Internet of Things for Hydrology: Potential and Challenges

Andrea Zanella
Department of Information Engineering
University of Padova
Padova, Italy
Email: andrea.zanella@unipd.it

Sergio Zubelzu
Departamento de Ingeniería Agroforestal,
Universidad Politécnica de Madrid
Madrid, Spain
Email: sergio.zubelzu@upm.es

Mehdi Bennis
University of Oulu
Oulu, Finland
Email: mehdi.bennis@oulu.fi

Martina Capuzzo
Department of Information Engineering
University of Padova
Padova, Italy
Email: martina.capuzzo@unipd.it

Paolo Tarolli
Department of Land Environment Agriculture and Forestry
University of Padova
Padova, Italy
Email: paolo.tarolli@unipd.it

***Abstract*—The management of water resources has always been important for the sustainability of our society and economy. This need has been further increased by climate change in recent years that, among other effects, has led to an increase in extreme events, such as prolonged droughts, severe storms, hurricanes, and so on. It is therefore urgent and critical to develop new and more sophisticated tools and methodologies to observe and possibly predict fundamental water processes. Internet of Things and machine learning can provide a significant contribution to this end, which requires bridging the gap that still exists between the communities of hydrologists, data scientists, and communications engineers. This article aims to help fill such a gap by introducing engineers to the challenges of hydrology, and reviewing existing solutions proposed in the literature to such challenges. Some results obtained from empirical data sets are used to illustrate the main concepts and corroborate the theoretical discussion with some practical examples. Finally, open problems and possible avenues for future research are discussed.**

I. Introduction

Recently, many European countries have experienced the effects of climate change in the form of a scarcity of drinking water resources, prolonged periods of drought, and extremely heavy rainfall, with unprecedented dramatic environmental, economic, and social costs. Therefore, understanding, modeling, and predicting the movement and distribution of water on Earth and effectively managing water resources is more than ever vital for agriculture, industry, and society at large. Unfortunately, hydrology involves atmospheric, surface, and underground water systems, which are difficult to model on their own, and even more so when considered as a whole. As a result, modern hydrology often relies on a number of mathematical and empirical models that focus on isolated portions of the whole water cycle, thus providing only partial, defective and oftentimes inconsistent information. Moreover, such models are either based on complex physical theories that involve a large number of variables and parameters, which are often difficult to observe in practice and do not have a clear physical meaning, or are empirically obtained from observations, thus lacking generality, adaptability and interpretability.

In this scenario, the Internet of Things (IoT) technologies, together with machine learning (ML) methodologies, are instrumental in developing more accurate hydrological models and more effective water management strategies and interventions. More specifically, IoT and ML can bring innovations in three main directions, namely: (i) *data collection*, i.e., the measurement and gathering of relevant environmental and weather data; (ii) *data processing*, i.e., the analysis of the collected data to refine hydrological models and develop accurate predictions of extreme events; and (iii) *data visualization*, i.e., methodologies to ease the understanding of the models and related inferences by means of graphical tools.

In this paper, we discuss the state of the art in these three fundamental areas and the many open challenges that still need to be addressed to improve current hydrology models. We focus on the problem of determining the behavior of the water level along the main discharge river of a catchment during and after a storm. This turns out to be a critical aspect for the control of water locks in order to manage tributaries and the main channel and avoid, or limit, flood events. Thus, a proper modeling and prediction of this process is necessary. We stress out that this is a problem more complex than just precipitation estimation, which is only one of the elements that determine the level of the water in basins. Currently, one of the most effective techniques for weather prediction is based on weather radars, which provide models with high density spatial information on precipitation. However, they have the disadvantage of requiring high investment and operation costs and would only provide a partial answer to the estimation of the variability of hydrological processes. Indeed, since there is no equivalent measurement technique for the remaining processes causing spatio-temporal variability, as for example infiltration/runoff or flood routing, their benefits remain uncomplete.

 ISBN: 978-3-903176-56-0

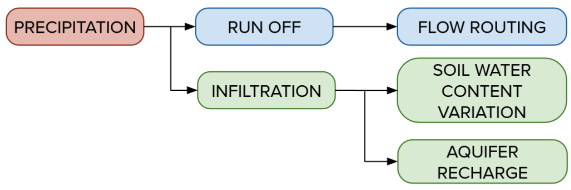


Fig. 1: Hydrological processes within a catchment.

This paper provides the following contributions:

(i) a high-level introduction to the fundamentals of hydrology to equip ICT practitioners with the basic knowledge to understand (and appreciate) the various challenges offered by the hydrological domain;
(ii) a review of the main solutions proposed in the literature regarding the problems of hydrological data detection, processing, and visualization;
(iii) a selection of results obtained in real-world scenarios to ground the discussion and exemplify possible solutions;
(iv) a reasoned discussion of the open challenges to be addressed and the possible approaches that can be applied to tackle such problems.

