenario. Furthermore, decision makers may trust all models or just one of them: the most pessimistic model in case of full pessimistic attitude or the most optimistic one in case of full optimistic attitude.

The automatic algorithm is structured into two phases and we will illustrate how it can be implemented in a distributed processing framework. Finally the proposal will be exemplified by a case study for mapping standing water from remote sensing.

1) *Ordered Weighted Averaging Operators* (*OWA*) *and Decision attitude*

The seminal paper [34], stemming from the consideration that "*the efficient use of decision support systems* (*DSSs*) *is to assist and help humans arrive at a proper decision, but by no means, to replace humans*" proposes to introduce some linguistic interaction between the human and machine. To this end, the author defines the fuzzy logic-based calculi of linguistically quantified propositions as a viable means for expressing human interpretable decisions [35].

In [36] the problem to define an overall decision function aggregating degrees of satisfaction of multiple criteria (in our context, results of the application of distinct models), was proposed based on the Ordered Weighted Averaging operators. An OWA of dimension N and weighting vector W, with $\sum_{i=1,\cdots,N} w_i = 1$, aggregates N values $[d_1,\cdots,d_N]$, and computes an aggregated value a in [0, 1] as follows [8]:

$$\text{OWA}:[0,1]^N \rightarrow [0,1]$$

such that

$$a = \text{OWA}\left([d_1,\cdots,d_N]\right) = \sum_{i=1,\cdots,N} w_i * g_i \tag{1}$$

in which g_i is the ith largest value of the d_1, d_N.

A fundamental aspect of the OWA is the reordering of its arguments so that the weight w_i is not associated with an argument d_i but rather with a particular rank of the arguments in decreasing order. OWA operators comprehend the Max, the Min and the arithmetic mean operators for the appropriate selection of the weighting vector W:

For $W = [1,0,\cdots,0]$, $\text{OWA}\left([d_1,\cdots,d_N]\right) = \max\left([d_1,\cdots,d_N]\right)$

For $W = [0,\cdots,0,1]$, $\text{OWA}\left([d_1,\cdots,d_N]\right) = \min\left([d_1,\cdots,d_N]\right)$

For $W = [1/N,\cdots,1/N]$, $\text{OWA}\left([d_1,\cdots,d_N]\right) = \frac{1}{N}\sum_{j=1}^{N} d_j$

It can be proved that OWA operators satisfy the commutativity, monotonicity and idempotency and are bounded by *Max* and *Min* operators [8]:

$$\min\left([d_1,\cdots,d_N]\right) \le \text{OWA}\left([d_1,\cdots,d_N]\right) \le \max\left([d_1,\cdots,d_N]\right)$$

To characterize the decision attitude modelled by an OWA operator with weighting vector W, two measures have been introduced in [8]: the measures of orness and of dispersion.

The measure of its $orness(W) \in [0,1]$ is defined as follows:

$$orness(W) = \frac{1}{N-1}\left(\sum_{j=1}^{N}(N-j)*w_j\right) \tag{2}$$

This measure characterizes the degree to which the aggregation is like an *OR* (*Max* operator).

It can be shown that, when the argument values $d_1,\cdots,d_N$ are degrees of partial evidence of an anomaly of an environmental phenomenon from N distinct sources, *i.e.*, the greater they are the more severe the anomaly, we have the following interpretations [24] [25] [36]:

- $orness[1,0,\cdots,0]=1$ indicates a pessimistic attitude applied when one wants to obtain the worst scenario by minimizing the risk of underestimating the effects of a critical phenomenon that can cause both danger to infrastructures and people's safety (*i.e.*, nothing is disregarded, any single source alone is trusted and taken into consideration to plan preparedness and mitigation interventions so as to minimize the occurrence of risky events);

- $orness[0,\cdots,0,1]=0$ indicates an optimistic attitude, applied when one wants the best scenario by minimizing false alarms due to overestimation of the effects of a critical phenomenon (*i.e.*, one wants to prioritize preparedness and mitigation interventions only to anomaly situations pointed out by all sources since any source alone is not trusted by itself);

- $orness[1/N,\cdots,1/N]=0.5$ indicates a balanced and neutral attitude towards risk-prone and risk-adverse.

Another measure can be defined to qualify the semantics of an OWA operator depending on the form of the weighting vector: the *dispersion* measure. This measure represents how much of the information in all the arguments is used by an OWA with weighting vector *W*. The idea behind its definition is that the greater the dispersion the more democratic is the aggregation of the correspondent OWA since it uses information from more sources. Several dispersion measures have been proposed, the first of which is based on the concept of entropy of *W*. Notably in [37] the following definition of dispersion of an OWA operator is defined in [0, 1]:

$$dispersion(W)=1-\max_{i=1,\cdots,N}(w_i) \tag{3}$$

We see that *dispersion*(*W*) is clearly symmetric. When it is zero it means that only one source is considered, *i.e.*, we have a monarchical aggregation, *i.e.*, the rule of one. The greater its value, the more the result is determined by more sources, and thus we have a more democratic aggregation.

In order to explicit the decision attitude modeled by an OWA aggregation with weighting vector *W* one computes the degrees of *orness*(*W*) and *dispersion*(*W*) as defined in formulae (2) and (3) respectively, and then can compare these values with the conditions in the rows and columns of Table 1 to identify the correspondent decision attitude. Conversely, in Table 2 a representative OWA operator can be selected to define an aggregation that models a desired decision attitude expressed linguistically.

In Table 1 and Table 2, the value of an argument to aggregate is a degree of evidence of a critical phenomenon; then, a high value is considered a pessimistic evaluation of what is occurring, while a low value is an optimistic evaluation. The rationale is that since the values to aggregate are evidence degrees of some undesired phenomenon, such as a flood occurrence, a wild fire occurrence, etc., high values/low values have a negative/positive flavor. This interpretation is opposite with respect to what happens in Multi Criteria Decision Making context in which high/low values are considered optimistic/pessimistic evaluations.

Table 1. Decision attitude defined by orness and dispersion in the case of aggregation of $N = 8$ models.

$N = 8$		Dispersion (W) (Δ)				
		0.00	<Δ<	0.44	<Δ<	0.88
Orness (W) (F)	0	*Monarchical & Optimistic*				
	<Φ<	*Monarchical & Towards Optimistic*	*Semi Monarchical & Towards Optimistic*	*Semi Monarchical/Democratic & towards Optimistic*	*Semi Democratic & Towards Optimistic*	*Democratic & Towards Optimistic*
	0.5	*Monarchical & Neutral*	*Semi Monarchical & Neutral*	*Semi Monarchical/ Democratic & Neutral*	*Semi Democratic & Neutral*	*Democratic & Neutral*
	<Φ<	*Monarchical & Towards Pessimistic*	*Semi Monarchical & Towards Pessimistic*	*Semi Monarchical/ Democratic & towards Pessimistic*	*Semi Democratic & Towards Pessimistic*	*Democratic & Towards Pessimistic*
	1	*Monarchical & Pessimistic*				

Table 2. OWA operators modeling a decision attitude satisfying conditions on the *orness* and *dispersion* as shown in **Table 1**.

