Kalman-Kriging Technique Applied to Space-aided Distributed Sensor System to Manage Critical Environmental Events

[1]Gregorio Andria, [1]Eugenio Di Sciascio, [1]Anna M. L. Lanzolla, [2]A. Lay-Ekuakille, [1]Michele Ruta
[1]Department of Electrical & Information Engineering (DEI), Politecnico di Bari, Via E. Orabona, 4 - 70125 Bari, Italy
[gregorio.andria; eugenio.disciascio; anna.lanzolla; michele.ruta]@poliba.it
[2]Dipartimento di Ingegneria dell'Innovazione, University of Salento, Via Monteroni, 73100 Lecce, Italy
aime.lay.ekuakille@unisalento.it

***Abstract* — This work deals with a useful joined application of Kalman filter and Kriging technique for a continuous and accurate environmental monitoring in a disaster scenario or generally in a critical event, necessary to assure an efficient and timely risk management. A suitable Decision Support Systems (DSS) is proposed to provide assistance for the "early warning" of a critical situation, so to improve the "first response" to the happened event. Using the proposed modelling techniques in data environmental analysis permits both the characterization and validation of all measured big data coming from a suitable Space-Aided Distributed Sensor System (SADSS). In particular, the proposed technique is able also to predict the values of the monitored environmental parameters, so it results a very useful analysis tool, especially when there are many missing ,erroneous or invalid data.**

Keywords — Sensor Systems; Decision Support Systems; Kalman filtering; Kriging algorithm; Big data evaluation; disaster scenario management

I. INTRODUCTION

The situations of growing alarming relevant for hydro-geological distress or in general for all critical environmental events require an efficient and reliable prevention or even management. It is noteworthy the large availability of data, in any place and under any meteorological condition, which can support decision-makers in assuming rapidly actions for protecting citizens.

During these last years, many remote sensing techniques have been developed to monitor physical or environmental conditions (such as temperature, sound, vibration, pressure, motion or pollutants) in order to study climate changes and to assess hydrogeological disasters and other dangerous environmental events [1]-[7]. This has led several countries in the world to establish severe laws and regulations in the environmental field, so many public administrations have installed several monitoring systems able to give suitable and rapid information about the environmental quantities and parameters/indexes to be continuously controlled. In particular, a considerable importance is turned to quality evaluation of both hydrogeological and geostatic parameters.

Moreover, it is generally acknowledged the need of suitable Decision Support Systems (DSSs) providing assistance for the "early warning" of a disaster or critical situation and able to improve the "first response" to the happened event. These systems can be based on a suitable integration of satellite big data and *in situ* information [8], [9].

The global system we envision here provides:

- a Natural Resource and Emergency Management and Organization designated to answer to the first warning of a disaster and immediately after it;
- ICT-based methods and techniques to improve understanding of the natural disaster processes aimed at mitigating the damage;
- a Support Service able to provide mechanisms for early warning for crisis management.

The system design follows the objective of providing the right information in the *right* time to the *right* person to support the *right* decision. The information management process is critical: an end-to-end *Quality of Information* must be pursued; a continuous *Context Extraction* must be performed both in terms of: (i) the place where data are generated, and (ii) the place where the data are consumed. The system incorporates a swarm-oriented data distribution service which gathers environmental measures as sensor output, validates them after suitable data processing with the proposed suitable Kalman-Kriging technique[10], [11]], and the consequent metrological characterization operations, replaces the values of the monitored environmental parameters when they are missing or erroneous or invalid, and finally supports decision-makers in their actions for the safety and citizen protection.

II. PROPOSED MEASUREMENT SYSTEM ARCHITECTURE

In our study, we propose the analysis of measured environmental data by the distributed sensor system, recoded in several fixed monitoring stations installed in Taranto area (South Italy). All the sources of environmental problems contribute to compromise the life quality of people, by introducing smog or damaging quality parameters of soil and water.