The rest of the paper is organized as follows. Sec. II provides a very compact introduction on basic hydrological processes. Then, Sec. III, IV and V discuss the data collection, processing, and visualization solutions, respectively. Finally, we conclude the paper with a discussion about possible approaches to tackle the open challenges in Sec. VI.

II. Basics on Hydrology

Predicting the water level along a river during and after a rainfall event requires to know a number of variables: how much water falls on different parts of the catchment; what fraction of such water will infiltrate the ground and what will flow to the surface (runoff) to the drainage channel; when and where the runoff water reaches the discharge channel. Fig. 1 exemplifies the main processes at catchment scale, with interest for events-based hydrology. A more detailed hydrological analysis, considering the specificity of the environment (topology, soil structure and absorbtion capabilities, ...) is out of the scope of this investigation, which aims to obtain more general considerations and methodologies.

The amount of rained water can be determined from the so-called *hyetograph*, i.e., the graphs of rainfall intensity over time in the catchment. Empirically, hyetographs can be built from sensors that measure the rain intensity during a storm. Fig. 2 shows an example of two hyetographs obtained in different days from the weather stations monitoring network of the Madrid City Council.[1] These data will be used also for the following examples, unless otherwise mentioned. In the figure, the hyetographs show storms with different behaviors: the graph on the left represents a storm that evolves with gradual intensity through several hours (9:00 – 24:00), reaching its peak in the late afternoon. The rain event represented by the

[1]https://www.mambiente.madrid.es/sica/scripts/index.php

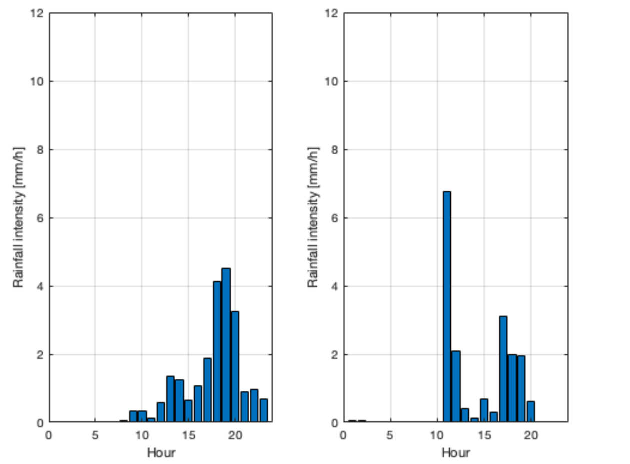
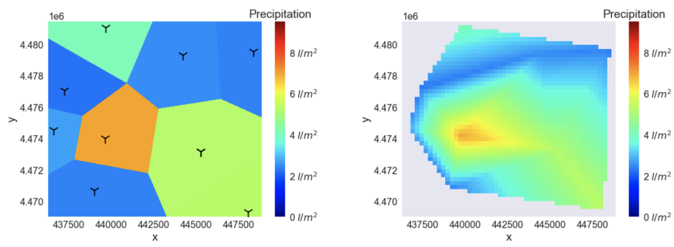


Fig. 2: Example of hyetographs for two different days in the Madrid basin. Each measurement refers to the amount of water falling in one hour.



(a) Thiessen polygons

(b) Barycentric interpolation

Fig. 3: Example of spatial interpolation techniques. Markers indicate the positions of the gauging stations in the Madrid catchment. The map on the left is obtained by using Voronoi tessellation, while that on the right is derived from a Barycentric interpolation.

image on the right, instead, reaches its maximum in the first hour of rain (11:00), then the rain almost stops and starts again in the late afternoon. Also from these examples, it is clear that the weather stations can only provide a discrete, sampled version of the two-dimensional time-varying continuous process of rainfall intensity. It is possible to obtain a measure of the total amount of water rained over the area during a given period by spatially interpolating the samples collected by different weather stations in the same basin, and then integrating over that period of time. However, the accuracy of the result depends on ow the interpolation is performed. In Fig. 3, for example, we report two interpolation maps obtained from the same dataset with different methods: since the computation of the interpolated values is different, the estimation of the total rained water obtained by integrating the values over the area of interest will yield to quite different results. Hence, choosing the proper interpolation techniques is critical.