Decision attitude	Representative weighting vector W of dimension N of the OWA operator
Monarchical & Optimistic	$w_N = 1$
Monarchical & Pessimistic	$w_1 = 1$
Monarchical & Neutral	for $N = 2k + 1$ $\quad w_{quotient(N/2)+1} = 1$
Democratic & Neutral	$w_i = 1/N \quad i = 1, \cdots, N$
Semi Monarchical & Neutral	$w_1 = w_N = 0.5$
Semi Democratic & Neutral	$w_i = 1/(N-2) \quad and \quad i = 2, \cdots, N-1$
Semi Monarchical/Democratic & Neutral	$w_1 = w_N = 0.25$ $w_i = 0.5/(N-2) \quad and \quad i = 2, \cdots, N-1$
Monarchical & Towards Optimistic	For $N = 2k + 1 \quad w_i = 1 \quad and \ quotient \ (N/2+1) < i < N$ (the greater the i the more the optimism) For $N = 2k \quad w_i = 1 \quad and \quad N/2 < i < N$ (the greater the i the more the optimism)
Monarchical & Towards Pessimistic	For $N = 2k + 1 \quad w_i = 1 \quad and \quad 1 < i < quotient(N/2+1)$ (the greater the i the more the optimism) For $N = 2k \quad w_i = 1 \quad and \quad 1 < i < N/2$ (the smaller the i the more the pessimism)
Semi Monarchical & Towards Optimistic	$w_1 = 1 - \alpha \quad w_N = \alpha \quad and \quad 0.5 < \alpha < 1$ the greater the α the more the optimism. An alternative choice can be $w_N = w_{N-1} = 0.5$
Semi Monarchical & Towards Pessimistic	$w_1 = \alpha \quad w_N = 1 - \alpha \quad and \quad 0.5 < \alpha < 1$ the greater the α the more the pessimism. An alternative choice can be $w_1 = w_2 = 0.5$
Democratic & Towards Optimistic	$w_i = 2i/(N(N+1))$
Democratic & Towards Pessimistic	$w_i = 2(N+1-i)/(N(N+1))$
Semi Democratic & Towards Optimistic	For $N = 2k \quad w_i = 2/N \quad i > N/2$ For $N = 2k + 1 \quad w_i = 2/(N+1) \quad i > N/2$
Semi Democratic & Towards Pessimistic	For $N = 2k \quad w_i = 2/N \quad i \le N/2$ For $N = 2k + 1 \quad w_i = 2/(N+1) \quad i \le N/2 + 1$

In Table 1 we can ideally imagine a triangle whose vertices connect the three extreme decision attitudes *Monarchical & Pessimistic*, *Monarchical and Optimistic* and *Democratic & Neutral*.

We can notice that a fully *pessimistic/optimistic* decision attitude is also *monarchical*, since it must select the worst/best scenario among those yielded by all models: *the rule of the most pessimistic/optimistic model*. In this case one can choose just one single OWA operator as shown in Table 2.

A *neutral* attitude must balance pessimism and optimism. Nevertheless, if one wants to be neutral one can have several choices of OWA operators with different blends of monarchical/democratic attitude: then one can refine the possible choices by deciding if he/she wants to be monarchical or democratic or in between.

Democratic and neutral is modelled by the arithmetic average: "*one head one vote*".

On the other extreme, *Monarchical and neutral* can be modelled in the case of odd number of arguments by the median: "*the right mean rule*".

Semi monarchical and neutral can be modelled by the Hurwicz's criterion equally combining the most pessimistic and the most optimistic models. On the other side, *Semi democratic and neutral* is modelled by the arithmetic average truncated at the extremes, thus excluding the most pessimistic and optimistic models.

Semi Monarchical/Democratic & Neutral combines the previous two attitudes *Semi monarchical and neutral* and *Semi democratic and neutral*.

Monarchical & Towards Optimistic can be modelled by an OWA operator considering a single contribution of any model that provides a result smaller than that of the median and greater than that of the most optimistic one, *i.e.*, greater than the minimum. The degree of optimism is greater as the considered contribution is closer to the most optimistic model.

Monarchical & Towards Pessimistic can be modelled by a dual OWA operator considering a single contribution of any model that provides a result greater than the result of the median and smaller than the result of the most pessimistic one, *i.e.*, smaller than the maximum. The degree of pessimism is greater as the considered contribution is closer to the most pessimistic model.

Semi Monarchical & Towards Optimistic and *Semi Monarchical & Towards Pessimistic* can be modelled by a weighted Hurwicz's criterion combining the most optimistic (pessimistic) and the most pessimistic (optimistic) models with complementary weights so as to weigh more the most optimistic (pessimistic) model. Alternatively one can consider the first two most optimistic (pessimistic) models.

Democratic & Towards Optimistic and *Democratic & Towards Pessimistic* can be modelled by an OWA that performs the weighted average of all its arguments by associating a greater weight as the argument is more optimistic (pessimistic).

Finally, *Semi Democratic & towards Optimistic* and *Semi Democratic & Towards Pessimistic* can be modelled by an OWA that performs the weighted average of half of its most optimistic (pessimistic) arguments.

2) *A scalable approach*

The ESI computation described in the previous section can be implemented in a distributed processing framework represented by the schema depicted in **Figure 1**.

The ESI computation is performed independently for each spatial unit and is organized in two subsequent phases. Thus we can implement it in a single round of a map-reduce framework [3]. Note that spatial units can be either pixels in an image, thus performing a pixel-based mapping, or spatial objects defined by a closed boundary, thus performing an object-based mapping.

The map-reduce framework is inspired by the "map" and "reduce" functions used in functional programming. Computational processing occurs on data stored in a distributed file system or within a database, which takes a set of input key-values pairs and produces a set of output key-values pairs.

A mapper M is a Turing machine $M(\langle k,v\rangle)\to(\langle k1',v1'\rangle,\cdots,\langle ks',vs'\rangle)$, which accepts as input a single key-value pair $\langle k,v\rangle$ and produces a list of key-value pairs $\langle k1',v1'\rangle,\cdots,\langle ks',vs'\rangle$.

A shuffle is performed on the outputs of the mappers so as to group the values with the same key:

$$\langle k1',v1',\cdots,vr1'\rangle,\cdots,\langle kR',v1',\cdots,vrR'\rangle.$$

A reducer R is a Turing machine R: $\langle k',v1',\cdots,vr'\rangle\to\langle k',v''\rangle$, which accepts as input a pair $\langle k',v1',\cdots,vr'\rangle$ and produces as output the same key k' and a new value v''.
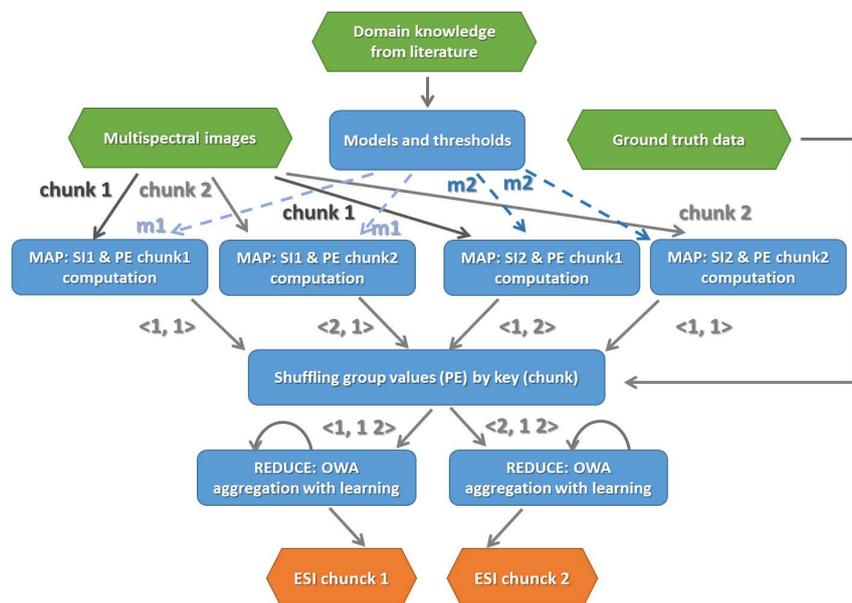
Figure 1. Distributed process computation of the ESI map.