978-1-4673-8291-5/16/$31.00 ©2016 IEEE

Fig. 1: Plant of Taranto area with the network of monitoring stations. The circle bounds the stations analysed in Kalman-Kriging technique

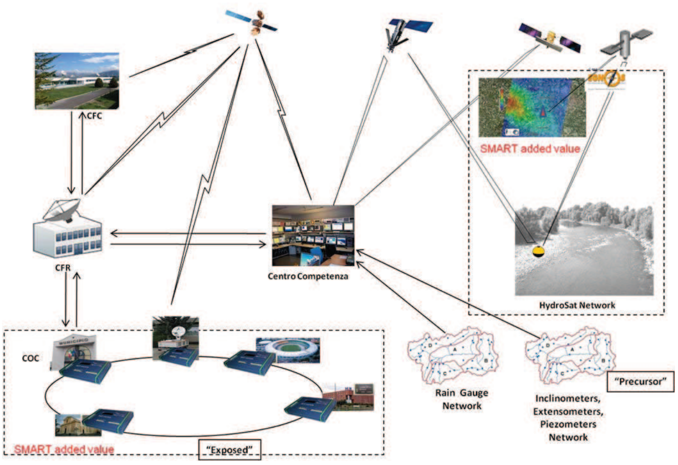

Fig.2: Integrated space-aided and *in situ* autonomous sensor system [11], whose measurements are processed in Kalman-Kriging technique

The *in situ* sensor network consists of a number of automatic acquisition and recoding stations able to measure both chemical substances and meteorological quantity. These fixed sensor stations are placed on different sites (shown in Fig.1), according to the national law prescriptions.

Other sensors are distributed in a swarm network system, in order to measure the quantities of interest and to allow a suitable management operation in case of critical events. Fig.2 shows an integrated space-aided and *in situ* autonomous sensor system, whose measurements can be easily processed via the proposed Kalman-Kriging technique. In particular, the proposed swarm system, named SMART (*Space for sMArt Resource managemenT for disaster early warning*) [12], aims at improving the quality of information to support (i) the plans addressing the management of disaster or critical event related to hydro-geological distresses or to generally critical events, and (ii) the "first response" to emergencies. The information management process must be carried out by always taking into account the end-to-end Quality of Information, after a suitable processing of the big data collected by all the measurements, by using the technique described in the next Section.

The proposed system exposes several Information Services each aiming at the provision of data about a specific issue of the natural resource protection process. The information provided by a generic Information Service is the outcome of a raw data processing specifically designed to obtain a high informational added value with reference to the original basic data. An Information Service may be requested by different actors, each with a different operational objective which requires different QoS contracts. An Integrated Information Service will decouple the Information Specific Services (*e.g., Landslide Information Service*, *Flooding Information Service* and *Field Information Service*), coming from differentiated end user communities.

Finally, a diversified packaged information provisioning is produced, aiming at satisfying all needs of (very) different users, through the public bodies. In addition, thanks to the adoption of advanced data-centric architectural models, the proposed solution has the great advantage of not interfere but only strengthen the procedures currently used by the Civil Protection to collect and manage the monitoring data. In what follows, some more details about the proposed system will be given.

III. Suggested Kalman-Kriging Method

An application of Kalman filter to analyse environmental data produced by the above mentioned sensor network is here described to overcome the possible presence of not complete time series of measured values. As well known [9], this recursive algorithm allows the data filtering and provides the best estimate of the analysed quantities, under a hypothesised steady-state or dynamic model of the observed quantities, even if they are unknown. In this way a considerable set of measured values can be analysed and all the desired information contained in big data can be successfully extrapolated, so to reduce the noise.

The application of this algorithm to measured data permits the user (*e.g.,* the environmental engineers) to study and characterise the dynamics of the observed critical quantities under observation by using a suitable mathematical models; hence initially known temporal relationships between different quantities can be accurately identified. For example, the daily averaged concentration of generic air pollutant p_i can be estimated by means of the averaged concentration values of other $n-1$ correlated pollutants, as follows [13]:

$$\hat{p}_i = k_1 p_1 + k_2 p_2 + \ldots + k_{i-1} p_{i-1} + k_{i+1} p_{i+1} + \ldots \\ + k_{n-1} p_{n-1} + k_n \qquad (1)$$

where the variables p_i represent the daily averaged values of all correlated quantities, and k_i are the coefficients of a suitable linear regression carried out by applying the Kalman algorithm to the measured values returned by the relevant sensors.