Only part of the rained water flows into the drainage river, while the remaining fraction is absorbed along the way by the ground. Knowing the absorption rate of the soil, it is virtually possible to find the rate of surface runoff at each point in the basin and, in turn, the net contribution to the discharge river. Unfortunately, the infiltration process is difficult to monitor, and can be extremely heterogeneous over time and space. In fact, not only the capability of the ground to absorb water depends on the use of the terrain and on its composition, but it also changes during a storm with the humidity of the terrain itself. Finally, the *time of concentration*, i.e., the time taken by water to flow from a certain point in a watershed to the watershed outlet, depends on topography, geology, and land use and, hence, is very difficult to compute or predict.

From this quick overview it should be clear that processes involved in atmospheric water, surface water and subsurface water subsystems jointly evolve and interact with high spatial and temporal variability, determining the whole hydrological system behaviour. Therefore, despite the numerous theoretical and empirical models that have been proposed to describe each of these phenomena, modeling their interactions in a real catchment remains a formidable challenge.

III. Measuring and collection of hydrological data

Finding strategies for the optimal sensor location and techniques to combine the measurements of different types of sensors is fundamental to provide the hydrological models with informative data. The problem of optimal sensor positioning and sampling rate has been addressed in other contexts (e.g., [1], [2] including water distribution systems or storm water infrastructures [3]), but it remains widely unexplored in hydrology, where the problem is exacerbated by the spatio-temporal dynamics of the observed variables. Some results based on mathematical theories for optimal process monitoring have been recently published in [4]–[6]. A number of studies on monitoring hydrological processes have been previously conducted (a detailed list can be found in [7]), but the monitoring network is either oversized (e.g., [8] deployed 300 sensors for a 1 km^2 catchment), or severely undersized, with just a few gauging stations for large areas [9]–[11].

Also, it is possible to exploit the temporal correlation of the monitored variables to optimize their sensing and storage. Some works have explored the possibility of using real-time weather forecasts and hydrologic conditions to autonomously adjust the measurement frequency, in order to intensify observations during interesting events [12], [13].

How to optimally merge remote sensing and ground-based measurements, paying attention both to spatial and temporal variability of the observed processes, remains an open problem. The optimal combination will have to cope with risks from potential statistical correlation between satellite and *in situ* information [14], and allow for a significant improvement in the speed and accuracy of the identification of relevant events, for a given measurement rate and density of sensing locations. Following the International Association of Hydrological Sciences (IAHS): "…*the ability of hydrologists to accurately measure processes and state variables is a priority for improving the understanding of processes…*", it seems evident that the accurate process monitoring has to be imperatively based on a rational strategy for optimal monitoring of hydrological processes as a whole. A comprehensive approach for determining the optimal monitoring strategy, i.e., what variables have to be measured, where, how often and what is the optimal spatial and temporal combination of remote sensing and ground-based sensors, is still lacking.

Moreover, it might be necessary to increase the density of weather stations and sensors on the ground to collect more accurate data and build precise maps of the hydrological processes. This requires the deployment of a large number of nodes, thus increasing the capital and operational expenditures. Such deployments can benefit from energy-neutral gauging stations to reduce infrastructure and maintenance costs. In this way, sensor nodes exploit energy harvesting capabilities to get the energy required for sensing, processing and transmitting operations. This makes it possible to realize more extended networks and place sensors that can not be powered by a power grid. Furthermore, when compared to the use of batteries, solutions with energy harvesting are more sustainable and eco-friendly, since they do not require battery replacement. Although, the performance of these solutions can strongly depend on the energy availability, and require an accurate design of sampling and communication protocols that increase the energy efficiency of data collection [15]–[17]. Some examples are applications where it is possible to harvest energy from the environment (e.g., solar light, wind) or from the monitored process itself (stream water flows, geothermal phenomena).

From this discussion, it is apparent that a trade-off is present for sensor density, quality of information and costs. Indeed, the deployed network and its configuration (i.e., sampling time, sensors position) should be tuned to provide reliable information, enough to provide accurate and precise results for the desired purpose but, at the same time, avoiding the collection of redundant data, which would lead to the useless waste of memory, computational, and economic resources.

IV. Data processing and forecasting

Hydrological theories provide tools to deal with complex random phenomena with high spatial and temporal variability, in which physical approaches require relevant simplifications and/or high computational efforts. On the other side, empirical models, much more frequently used, are more scenario-specific and lack generality. Instead, machine-learning algorithms are instrumental in automatically estimating the models' parameters based on the actual data collected from the environment, in order to make them more flexible and adaptable to the underlying system characteristics. Such an approach has been successfully applied in other disciplines, e.g., telecommunications [18] and energetic engineering [19], but has only recently been considered in hydrology. For example, Random Forest, Support Vector Machine, and Extreme Learning Machine have been used in [20] for the estimation of daily reference evapotranspiration in the presence of

TABLE I: Main statistical moments considering the entire storms dataset.