A mapper can be instructed by its input parameters to compute results from more models, hereafter named contributing factors, and to evaluate more constraints, *i.e.*, thresholds, on the same chunk. This consists in evaluating selection conditions defined by specifying thresholds on the domains of the results of the models so as to segment binary partial evidence maps. Such constraints can be also "soft" defining in this way fuzzy thresholds, in the case of imprecision and uncertainty on the exact segmentation rules. They are defined by membership functions on the domains of the contributing factors so as to compute real values in [0, 1], thus yielding gradual partial evidence maps. The input key *k* identifies either a single pixel or a spatial unit in a multispectral image chunk. The associated value *v* is the information associated with the input chunk (e.g., the bands of the image chunk, the values of the theme such as GTD), plus parameters and model to apply (the names of contributing factors, the algorithm or function of the model that the mapper has to compute) and finally the tuples (*a, b, c, d, e, f*) defining the membership functions of the soft constraints μ_C according to the following definition:

$$\mu_C(x) = \begin{cases} 0 & x < a, x > d \\ \left((x-a)/(b-a)\right)^e & a \leq x < b \\ 1 & b \leq x \leq c \\ \left((d-x)/(d-c)\right)^f & c < x \leq d \end{cases} \tag{4}$$

The definition of precise thresholds to segment a grey level image resulting from the application of a model can be defined as special case of soft constraints by setting $a = b = -\infty$ and $c = d$ or $a = b$ and $c = d = +\infty$ in the Formula (4). This way an image segmentation is performed highlighting pixels belonging/not belonging to the phenomenon under study. Such segmented images are considered as partial evidence maps yielded by the distinct models.

For each pixel in the input chunk a mapper can compute the key-value pairs $\langle k1', v1' \rangle, \cdots, \langle ks', vs' \rangle$, where ki' identifies the chunk and vi' are the computed degrees of partial evidence in the chunk.

Successively, the reducers execute the second phase by aggregating the partial evidence maps $v1', \cdots, vr1'$ of the same chunk ki' in parallel so as to compute the ESI map v'' for the chunk.

Chunks are finally recombined by mosaicking at the end of the process.

The values v'' are computed by applying in each pixel or spatial unit of the chunk the OWA operator learned by leveraging GTD in the area covered by the chunk. This way, each reducer can learn a distinct OWA operator in each chunk thus, adapting the ESI computation to the local context and GTD. Notice that the learning process is performed within each reducer module, which applies on its input chunk the OWA operator learned at time epoch *L* based on the subset of ground truth included in the input chunk. There is no need to upload the input at each epoch, since the evidence maps do not change from epoch to epoch.

Once the optimal OWA has been determined, *i.e.*, the OWA weights do not substantially vary by increasing the epoch, the ESI map can be computed and stored on disk. The number of epochs and learning rates can be set so as to take into account the accuracy and the efficiency of this training phase in relation with the needs of the application. When a near-real time processing is desired the number of epochs can be reduced by tolerating a less accurate result.

The learning mechanism takes in input the observations $a_1, \cdots, a_K$ assumed as ground truth and their geographic coordinates, and associates with each one the partial evidence values $[a_{i1}, \cdots, a_{iN}]$ having the same coordinates. This way the following antecedent-consequent rules must be satisfied:

$$\begin{aligned} a_{11}, \cdots, a_{N1} &\to a_1 \\ a_{1K}, \cdots, a_{NK} &\to a_K \end{aligned} \tag{5}$$

In principle $a_1, \cdots, a_K$ can be specified on a continuous scale [0, 1] to quantify the extent of the phenomenon in the specific location; nevertheless, in practical situations, a discrete scale such as {0, 0.5, 1}, or even a binary scale {0, 1}, is used where 0 means absence of the phenomenon and 1 represents presence of it.

The learning mechanism starts at epoch $L = 0$ by assuming the weighted average (balanced and neutral attitude) as initial OWA_0 operator, which is defined with weighting vector $W_0 = [1/N, \cdots, 1/N]$. Then, at each epoch L, it iteratively determines the weighting vector $W_L = [w_{1L}, \cdots, w_{NL}]$ of OWA_L that minimizes the error existing between the results of its application to all the antecedents of the rules in (5) and the GTD *i.e.*, the consequents of the rules.

Formally, this is equivalent to applying the following rule:

$$\text{Select } W_L \text{ such that } |\Lambda_i(L) - \Lambda_i(L+1)| < \varepsilon \approx 0 \text{ or } L = L_{\max} \tag{6}$$

where

$$\begin{aligned} \Lambda_i(L+1) = \Lambda_i(L) - \beta w_{iL} \left(\arg\max_i(a_{1k}, \cdots, a_{Nk}) - OWA_L(a_{1k}, \cdots, a_{Nk}) \right) \\ * \left(OWA_L(a_{1k}, \cdots, a_{Nk}) - a_k \right) \end{aligned} \tag{7}$$

in which $\beta \in (0,1]$ is a learning rate parameter and the *ith* weighting vector element at epoch L is defined as follows:

$$w_{iL} = e^{\Lambda_i(L)} \Big/ \sum_{j=1,\cdots,N} e^{\Lambda_j(L)} \quad \forall i = 1, \cdots, N \tag{8}$$

4. Case Study: Mapping Standing Water Areas

1) *Data and Sources*

The case study is located in a territory in Northern Italy and is relative to monitoring standing water, which can occur due to controlled inundations (irrigation), extreme event floods and natural water reservoirs. Specifically, the three sites shown in **Figure 2** were selected as ROIs to cover different conditions of standing water in order to capture variable spectral characteristics: flooded area due to extreme heavy rainfall (ROI_1 Emilia area), river bed (ROI_2 Po Valley) and flooded rice fields (ROI_3 rice paddies) (see **Table 3**).
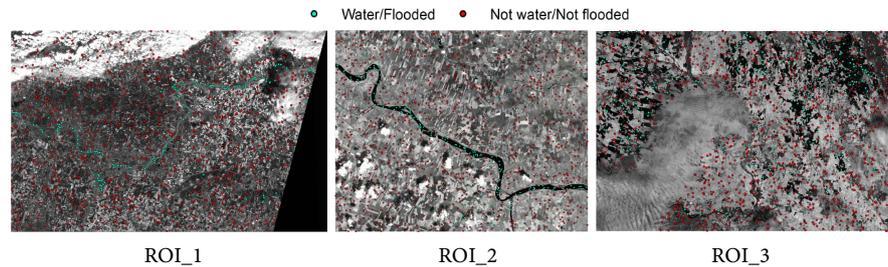
Figure 2. Study sites in Northern Italy with the ground truth points (blue: water and red: non-water) used for: 1) learning ordered weight averaging (OWA) operators and 2) validation of the algorithm. ROI: region of interest.

Table 3. Location/extent of the study sites and characteristics/conditions of the surface water areas. ROI: region of interest.