Considering a null initial condition to get started the Kalman filter, the coefficients k_i can be calculated by means of the following iterative expression:

$$\begin{bmatrix} \hat{k}_{1,j} \\ \hat{k}_{2,j} \\ \\ \hat{k}_{n,j} \end{bmatrix} = \begin{bmatrix} \hat{k}_{1,j-1} \\ \hat{k}_{2,j-1} \\ \\ \hat{k}_{n,j-1} \end{bmatrix} + + K_j \left\{ p_{i,j} - \begin{bmatrix} p_{1,j} & p_{2,j} & \cdots p_{n-1,j} & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{k}_{1,j-1} \\ \hat{k}_{2,j-1} \\ \\ \hat{k}_{n,j-1} \end{bmatrix} \right\} \quad (2)$$

where $\hat{k}_{r,j}$ and $\hat{k}_{r,j-1}$ are the estimates of coefficients at *j-th* and $(j-1)$*-th* step respectively, K_j is the *Kalman gain* at *j-th* step, and $p_{r,j}$ represent the *j-th* daily averaged values of all quantities under observation, respectively.

The above model can be applied to all kind of measured data for each one of fixed or mobile sensor of the network system. With the aim to improve the algorithm performance, one can consider also taking into account the data coming from both the other fixed stations of monitoring network and the space swarm sensor system of Fig.2, in order to identify in real time the spatial relationships between the environmental parameters of quantities under observation. To reach this objective, we suggest to apply to the measured data the Kriging technique, a specific geostatistical algorithm of interpolation that estimates quantity values in unknown area considering both the distances and the variations between the local values of the measured environmental quantities as well [13]. The method employs the *variogram* function representing the spatial variation and minimizes the estimated errors of predicted values carried out from the spatial distribution, by means of a suitable linear combination of the measured data located around the unknown data point to be estimated.

Being the Kriging algorithm based on the hypothesis of stationary data, while in general the behaviour of daily averaged values of the interest quantities is time-varying, it is necessary to split the total observation interval in a number of different classes where data can be considered quite stationary in each class, so to easily calculate the relevant variogram per class, defined as [13]:

$$\gamma_k(h) = \frac{1}{2N_k} \sum_{j=1}^{N_k} \left(Z_{kj}(0) - Z_{kj}(h) \right)^2 \quad (3)$$

where N_k is the number of measurement data in class k, $Z_{kj}(0)$ and $Z_{kj}(h)$ represent the *j-th* values of daily averaged of a generic environmental quantity relevant to class k, measured by sensor pairs with reciprocal distance h. The function that best fits the calculated experimental variogram is the exponential one, defined as follows:

$$\gamma_k(h) = c_k \left(1 - e^{-3h/a_k} \right) \quad (4)$$

where c_k and a_k are the *sill* and *range* values for class k, respectively. The function permits the estimation of the desired quantities, in each class k, in every point at distance h from the sensor of reference.

Finally, it can be applied a hybrid model based on the joint application of both Kalman and Kriging techniques. It is possible to rebuild the behaviour of daily averaged values of quantities relevant to a generic point without measured data, by using the Kriging model applied to the data coming from two measurement point near. Then, it can estimate the daily averaged values of a generic quantity or parameter by means of the eq. (1).

As an example of the application of the suggested method, in Fig.3 the daily averaged values of benzene concentrations in air of their relevant estimates obtained using Kriging model with their uncertainty bands are shown. By tacking a glance to this figure, it is possible to highlight the estimated values are always included within the uncertainty band.

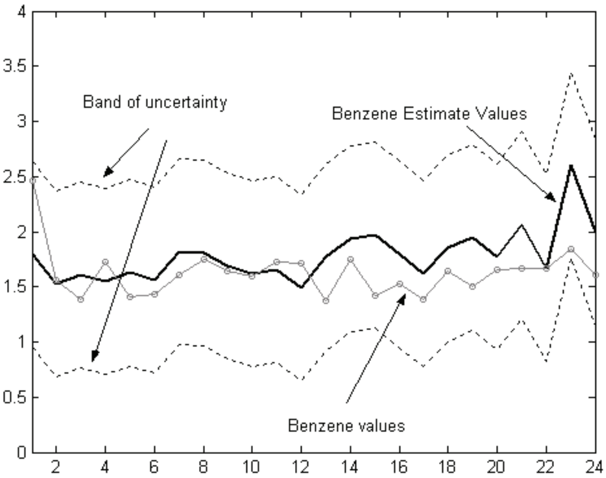

Fig.3: Behaviors of daily averaged values of benzene concentration (continuous bold line) of its estimate obtained using Kriging model (line with circle marker), and of uncertainty bands (dashed lines).