Metric	Duration [h]	Volume [mm]	Average rainfall intensity [mm/h]	Maximum rainfall intensity [mm/h]
Maximum	21	47.1	8.07	23.7
Minimum	1	0.1	0.1	0.1
Average $\pm$ std	2.56 $\pm$ 3.04	1.69 $\pm$ 4.3	0.3 $\pm$ 0.65	0.81 $\pm$ 2.03
Mode	1	-	-	-

TABLE II: Main statistical moments considering storms generating runoff according to the Green-Ampt model.

Metric	Duration [h]	Volume [mm]	Average rainfall intensity [mm/h]	Maximum rainfall intensity [mm/h]	Runoff volume [mm]	Maximum runoff runoff intensity[mm/h]
Maximum	21	47.1	7.35	21.4	24.3	24.63
Minimum	2	5.7	0	1.9	0.1	0.17
Average $\pm$ std	8.44$\pm$0.9	14.54$\pm$1.62	1.82$\pm$2.93	5.81$\pm$2.12	5.55$\pm$1.74	5.2$\pm$1.71
Mode	5.44	8.75	1.34	4.25	5.97	5.59

meteorological data, while in [21], the authors developed a study for the prediction of water flow using artificial neural networks. In [22] Xu *et al.* integrated deep learning and the fire-fly algorithm for optimizing the parameters of a Support Vector Regression method to predict the hydrological process of Huangfuchuan in Fugu County, China. Multiple linear regression, K-nearest neighbors, Support Vector Regression, Cubist, Random Forest and Artificial Neural Networks have been widely used for mapping soil physical properties [23]. Furthermore, ML and deep learning techniques have been recently applied for water flow and flood prediction. In [24] the authors compare the performance of different types of neural networks for monthly flow prediction. However, the best solution turns out to be dependent on the specific case study, amount of available features and training observations. In [25] the authors couple convolutional neural networks with transfer learning to reduce the training time of the neural network in flood prediction problems.

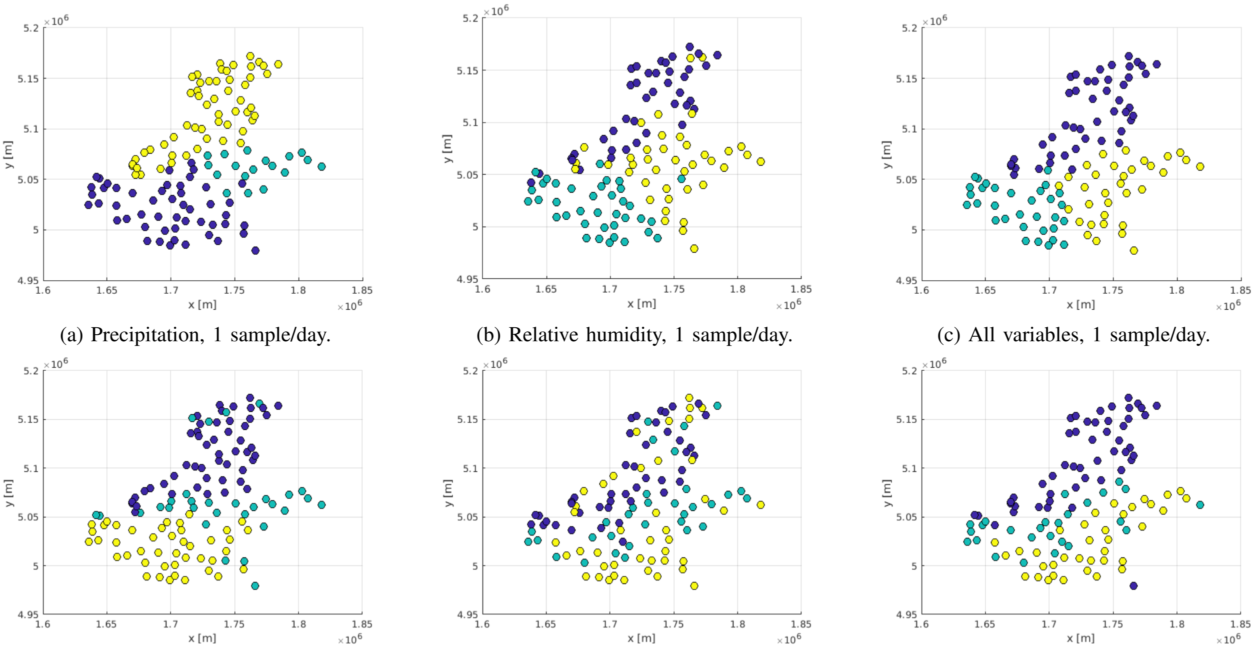
Other ML techniques have been applied to specific problems (signals decomposition, model parameters estimation, model matching, etc), overlooking the general problem of modelling the whole water cycle. One of the reasons is that the complexity of such general problem overwhelms the capability of classic, centralized ML approaches, which are not able to scale up with the number, high-dimensionality and heterogeneity of the input variables involved in the general model. Recently, promising results have been obtained by pre-processing the input signals to make them more informative, before feeding them into ML algorithms [26], [27]. However, no attempt has been made to model the whole water cycle, given its huge complexity, and the variety of input processes it hinges upon.