Site Name	Location (Italy)	Latitude (North)	Longitude (West)	Surface Water Conditions	Dimension (km²)
ROI_1	Emilia (IT)	44.968861	10.649674	Flooded areas due to severe rainfall events	2090
ROI_2	Po Valley (IT)	44.992491	11.377019	River in standard conditions	546
ROI_3	Rice paddies (IT)	45.278927	8.527552	Flooded rice fields	1937

The latter site was selected, although flooding was not due to a natural event, to train and validate the algorithm over heterogeneous conditions of a shallow water surface (<50 cm) mixed with soil patches and vegetation (emerging rice plants).

In the study, several information sources were used. Ground truth is either created by agronomists *in situ* by the use of a mobile application (in ROI_3) or by trained photointerpreters using a Geographic Information System (GIS) in which distinct layers of information are displayed to help them (in ROI_1 and ROI_2). The available ground truth set in each ROI was partitioned into two subsets and used as a training set for learning the OWA operator and as a reference set for validation.

The remote sensing data source used in all sites is Sentinel 2 (S2) (https://earth.esa.int/web/sentinel/home). The S2 mission operates as part of a two-satellite system (A and B) providing high-resolution multispectral optical imagery since June 2015 (A) and March 2017 (B). The S2 multispectral instrument (MSI) measures the Earth's reflected radiance in 13 spectral bands from VIS/NIR to SWIR with a spatial resolution ranging from 10 m to 60 m. Level-2A S2 images were downloaded and preprocessed with a sen2r toolbox [38]. The details of the preprocessing operations are described in [39]. For ROI_1 and ROI_2, Level-2A S2 imagery was downloaded as the bottom of atmosphere (BOA) reflectance through the Copernicus Open Access Hub. Preprocessing consisted of clipping images to our area of interest and masking clouds using the scene classification (SC) product; pixels classified as high and medium cloud probability were masked out, while pixels belonging to different classes were retained to

avoid masking-out water pixels. For ROI_3, a BOA image was not available at the desired dates of the field survey in the Copernicus archive, so it was necessary to download the top of atmosphere Level-1C products and apply atmospheric correction by using the Sen2Cor algorithm of the sen2r toolbox library [38].

Table 4 reports for each site's EO satellite data and acquisition dates the number of ground truth pixels (w/nw stand for water/not water) used for learning the OWA operator (L) in phase two of the algorithm and for validation (V) of the computed ESI maps. At each validation epoch, 10% (90%) of the ground truth pixels were randomly selected for (L), and the remaining 90% (10%) were used for (V) in the typical (atypical) validation settings, respectively as explained in next sections.

2) *The input models for mapping standing water*

For mapping standing water from optical remote sensing images several models have been defined in the literature. In the case study we identified and selected eight different models: seven of which propose to compute a spectral index (SI) enhancing a specific characteristic shape of the spectral signature of standing water. These models also define thresholds on the SI domain to segment standing water. This way we can compute binary partial evidence maps of standing water. Besides, we selected a last model defined in [22] that maps the optical image into hue (*H*) and value (*V*) dimensions of the *HSV* color space, derived by transforming the components *SWIR*2, *NIR* and *RED*. In this case standing water surfaces can be separated from land surfaces by means of two empirical thresholds on *V* and *H* domains as defined in [22].

The transformation function *f*: *SWIR*2 × *NIR* × *RED* → *H* × *V* is a standardized colorimetric transformation from *RGB* to *HV* components of the *HSV* color space, where *SWIR*2 = *R*, *NIR* = *G* and *RED* = *B* respectively, defined as in [40]:

$$
f(R,G,B)
\begin{cases}
V = \max(R,G,B) \\[2mm]
H = \begin{cases}
\left(\left(60° * \dfrac{G-B}{V-\min(R,G,B)}\right) + 360°\right) \bmod 360° & \text{if } V = R \\[4mm]
\left(\left(60° * \dfrac{B-R}{V-\min(R,G,B)}\right) + 120°\right) & \text{if } V = G \\[4mm]
\left(\left(60° * \dfrac{R-G}{V-\min(R,G,B)}\right) + 240°\right) & \text{if } V = B
\end{cases}
\end{cases}
\tag{9}
$$

Table 4. Number of pixels (w/nw stand for water/not water) for ROI used for learning the weights of the OWA operator (L) in the 10-fold cross typical/atypical validations (V).

Name	Dates	# ground truth pixels (w/nw) for (L)	# ground truth pixels (w/nw) for (V)
ROI_1	S2A 2017-12-13	87 - (16/71)	779 - (141/638)
ROI_2	S2A 2017-05-14	113 - (19/94)	1018 - (173/845)
ROI_3	S2A 2016-04-22	102 - (17/85)	921 - (153/768)

All crisp thresholds suggested in the literature for each SI and the two thresholds on HV were used as parameters to define the crisp constraints based on Formula (4). All the models used in the case study are reported in Table 5 [19] [20] [21] [22] [39].

3) *Results of the synthesis of the models*

The validation experiment was designed with the following objectives:

a) to compare the accuracy of the proposal with respect to the results yielded by the single models;

b) to investigate the robustness, *i.e.*, stability of results with respect to changing the ROI;

c) to investigate the performance when downscaling the training set.

In phase 1 of the algorithm binary partial evidence (PE) maps are computed by applying each model using as input the preprocessed multispectral images, and the crisp thresholds suggested in the literature and reported in Table 5. The PE maps are successively used by phase 2 together with GTD subsets to the aim of learning the OWA operator and then computing the overall ESI map. This consists in aggregating PE maps by applying the OWA operator learned by the iterative process exploiting GTD. Outputs of this phase are: the ESI map, the weighting vector of the OWA operator, its *orness* and *dispersion* measures and the correspondent decision attitude (as defined in Table 1) modeled by the OWA operator. Note that, while the PE maps are binary with pixel values defined in {0, 1} in which 1 represents standing water evidence and 0 no evidence, the ESI map computed by aggregating them have pixel values defined in [0, 1], in which 1 means maximum evidence of standing water while 0 no evidence at all, and intermediate values represent gradual evidence.

Table 5. Selected Models and the thresholds suggested by the literature to segment flooded/not flooded areas.

Model	Formula	Category	Crisp Threshold to segment standing water	Reference
AWEI	C1 * (GREEN − SWIR1) − (C2 * NIR + C3 * SWIR2)	Water SI	AWEI > 0	[19]
AWEIsh	BLUE + D1 * GREEN − D2 * (NIR + SWIR1) − D3 * SWIR2	Water SI	AWEIsh > 0	[19]
mNDWI	(GREEN − SWIR1)/(GREEN + SWIR1)	Water SI	mNDWI > 0	[23]
NDWI	(GREEN − NIR)/(GREEN + NIR)	Water SI	NDWI > 0	[21]
NDFI	(RED − SWIR2)/(RED + SWIR2)	Flooding SI	NDFI > 0.32	[18]
SAVI	(1 + L) * (NIR − RED)/(NIR + RED + L)	Vegetation SI	SAVI < −0.25 or SAVI < −0.27 (in the present paper we used −0.25)	[20]
WRI	(GREEN + RED)/(NIR + SWIR2)	Water SI	WRI > 1	[41]
HV	*f*(SWIR2, NIR, RED)	Water indicator	$H < (−0.3593 * V) + 360$ or $H > (0.0363 * V) + 75.853$ or $H > (0.054 * V) + 150.04$ and $V < 5500$	[22]

Where C1 = 4; C2 = 0.25; C3 = 2.75; D1 = 2.5; D2 = 1.5; D3 = 0.25; L = 0.5 and S2 MSI bands are: BLUE = band 2 (490 nm), GREEN = band 3 (560 nm), RED = band 4 (665 nm), NIR = band 8 (842 nm), SWIR1 = band 11 (1610 nm) and SWIR2 = band 12 (2190 nm) as defined in the cited literature.