IV. Application of Ubiquitous Knowledge Base Model

In this Section we give some information about a transparent and suitable access to information associated to objects scattered in the environment, after the overall measurement process and by leveraging a semantic-based enhancement of them.

A. Bee Data Distribution Service (Bee-DDS)

Each agent in the swarm system of environmental sensors aims at the *provision of data referred to a specific item and/or subject of the natural resource protection process*. Moreover, a given informative service may be requested to the above mentioned DSS by different actors, with a different operational objective each. *Bee Data Distribution Service* (Bee-DDS), a *Leonardo Company™* message-oriented platform based on the publish-subscribe model, is the one adopted for the swarm setting. It provides affordable communication among loosely-coupled agents and nodes to support functionalities of monitoring, safety and recovery.

B. Ubiquitous Knowledge Base model

The *Ubiquitous Knowledge Base* (*u-KB*) model [14] grants transparent access to information embedded in semantic-enabled objects scattered in a given environment (see Fig.4, representing the proposed DDS Layered Architecture).

A KB is composed by a Terminological Box (TBox or *ontology*), *i.e.,* the formal representation of the conceptual model of the studying domain through a taxonomy of *classes* and *properties* 0 and an Assertion Box (ABox) specifying the factual knowledge concerning a specific problem, with *individuals* as instances of classes. Software tools called *reasoners* allow to derive implicit knowledge from what explicitly stated in a KB. In a u-KB, individuals are physically associated to distinct devices in a given environment (sensors, actuators, users). The ontology is fragmented in one or more *chunks* scattered across the network nodes.

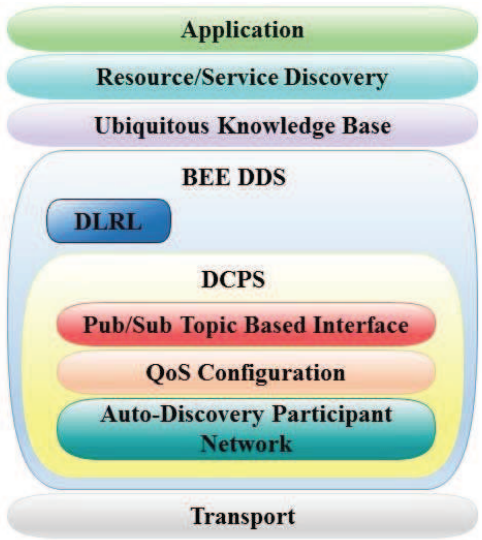

Fig.4: DDS Layered Architecture

C. Reasoning for critical environmental scenarios

Each of the nodes scattered in the environment manages a cache of *ontology chunks*. Every cache contains a static part of the ontology (named *Upper Ontology*, UO) as well as the chunk(s) required by semantic annotations. A *semantic service/resource request* is a logic-based annotation expressed w.r.t. a reference ontology.

After the metrological characterization of gathered data, an annotation phase is well suited, adopting a threshold decision system. It associates logic-based descriptions to (set of) raw data coming from the deployed sensor network. Furthermore, the *semantic-based reasoning* allows an *advanced decision support* as it enables novel peculiarities of resource discovery and integration:

- a node in Data Distribution System provides sensed data referred to a particular monitored area;
- this information could be automatically annotated with respect to a reference ontology;
- it can be exploited as request *to build a recovery team* composed in the best possible way.

V. CONCLUSIONS AND FUTURE WORK

This paper presented a proposal for a space-based sensor network system used as disaster management through “early warning”. This system employs a suitable joined application of Kalman filter and Kriging technique to the big data collected from both *in-situ* and mobile sensor networks, very useful for the environmental engineers and in general for the Civil Protection.