Existing hydrological models can be improved in different ways. First, simplifying the existing models can provide a better understanding of the most important parameters to be accurately estimated, helping to fit the model to the actual geographical scenario. Second, the exploitation of suitable ML techniques, specifically selected to the purpose, can contribute to automatically configure such parameters according to the characteristics of the region, using satellite maps and measurements provided by on-ground sensors.

A hybrid approach can be combining hydrological models with data-driven techniques. For example, Tab. I shows the statistics of all rain phenomena in the Madrid dataset. By considering [28], it is possible to obtain the soil parameters for the catchment and to apply the Green-Ampt model [29] to estimate the occurrence of runoff. This makes it possible to select the events generating the runoff and, then, applying ML techniques to investigate the behavior and the impact of other variables (e.g., pressure, relative humidity, temperature). Tab. II shows the statistical values when considering only the subset of precipitation events generating the runoff. As expected, these include heavier storm events, corresponding to longer storm duration and more intense rainfall.

It is worth noting that ML algorithms themselves may suffer from the *a priori* choice of the measured variables and sampling frequency. As an example, we clusterized metereological stations in the Veneto region, in the north of Italy,[2] by considering different combinations of the available variables and different sampling frequencies, employing the data collected in year 2017, with the purpose of identifying areas having homogeneous weather characteristics. Fig. 4 shows the results for different combinations when fixing the number of clusters to 3 and using K-medoids clustering algorithm. Timeseries of the values of the variables of interest were given as input features characterizing each weather station. It is apparent that, although the groups of stations are similar, roughly corresponding to the three geographical regions (mountains, plain, Venetian lagoon/coast), the actual clusters are different. For example, while the clusters obtained considering only the precipitation (Figs. 4a, 4d) are geographically compact, those obtained with the relative humidity (Figs. 4b, 4e) are more spread. Then, considering Figs. 4c, 4f, where all the variables are considered, we can observe that also the sampling frequency plays a role in the cluster determination. This example, although simple, shows how the choice of the input data, previously discussed in Sec. III, can impact the result of ML approaches, thus proving the importance of proper data selection.

[2]ARPAV dataset: https://www.arpa.veneto.it/bollettini/storico/Mappa_2022_TEMP.htm.



(a) Precipitation, 1 sample/day. (b) Relative humidity, 1 sample/day. (c) All variables, 1 sample/day.

(d) Precipitation, 4 samples/hour. (e) Relative humidity, 4 samples/hour. (f) All variables, 4 samples/hour.

Fig. 4: Clustering of meteorological stations in the Veneto region (IT) using as features different combinations of environmental variables (precipitation, relative humidity, air temperature, solar radiation) and sampling periods (24 h and 15'). Clusters are identified by different colors.

We stress out that the aim of data-driven approaches should not be too specific to the addressed scenario, but rather provide a framework that can be applied to different cases. For example, unsupervised deep learning and sensitivity/uncertainty analysis techniques for identifying the most relevant parameters for the accuracy of the model should be considered, as proposed in [26]. These methodologies, therefore, acquire general purpose and applicability, and can become an effective tool to study an area of interest. Once developed, the general ML architecture can thus be fed offline with historical data collected from the area to be studied, and be employed as a *black-box* tool to estimate the soil parameters and predicting the water distribution and movement according to the measured variables. Because of its use and possible computational cost, it is expected that such a framework will run offline and in a centralized manner, employing the available data collected previously.

V. Visualization tools

Currently, hydrological models are accessible only to experts, which know the meaning and role of the different models parameters, and have the skills to correctly read and interpret the models' outcome, which may be given in the form of probability distributions or time-series with uncertainty intervals. However, decision making in the water management domain involves a number of stakeholders, not all equipped with the technical background to correctly use the model, or interpret its outcomes, especially thematic maps.

It is hence fundamental to develop new techniques to visualize the output of the hydrological models in a graphical format, making it possible to "observe" the processes in the different subsystems and the water fluxes and storage within and among hydrological subsystems. Furthermore, the tool should make it possible to observe how the water level in the catchments is predicted to change when varying some of the system input variables (e.g., rainfall/sun/wind intensity, geomorphology, land uses, water allocation within the catchment). The uncertainty of the estimation should also be graphically represented, to provide an indication of the reliability of the prediction. Digital visualization tools in Geographic Information System (GIS) can be used to represent the results (e.g., surface morphology of terrain, soil moisture, surface water flow) [30], [31].