The algorithm phase 1 was executed on the three ROIs. Phase 2 was executed several times with distinct GTD subsets, as described in the following subsection; in facts by changing GTD, different ESI maps can be computed for the same ROI.

Evaluation of performance of each single model (contributing factor) in mapping standing water was accomplished by a site-specific assessment adopting the error matrix approach [42]. From the error matrix, where columns represent reference classes, while rows represent the model predictions, and true positives (*TP*), true negatives (*TN*), false positives (*FP*) and false negatives (*FN*), the following accuracy metrics were derived: *Precision* (User Accuracy) and *Recall* (Producer Accuracy) and *F-score*. Precision represents User Accuracy or map user's viewpoint and is determined as:

$$Precision = 1 - CE = 1 - FP/(FP + TP)$$

where *CE* is Commission Error.

Recall represents Producer Accuracy or the map producer's point of view and is defined as:

$$Recall = 1 - OE = 1 - FN/(FN + TP)$$

where *OE* is Omission Error.

F-score is a synthetic measure defined as the harmonic mean:

$$F\text{-}score = 2*(Prec * Recall)/(Prec + Recall) = TP/(2TP + FN + FP) \quad (10)$$

Table 6 reports the omission error *OE*, commission error *CO* and *F-score* measure in the three ROIs (shown in **Figure 2**) obtained by applying the single models with the thresholds specified in **Table 5** generating distinct SI maps. Ground truth for the validation is composed of around 1000 GTD independent elements in each ROI, as reported in the fourth column of **Table 4**. Pixels with values exceeding the threshold are considered as "standing water".

It can be observed that in the three ROIs which are characterized by distinct land covers and water conditions (water depth, color, fractional cover, plant/soil patches presence, etc.), a different model presents the best performance (greatest *F-score*). This confirms our intuition that a single model cannot capture all types of standing water conditions.

In the Emilia area, AWEIsh, mNDWI, NDFI and WRI have the best comparable performance; in the Po Valley area and Rice paddies site the HV has the best performance.

Furthermore we can observe that some of the models generate omission errors in all the three sites, like SAVI and secondly NDWI, while the other models do not behave the same in the three sites. Thus by choosing a single model one cannot be sure of the decision attitude it applies since this depends on the characteristics of the area. By averaging the results on all the three sites we can see that the best average models are mNDWI and WRI, both characterized by a greater average omission error than average commission error. So by choosing them we do not have flexibility to model decision attitudes that are risk-adverse.

Table 6. Performance of the single models over the three validation sites.

		AWEI	AWEIsh	mNDWI	NDFI	NDWI	SAVI	WRI	HV
Emilia area	*Omission*	0.01	0.00	0.01	0.01	0.74	0.82	0.00	0.01
	Commission	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17
	F-score	0.99	**1.00**	**1.00**	**1.00**	0.41	0.30	**1.00**	0.90
Po valley	*Omission*	0.17	0.06	0.06	0.23	0.05	1.00	0.04	0.01
	Commission	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	F-score	0.91	0.97	0.97	0.87	0.98	0.00	0.98	**0.99**
Rice paddies	*Omission*	0.27	0.23	0.10	0.23	0.62	1.00	0.11	0.02
	Commission	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	F-score	0.84	0.87	0.95	0.87	0.55	0.00	0.94	**0.99**
Average on the three sites	*Omission*	0.15	0.10	0.06	0.16	0.47	0.94	0.05	0.01
	Commission	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06
	F-score	0.92	0.95	**0.97**	0.91	0.64	0.10	**0.97**	0.96

This motivates our approach that allows either to learn the best aggregation in each site or to choose a desired decision attitude for computing synthetic ESI maps.

In order to pursue the validation of the approach when applying the learning mechanism of the aggregation with a traditional setting of the training set and with a small training set, we designed two kinds of k-fold cross-validation experiments.

We recall that a k-fold cross validation is a statistical method aimed at evaluating the performance of a learning algorithm by changing the training set; in doing so, it is possible to compute both average performance metrics and the standard deviation to assess its sensitivity.

In each experiment, phase 2 was executed 10 times ($k = 10$), thus generating 10 weighting vectors of the OWA and, consequently, 10 distinct ESI maps for the site. At each run, a different subset of both GTD for learning and testing were selected by applying stratified random sampling. In the first kind of validation experiments, we used 90% of GTD elements for learning the OWA aggregation and 10% for testing, as in the standard validation methods of machine-learning. These experiments are named typical (T) k-fold cross validations.

To test the algorithm with a small training set, we performed the 10-fold cross validations with a different proportion of the learning and testing sets. Differently than in the typical validations, this time we used only 10% of the available ground truth pixels for training, while we used the remaining 90% for testing. Stratified random sampling was applied to select the two subsets. This validation is called atypical (AT) and was aimed at investigating the stability of the results when simulating a realistic situation with a small set of GTD.

Performances achieved on each ROI by the typical and atypical 10-fold cross validations are shown in **Figure 3**; the 10 *F-score* diagrams in each area are relative to the 10 ESI maps produced as a result of the 10 executions of the algorithm phase 2. The *F-score* values of the diagrams in correspondence with a threshold on the ESI are computed by considering as "watered" the values of pixels above the threshold. By increasing the threshold we become more strict on the definition of "watered" area.

Table 7 summarizes average performances of the algorithm over all runs and all thresholds in both the typical and the atypical validations in each ROI. **Table 8** reports the learned OWA operator, averaged over the 10 runs in both the typical (T) and atypical (AT) validation settings.

Finally, **Figure 4** illustrates for each ROI the ESI maps obtained by applying OWA operators modeling distinct decision attitudes.
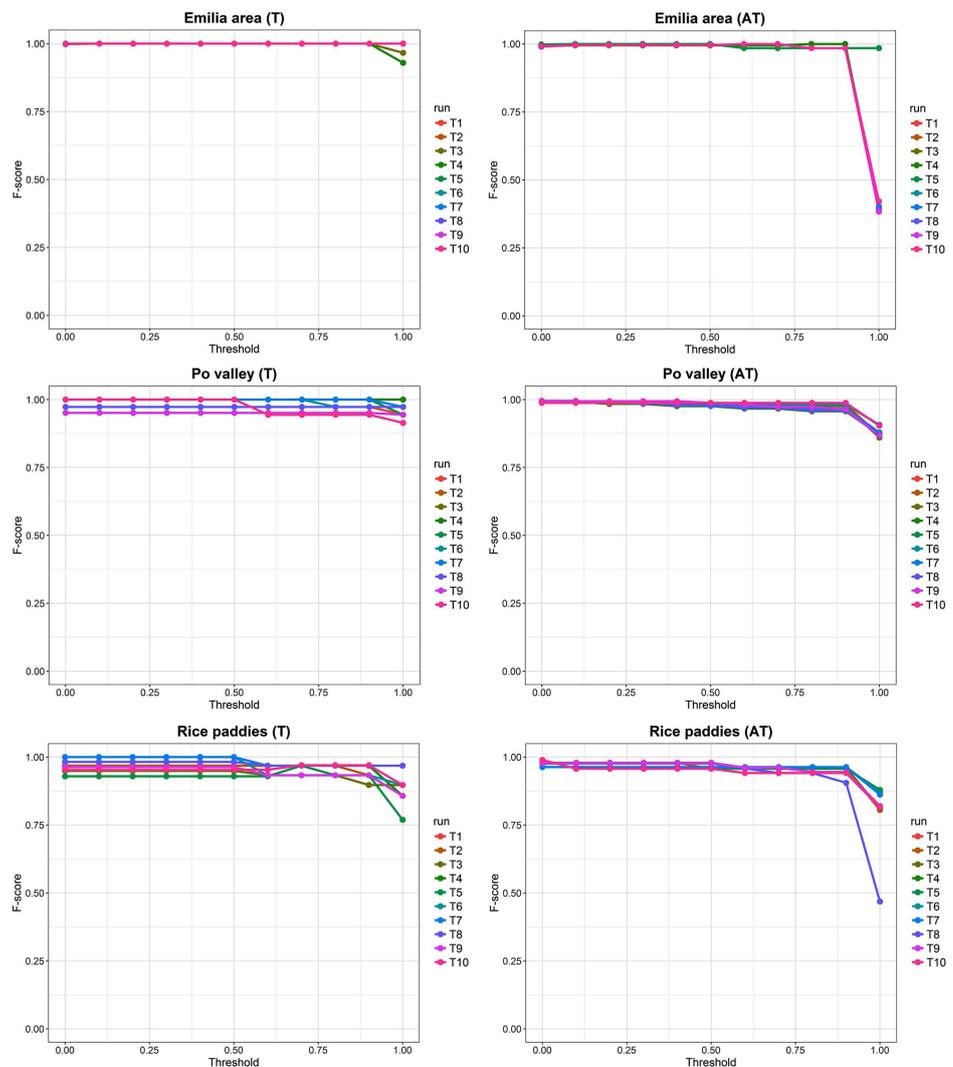
Figure 3. *F-score* diagrams on the three ROIs in the typical (T) and atypical (AT) 10-fold cross validations. Parameters used for *k*-fold cross validations: *k* = 10, learning rate = 0.5 and number of epochs = 500.