A continuous and accurate environmental monitoring in a disaster scenario or generally in a critical event is needed to grant an efficient and quick risk management. The suggested data processing technique allows to carry out a good estimate of analysed substances and to rebuild the possible missing data in time and space domains. By using this method, it is possible to work with a continuous and valid dataset, with the further aim of reducing the measurement uncertainty.

It has been shown also how the metrological characteristics of the whole sensor network could be further enhanced if provided data are annotated in a logic-based formalism and a deductive reasoning is performed on them, so that the whole proposed system should allow a good improvement of performance and reliability with respect to competitor frameworks. This measurement and management system is now under implementation, on a suitable publisher/subscriber Message-Oriented Middleware named Bee-DDS as reference architecture. As a future work, test results will be presented, with the aim of assessing feasibility and measuring performance.

REFERENCES

[1] D. M. Tralli, R. G. Blom, V. Zlotnicki, A. Donnellan, D. L. Evans “Satellite remote sensing of earthquake, volcano, flood, landslide and coastal inundation hazards” , ISPRS Journal of Photogrammetry and Remote Sensing, vol. 59, no.4, pp185-198, June 2005.

[2] F. Adamo, F. Attivissimo, C. Guarnieri Calò Carducci, A.M.L. Lanzolla, “A Smart Sensor Network for Sea Water Quality Monitoring”, *IEEE Sensors Journal*, vol. 15, no. 5, pp. 2514-2522, May 2015.

[3] K. E. Joyce , S. E. Belliss, S. J. McNeill, P. J. Glassey “A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters”, Prog. in Phys. Geo., June 2009.

[4] F.Adamo, C. De Capua, P. Filianoti, A. M. L. Lanzolla, R. Morello, “A Coastal Erosion Model to Predict Shoreline Changes”, Measurement Journal of Imeko, vol. 47 no.1, pp. 734-740, January 2014.

[5] Di Keith Smith, Environmental Hazards: Assessing Risk and Reducing Disaster, Routledge, 2013.

[6] V.Pelillo, L. Piper, A. Lay-Ekuakille, G. Griffo, A. Lanzolla, G. Andria, R. Morello “Geostatistical approach for validating contaminated soil measurements”, Measurement vol. 47, pp. 1016-1023, 2013, January 2014[13]

[7] S. Nazlibilek, O. Kalender, Y. Ege “Mine Identification and Classification by Mobile Sensor Network Using Magnetic Anomaly”, IEEE Transactions on Instrumentation and Measurement, vo 60 , no.3, , pp. 1028 – 1036, March 2011

[8] Jongeun Choi, Songhwai Oh, Roberto Horowitz, “Distributed learning and cooperative control for multi-agent systems”, Automatica, vol. 45, no. 12, pp. 2802–2814, December 2009.

[9] Hairong Qi, S. Iyengar;K. Chakrabarty "Multiresolution data integration using mobile agents in distributed sensor networks", IEEE Trans. on Systems, Man, and Cybernetics, vol 31, no.3, pp.383-391, August 2001.

[10] H. W. Soreson: "Kalman Filtering: Theory and Application", IEEE Press, New York, 1985.

[11] P. Brooker: "Kriging", Engineering and Mining Journal, Vol. 180, No. 9, 1979.

[12] S. Cacucci, F. Fedi, "Space-aided Autonomous Sensor System for Hydrogeological Disaster Management", Proc. of 2nd International Workshop on Metrology for Aerospace, pag. 240-244, Benevento, Italy, 4-5 June 2015.

[13] G. Andria, G Cavone, V, Di Lecce, A.M.L. Lanzolla: "Measurement and Characterization of Environmental Pollutants via data Correlation of Sensor outputs", Proc. the IMTC/03 Instr. & Meas. Techn.Conf., Vail-CO, USA, May 2003.

[14] M. Ruta, F. Scioscia, and E. Di Sciascio, "Enabling the semantic web of things: framework and architecture", in *Sixth IEEE International Conference on Semantic Computing (ICSC 2012)*. IEEE, 2012, pp. 345-347, doi: 10.1109/ICSC.2012.42

[15] S. Staab, R. Studer, "Handbook on Ontologies", *International Handbooks on Information Systems*, Springer, 2004

.