VI. Way forward

In this section we give an overview of the tools and research directions that can be explored to tackle the problem of water management.

In complex systems, including hydrologic ones, macroscopic effects arise from the interactions among lower-layer processes. To deal with such systems, it is hence natural to focus on the individual underlying elements, before attacking the system as a whole. In the case of hydrology, the macroscopic effects at basin level emerge from the interactions among atmospheric (rainfall), surface and subsurface water subsystems. In order to tame this complexity, it is worth considering each of the subsystems individually, before making an attempt to

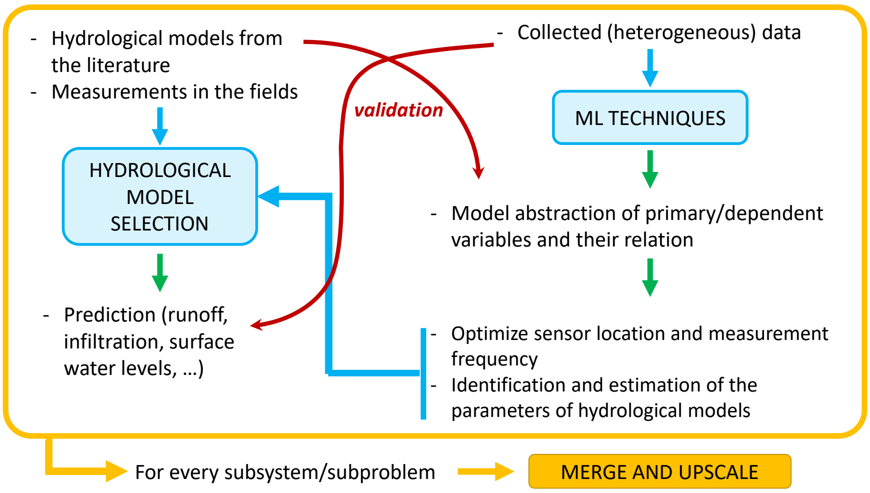


Fig. 5: Diagram representing the relations and dependencies among the various methodology steps.

derive a more comprehensive model for the whole water cycle in a catchment.

For the rainfall process, models should focus on both short-term forecasting, seeking to model individual storm events, and medium/long term for detecting trends relevant for water management and allocation. Conversely to current approaches in the literature, an attempt should be made to model individual storm events focusing on physical laws instead of defining purely autoregressive or model-free solutions. This "white box" approach, if successful, will not only provide efficient models to predict storm events, but also shed light on the parameters that have a stronger impact on the physical models, thus contributing to a better understating of the theoretical aspects of hydrology. Algorithms for detecting and predicting medium/long term trends are also needed to guide water management decisions, plans and policies. Satellite information can help with medium/long term forecasts, while short term characterisation and forecast can be aided by both ground-based sensors or radar measurements. In addition, portable radars can be used for retrieving highly detailed spatial information (for example portable X-band weather radars) to benefit from accurate and high spatial resolution measurements.

On atmospheric water subsystems, theoretical aerodynamic, energy balance or combined approaches, should be explored for modelling both evaporation and evapotranspiration.

Surface water subsystem encompasses infiltration and runoff-related processes. Remote sensing has been proved to be valid for directly measuring water content at upper soil layers, but the required temporal resolution for estimating infiltration evolution during storm events is not ensured by satellite images. For this reason, infiltration should be measured using ground-based sensors. On infiltration models, physically based approaches as those proposed by Richards, Green-Ampt and Soil Conservation System Curve Number method can be of use.

Infiltration phenomena are also linked to the groundwater dynamics, which affect the subsurface water subsystem. The movement of groundwater can be expressed by a variation of Darcy's law dependent on permeability aquifer's coefficient (or complementarity by transmissivity aquifer's rate) which in general terms can be easily determined by field measurements paying particular attention to soil spatial heterogeneity.

In summary, to address the challenges described in this paper, we recommend the following general methodology, whose main steps are represented in Fig. 5 along with their dependencies.

- Select reliable and accurate hydrologic models of the individual water subsystems whose parameters can be more easily estimated from environmental data.
- Exploit ML techniques to extract abstract models of heterogeneous sensing data, including remote sensing (satellite and aerial 2D/3D images), and ground measurements of primary variables (solar radiation, temperature, humidity, soil water content, water table level, etc.) and dependent variables (evaporation, evapotranspiration, precipitation, infiltration, surface water levels, velocity, depletion levels, etc.)
- Use such models to optimise sensor location and measurement frequency, in order to improve the quality (accuracy, reliability, information provided) of the data and reduce monitoring costs.
- Develop techniques to automatically configure the hydrologic models parameters based on the information obtained from the data, in order to improve their accuracy while reducing complexity.
- Merge and upscale the previous models to come up with a comprehensive model for the whole hydrological system, following the Reynold Transport Theorem, and solve the system by applying distributed learning techniques or PDE approximation methods based on neural networks with prior knowledge of the physical world.