Table 7. Average *OE*, *CE*, *F-score* and standard deviation values over all 10 runs of the algorithm and all thresholds in the typical (T) and atypical (AT) validations in each sites and average of the same measures over the three sites.

	Average over 10 runs	(T)	(AT)
	Omission	0.00	0.06
	Commission	0.00	0.01
Emilia area	F-score	**1.00**	0.95
	Std Dev	0.00	0.02
	Omission	0.02	0.05
	Commission	0.01	0.00
Po valley	F-score	**0.98**	0.97
	Std Dev	0.02	0.00
	Omission	0.06	0.07
	Commission	0.02	0.03
Rice paddies	F-score	**0.96**	0.95
	Std Dev	0.02	0.01
	Omission	0.03	0.06
	Commission	0.01	0.01
Average on the three sites	F-score	**0.98**	0.96
	Std Dev	0.02	0.02

Table 8. Learned weighting vectors of the OWA operator in each ROI averaged over the 10 runs of both the typical (T) and atypical (AT) 10-fold cross validations. The table also reports the values of the approximated weighting vector, the average ORness (Φ), the dispersion (Δ) and correspondent decision attitude.

	w1	w2	w3	w4	w5	w6	w7	w8	Φ	Δ	Decision attitude
Emilia area (T)	0.001	0.008	0.008	0.008	0.964	0.009	0.001	0.001	0.43	0.04	*Semi Monarchical & Towards Optimistic*
Emilia area (AT)	0.010	0.171	0.171	0.171	0.263	0.199	0.007	0.007	0.55	0.74	*Semi Democratic & Towards Pessimistic*
Po valley (T)	0.734	0.041	0.200	0.011	0.007	0.003	0.003	0.001	0.92	0.27	*Semi Monarchical & Towards Pessimistic*
Po valley (AT)	0.389	0.248	0.167	0.107	0.063	0.012	0.010	0.005	0.81	0.61	*Semi Democratic & Towards Pessimistic*
Rice paddies (T)	0.563	0.272	0.148	0.007	0.004	0.003	0.002	0.001	0.91	0.44	*Semi Monarchical/ Democratic & Towards Pessimistic*
Rice paddies (AT)	0.589	0.276	0.091	0.022	0.008	0.006	0.004	0.003	0.91	0.41	*Semi Monarchical & Towards Pessimistic*
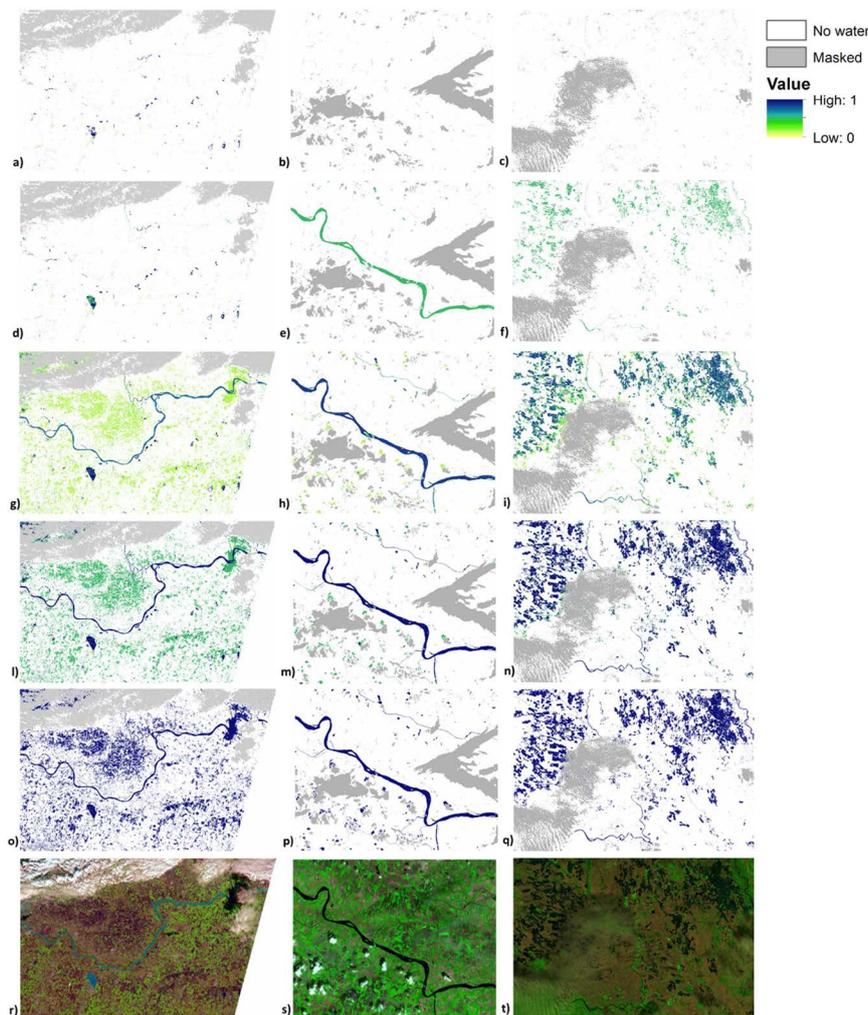
Figure 4. The global evidence degree over Emilia area (left), Po valley (middle) and Rice paddies area (right) obtained with OWA with *Monarchical & Optimistic* attitude (a, b, c), *Semi Democratic toward Optimistic* attitude (d, e, f), *Democratic and neutral* attitude (g, h, i), *Semi Democratic toward Pessimistic* attitude (l, m, n) and *Monarchical & Pessimistic* attitude (o, p, q). In the bottom row (r, s, t) S2 images as false colour composites (R = SWIR/band 12, G = NIR/band 8, B = RED/band 4).