The fusion of different types of measurements is hence a key characteristic of the proposed approach: the underlying assumption is that the ensemble of remote, aerial, and on-ground measurements contains all the information needed to correctly set the parameters of the selected hydrological models, and that such information can be automatically extracted from the data by means of proper ML approaches.

For validation and calibration of the developed ML models, physically based theories can be used namely: Reynolds' transport theorem for defining mass and energy fluxes at the whole hydrological system; theoretical aerodynamic, energy balance or combined approaches, for modelling both evaporation and evapotranspiration; precipitable water column, thunderstorm cell models, frontal storm, mesoscale convective systems for rainfall estimate; Richards [32], Green-Ampt [29], and SCS Curve number methods for estimating infiltration and soil water dynamics; unit hydrographs, lumped flow routing models and distributed flow methods based on Saint Venant's equation for surface water dynamics; and modified Darcy´s law for groundwater dynamics.

VII. Conclusions

An effective management of water resources is precious for social, economical and safety reasons. Addressing this problem is even more urgent when considering the presence of critical events produced by climate change. However, as explained and exemplified in the previous sections, the methods that are currently used (i.e., hydrological models) rely on many parameters and variables, whose meaning is not always related to physical quantities. Furthermore, they often need to be calibrated and fed with values coming from estimated or measured values that describe the soil characteristics, which prevents them from being generally applicable and requires extensive measurement campaigns on the field.

Current ICT solutions, such as sensor networks and artificial intelligence can be useful in the model definition and parameter tuning, based on measured data. Though, as explained in this work, modeling and predicting water flows and distribution means considering a complex system, which is affected by many variables and elements (e.g., soil characteristics and use, weather conditions). Hence, to attack the problem, there is a need to gain a better understanding of several aspects, related to the observation, processing, and interpretation of the hydrological processes. How many gauging stations are needed to correctly estimate the amount of rained water in a catchments? Which type of sensors should be provided to such stations? What should be the sampling period for the different environmental variables? How to avoid bias in the collected data? To what extent is it possible to predict events? Which hydrological theoretical/empirical models are more suitable to be combined with ML-aided parameters estimate? These are only a few of the open questions that call for a stronger cooperation between ICT and hydrology experts.

Acknowledgment

This research was funded by the CHIST-ERA grant CHIST-ERA-19-CES-002 and by MIUR (Italian Ministry for Education and Research) under the "Departments of Excellence" initiative (Law 232/2016). In Spain, this paper belongs to PCI2020-120694-2 project funded by MCIN/AEI/10.13039/501100011033 and the European Union "NextGenerationEU"/PRTR.

References

[1] D. J. Chmielewski, T. Palmer, and V. Manousiouthakis, "On the theory of optimal sensor placement," *AIChE journal*, vol. 48, no. 5, pp. 1001–1012, 2002.

[2] A. A. Alonso, I. G. Kevrekidis, J. R. Banga, and C. E. Frouzakis, "Optimal sensor location and reduced order observer design for distributed process systems," *Computers & chemical engineering*, vol. 28, no. 1-2, pp. 27–35, 2004.

[3] M. Bartos and B. Kerkez, "Hydrograph peak-shaving using a graph-theoretic algorithm for placement of hydraulic control structures," *Advances in Water Resources*, vol. 127, pp. 167–179, 2019.

[4] S. Zubelzu, L. Rodríguez-Sinobas, D. Segovia-Cardozo, and A. Díez-Herrero, "Optimal locations for flow and velocity sensors along a river channel," *Hydrological Sciences Journal*, vol. 65, no. 5, pp. 800–812, 2020.

[5] W. H. Asquith, "The use of support vectors from support vector machines for hydrometeorologic monitoring network analyses," *Journal of Hydrology*, vol. 583, p. 124522, 2020.

[6] L. Zhuo, Q. Dai, B. Zhao, and D. Han, "Soil moisture sensor network design for hydrological applications," *Hydrology and Earth System Sciences*, vol. 24, no. 5, pp. 2577–2591, 2020.