4) *Comparison with results of single models*

As shown in Table 7 the average *F-score* on the three sites of our proposal in the typical validation are superior with respect to the average results of all the single models shown in Table 6 (*F-score* equal to 0.98 with respect to 0.97 of the average best models).

In the atypical validation our proposal yields an average *F-score* of 0.96 that is still superior or equal with respect to six of the 8 single models.

5) *Stability of the Results by Changing ROI*

Figure 3 shows that in all ROIs and in the typical validation setting, *F-score* diagrams are quite stable and maintain high values for all thresholds.

In the atypical setting, the *F-scores* for some of the 10 runs decrease for high thresholds (above 0.9), especially in the Emilia ROI. This may depend on the small

dimensions of the learning set used, which for some runs may not represent well all types of surface water.

This stability of the performance when changing the threshold on the ESI is an advantage with respect to the single models which conversely need an accurate calibration of the thresholds to yield accurate results.

6) *Synthesising models based on a decision attitude*

Our last experiment has been to compute the ESI map by applying an OWA operator modeling a desired decision attitude. Figure 4 shows the ESI maps over the three ROIs obtained with OWA operators associated with distinct decision attitudes. It can be noticed that the ESI maps (panels a, b, c) computed by the *Monarchical & Optimistic* attitude exhibit the best and less alarming scenarios; thus the decision maker takes the risks of underestimating the critical situation since the extent of flooded areas is at the minimum in the three sites. Conversely, the ESI maps (panels o, p, q) computed by the *Monarchical & Pessimistic* attitude exhibit the worst scenarios since the extent of the flooded areas is maximum in the three sites. These results are coherent with the attitudes of the two OWA operators used to compute them which model prioritization of interventions by minimizing false alarms and risk aversion by not underestimating any hint of floods, respectively. The *Democratic & Neutral* synthesis (panels g, h, i) is the decision attitude right in the middle that balances pessimism and optimism democratically, by considering all models equally. The *Semi Democratic toward Optimistic* attitude (panels d, e, f) and the *Semi Democratic toward Pessimistic* attitude (panels l, m, n) have been defined with OWA operators that equally weigh the contributions of the two most optimistic (pessimistic) models respectively. It can be observed that the best visual results for Po valley and Rice paddies are those in panels (m, n) obtained with a *Semi Democratic toward Pessimistic attitude*, that is characterized by an OWA operator with dispersion = 0.5 and orness = 0.93. In facts, in Po valley and Rice paddies, we obtain average OWA operators, considering typical and atypical validations, with average orness equal to 0.87 and 0.91 which both correspond to *towards pessimistic* and average dispersion equal to 0.44 and 0.42 which correspond to *Semi Monarchical/Democratic* and *Semi monarchical* respectively, as it can be easily verified by looking at Table 1. In the case of Emilia area the best visual result is the one in panel (g) corresponding with *Democratic & Neutral*: in facts the average learned OWA operator has an average orness of 0.49, even if the average dispersion is 0.39.

5. Conclusions

The literature on remote sensing image analysis and interpretation to map the status of a specific environmental problem is extensive and many models have been defined for distinct phenomena. Such models when applied in a different (same) area using the same (different) data sources can yield inconsistent results. Flexible methods, which are capable to integrate the outputs of the models in a flexible way depending on the context, are needed. In fact in some areas the phe-

nomenon can have multiple aspects and thus a complementarity of models should be exploited to generate the synthesis. In other situations, a synergic aggregation of the models is more suitable to capture their consensual evaluations of the phenomenon.

In this paper we applied a method defined within soft computing in [10] to exploit both ill-defined knowledge of remote sensing experts (or products affected by uncertainty because generated by models calibrated in a different context) and *in situ* observations to yield ESI maps.

The method is here used to synthesize the results of distinct models which compute binary images to map the status of phenomena from the analysis of multispectral remotely sensed images. The described case study had the objective of mapping standing water areas as an indicator of flooding condition in natural hazard management scenario.

The results show that the proposal is characterized by two levels of adaptation: it can learn a distinct aggregation operator depending on the environmental conditions of a given area and it can select the best models in each spatial unit to determine the final result. It has been shown in the case study that the approach can produce better results than any single model used. Finally, it is described how the semantics of the aggregation can be interpreted as modeling a given decision attitude towards facing risks. This confers human interpretability to the aggregation that can thus be specified linguistically by a decision maker in order to generate ESI maps that minimize either False Positives or False Negatives depending on the decision objectives.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Torlone, R. (2020) Hadoop & Map-Reduce.
http://torlone.dia.uniroma3.it/bigdata/L2-Hadoop.pdf

[2] Dean, J. and Ghemawat, S. (2008) Mapreduce: Simplified Data Processing on Large Clusters. *Communications of the ACM*, **51**, 107-113.
https://doi.org/10.1145/1327452.1327492

[3] Karloff, H., Suri, S. and Vassilvitskii, S. (2010) A Model of Computation for Map Reduce. *Proceedings of the* 21*st Annual ACM-SIAM Symposium on Discrete Algorithms*, Austin, 17-19 January 2010, 938-948
https://doi.org/10.1137/1.9781611973075.76

[4] Atzeni, P., Jensen, C.S., Orsi, G., Ram, S., Tanca, L. and Torlone, R. (2013) The Relational Model Is Dead, SQL Is Dead, and I Don't Feel So Good Myself. *SIGMOD Record*, **42**, 64-68. https://doi.org/10.1145/2503792.2503808

[5] OGC Standards (2020). https://www.ogc.org/docs/is

[6] Sterlacchini, S., Bordogna, G., Cappellini, G. and Voltolina, D. (2018) SIRENE: A Spatial Data Infrastructure to Enhance Communities' Resilience to Disaster-Related

Emergency. *International Journal of Disaster Risk Science*, **9**, 129-142.
https://doi.org/10.1007/s13753-018-0160-2

[7] Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V. (2017) Critical Analysis of Big Data Challenges and Analytical Methods. *Journal of Business Research*, **70**, 263-286. https://doi.org/10.1016/j.jbusres.2016.08.001

[8] Yager, R.R. (1988) On Ordered Weighted Averaging Aggregation Operators in Multi-Criteria Decision Making. *IEEE Transactions on Systems*, *Man, and Cybernetics*, **18**, 183-190. https://doi.org/10.1109/21.87068

[9] Filev, D.P. and Yager, R.R. (1998) On the Issue of Obtaining OWA Operator Weights. *Fuzzy Sets and Systems*, **94**, 157-169.
https://doi.org/10.1016/S0165-0114(96)00254-0

[10] Goffi, A., Bordogna, G., Stroppiana, D., Boschetti, M. and Brivio, P.A. (2020) Knowledge and Data-Driven Mapping of Environmental Status Indicators from Remote Sensing and VGI. *Remote Sensing*, **12**, 495. https://doi.org/10.3390/rs12030495

[11] Lecun, Y., Bengio, Y. and Hinton, G.E. (2015) Deep Learning. *Nature*, **521**, 436-444.
https://doi.org/10.1038/nature14539

[12] Reichstein, M., Camps-Valls, G., Stevens, B., *et al.* (2019) Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature*, **566**, 195-204.
https://doi.org/10.1038/s41586-019-0912-1

[13] Zhang, L., Zhang, L. and Du, B. (2016) Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, **4**, 22-40. https://doi.org/10.1109/MGRS.2016.2540798

[14] Castelluccio, M., Poggi, G., Sansone, C. and Verdoliva, L. (2015) Land Use Classification in Remote Sensing Images by Convolutional Neural Networks.

[15] Larochelle, H., Bengio, Y., Louradour, J. and Lamblin, P. (2009) Exploring Strategies for Training Deep Neural Networks. *Journal of Machine Learning Research*, **10**, 1-40.

[16] Han, X., Zhong, Y., Cao, L. and Zhang, L. (2017) Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification. *Remote Sensing*, **9**, 848.
https://doi.org/10.3390/rs9080848

[17] Acharya, T.D., Subedi, A. and Lee, D.H. (2018) Evaluation of Water Indices for Surface Water Extraction in a Landsat 8 Scene of Nepal. *Sensors*, **18**, 2580.
https://doi.org/10.3390/s18082580

[18] Boschetti, M., Nutini, F., Manfron, G., Brivio, P.A. and Nelson, A. (2014) Comparative Analysis of Normalised Difference Spectral Indices Derived from MODIS for Detecting Surface Water in Flooded Rice Cropping Systems. *PLoS ONE*, **9**, e88741.
https://doi.org/10.1371/journal.pone.0088741

[19] Feyisa, G.L., Meilby, H., Fensholt, R. and Proud, S.R. (2014) Remote Sensing of Environment Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery. *Remote Sensing of Environment*, **140**, 23-35.
https://doi.org/10.1016/j.rse.2013.08.029

[20] Huete, A.R. (1988) A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, **25**, 295-309. https://doi.org/10.1016/0034-4257(88)90106-X

[21] McFeeters, S.K. (1996) The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *International Journal of Remote Sensing*, **17**, 1425-1432. https://doi.org/10.1080/01431169608948714

[22] Pekel, J.F., Vancutsem, C., Bastin, L., Clerici, M., Vanbogaert, E., Bartholomé, E. and Defourny, P. (2014) A Near Real-Time Water Surface Detection Method Based

on HSV Transformation of MODIS Multi-Spectral Time Series Data. *Remote Sensing of Environment*, **140**, 704-716. https://doi.org/10.1016/j.rse.2013.10.008

[23] Xu, H. (2006) Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, **27**, 3025-3033. https://doi.org/10.1080/01431160600589179

[24] Bordogna, G., Pagani, M. and Pasi, G. (2006) A Flexible Decision Support Approach to Model Ill-Defined Knowledge in GIS. *NATO Workshop on Environmental Impact Assessment*, Kiew, June 2006, 133-152. https://doi.org/10.1007/978-1-4020-6438-8_8

[25] Yager, R.R. (1998) New Modes of OWA Information Fusion. *International Journal of Intelligent Systems*, **13**, 661-681. https://doi.org/10.1002/(SICI)1098-111X(199807)13:7<661::AID-INT5>3.0.CO;2-I

[26] Bone, C., Dragicevic, S. and Roberts, A. (2005) Integrating High Resolution Remote Sensing, GIS and Fuzzy Set Theory for Identifying Susceptibility Areas of Forest Insect Infestations. *International Journal of Remote Sensing*, **26**, 4809-4828. https://doi.org/10.1080/01431160500239180

[27] Chanussot, J., Mauris, G. and Lambert, P. (1999) Fuzzy Fusion Techniques for Linear Features Detection in Multitemporal SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 1292-1305. https://doi.org/10.1109/36.763290

[28] Jiang, H. and Eastman, J.R. (2000) Application of Fuzzy Measures in Multi-Criteria Evaluation in GIS. *International Journal of Geographical Information Science*, **14**, 173-184. https://doi.org/10.1080/136588100240903

[29] Malczewski, J. (2006) GIS-Based Multicriteria Decision Analysis: A Survey of the Literature. *International Journal of Geographical Information Science*, **20**, 703-726. https://doi.org/10.1080/13658810600661508

[30] Robinson, P.B. (2003) A Perspective on the Fundamentals of Fuzzy Sets and Their Use in Geographic Information Systems. *Transactions in GIS*, **7**, 3-30. https://doi.org/10.1111/1467-9671.00127

[31] Brivio, P.A., Boschetti, M., Carrara, P., Stroppiana, D. and Bordogna, G. (2006) Fuzzy Integration of Satellite Data for Detecting Environmental Anomalies across Africa. In: Hill, J. and Roeder, A., Eds., *Recent Advances in Remote Sensing and Geoinformation Processing for Land Degradation Assessment*, Taylor & Francis, London, 147-160.

[32] Carrara, P., Bordogna, G., Boschetti, M., Brivio, P.A., Nelson, A.D. and Stroppiana, D. (2008) A Flexible Multi-Source Spatial-Data Fusion System for Environmental Status Assessment at Continental Scale. *International Journal of Geographical Information Science*, **22**, 781-799. https://doi.org/10.1080/13658810701703183

[33] Bordogna, G., Boschetti, M., Brivio, P.A., Carrara, P., Pagani, M. and Stroppiana, D. (2011) Fusion Strategies Based on the OWA Operator in Environmental Applications. In: Kacprzyk, J., Yager, R.R. and Beliakov, G., Eds., *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*, Sprinter Verlag, Berlin, Vol. 265, 298.

[34] Kacprzyk, J. (1988) Fuzzy Logic with Linguistic Quantifiers: A Tool for Better Modeling of Human Evidence Aggregation Processes? *Advances in Psychology*, **56**, 233-263. https://doi.org/10.1016/S0166-4115(08)60487-2

[35] Zadeh, L.A. (1983) A Computational Approach to Fuzzy Quantifiers in Natural Languages. *Computers & Mathematics with Applications*, **9**, 149-184. https://doi.org/10.1016/0898-1221(83)90013-5

[36] Yager, R.R. (1996) Quantifier Guided Aggregation Using OWA Operators. *Interna-

tional Journal of Intelligent Systems, **11**, 49-73.
https://doi.org/10.1002/(SICI)1098-111X(199601)11:1<49::AID-INT3>3.0.CO;2-Z

[37] Yager, R.R. (2009) On the Dispersion Measure of OWA Operators. *Information Sciences*, **179**, 3908-3919. https://doi.org/10.1016/j.ins.2009.07.015

[38] Ranghetti, L. and Busetto, L. (2019) Sen2r: An R Toolbox to Find, Download and Preprocess Sentinel-2 Data. R Package Version 1.0.0. http://sen2r.ranghetti.info

[39] Goffi, A., Stroppiana, D., Brivio, P.A., Bordogna, G. and Boschetti, M. (2020) Towards an Automated Approach to Map Flooded Areas from Sentinel-2 MSI Data and Soft Integration of Water Spectral Features. *International Journal of Applied Earth Observation and Geoinformation*, **84**, Article ID: 101951. https://doi.org/10.1016/j.jag.2019.101951

[40] Smith, A.R. (1978) Color Gamut Transform Pairs. *Proceedings of the 5th Annual Conference on Computer Graphics and Interactive Techniques*, New York, 23-25 August 1978, 3, 12-19. https://doi.org/10.1145/800248.807361

[41] Shen, L. and Li, C. (2010) Water Body Extraction from Landsat ETM + Imagery Using Adaboost Algorithm. *Proceedings of the 2010 18th International Conference on Geoinformatics*, Beijing, 18-20 June 2010, 1-4. https://doi.org/10.1109/GEOINFORMATICS.2010.5567762

[42] Congalton, R.G. and Green, K. (2008) Assessing the Accuracy of Remotely Sensed Data. 2nd Edition, CRC Press, Boca Raton, 346. https://doi.org/10.1201/9781420055139