[7] F. Tauro, J. Selker, N. Van De Giesen, T. Abrate, R. Uijlenhoet, M. Porfiri, S. Manfreda, K. Caylor, T. Moramarco, J. Benveniste *et al.*, "Measurements and Observations in the XXI century (MOXXI): innovation and multi-disciplinarity to sense the hydrological cycle," *Hydrological sciences journal*, vol. 63, no. 2, pp. 169–196, 2018.

[8] B. Kerkez, S. D. Glaser, R. C. Bales, and M. W. Meadows, "Design and performance of a wireless sensor network for catchment-scale snow and soil moisture measurements," *Water Resources Research*, vol. 48, no. 9, 2012.

[9] E. Varouchakis, D. Hristopulos, and G. Karatzas, "Improving kriging of groundwater level data using nonlinear normalizing transformations—a field application," *Hydrological Sciences Journal*, vol. 57, no. 7, pp. 1404–1419, 2012.

[10] A. Nebenzal, B. Fishbain, and S. Kendler, "Model-based dense air pollution maps from sparse sensing in multi-source scenarios," *Environmental Modelling & Software*, vol. 128, p. 104701, 2020.

[11] S. Grimaldi, A. Petroselli, and F. Serinaldi, "A continuous simulation model for design-hydrograph estimation in small and ungauged watersheds," *Hydrological Sciences Journal*, vol. 57, no. 6, pp. 1035–1051, 2012.

[12] B. P. Wong and B. Kerkez, "Real-time environmental sensor data: An application to water quality using web services," *Environmental Modelling & Software*, vol. 84, pp. 505–517, 2016.

[13] ——, "Adaptive measurements of urban runoff quality," *Water Resources Research*, vol. 52, no. 11, pp. 8986–9000, 2016.

[14] J. Dong, L. Wei, X. Chen, Z. Duan, and Y. Lu, "An instrument variable based algorithm for estimating cross-correlated hydrological remote sensing errors," *Journal of Hydrology*, vol. 581, p. 124413, 2020.

[15] A. Biason, C. Pielli, A. Zanella, and M. Zorzi, "Access control for IoT nodes with energy and fidelity constraints," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3242–3257, 2018.

[16] C. Pielli, D. Zucchetto, A. Zanella, and M. Zorzi, "An interference-aware channel access strategy for WSNs exploiting temporal correlation," *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8585–8597, 2019.

[17] M. Capuzzo, C. Delgado, J. Famaey, and A. Zanella, "An ns-3 implementation of a battery-less node for energy-harvesting internet of things," in *Proceedings of the Workshop on ns-3*, 2021, pp. 57–64.

[18] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.

[19] X. Chen, K. Yu, W. Du, W. Zhao, and G. Liu, "Parameters identification of solar cell models using generalized oppositional teaching learning based optimization," *Energy*, vol. 99, pp. 170–180, 2016.

[20] J. Fan, W. Yue, L. Wu, F. Zhang, H. Cai, X. Wang, X. Lu, and Y. Xiang, "Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China," *Agricultural and forest meteorology*, vol. 263, pp. 225–241, 2018.

[21] M. Kovačević, N. Ivanišević, T. Dašić, and L. Marković, "Application of artificial neural networks for hydrological modelling in karst," *Graevinar*, vol. 70, no. 01., pp. 1–10, 2018.

[22] L. Xu, J. Zhao, C. Li, C. Li, X. Wang, and Z. Xie, "Simulation and prediction of hydrological processes based on firefly algorithm with deep learning and support vector for regression," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 35, no. 3, pp. 288–296, 2020.

[23] Y. Khaledian and B. A. Miller, "Selecting appropriate machine learning methods for digital soil mapping," *Applied Mathematical Modelling*, vol. 81, pp. 401–418, 2020.

[24] Z. Alizadeh, J. Yazdi, J. H. Kim, and A. K. Al-Shamiri, "Assessment of machine learning techniques for monthly flow prediction," *Water*, vol. 10, no. 11, p. 1676, 2018.

[25] N. Kimura, I. Yoshinaga, K. Sekijima, I. Azechi, and D. Baba, "Convolutional neural network coupled with a transfer-learning approach for time-series flood predictions," *Water*, vol. 12, no. 1, p. 96, 2020.

[26] M. Zorzi, A. Zanella, A. Testolin, M. D. F. De Grazia, and M. Zorzi, "Cognition-based networks: A new perspective on network optimization using learning and distributed intelligence," *IEEE Access*, vol. 3, pp. 1512–1530, 2015.

[27] M. Rezaie-Balf, O. Kisi, and L. H. Chua, "Application of ensemble empirical mode decomposition based on machine learning methodologies in forecasting monthly pan evaporation," *Hydrology Research*, vol. 50, no. 2, pp. 498–516, 2019.

[28] R. F. Carsel and R. S. Parrish, "Developing joint probability distributions of soil water retention characteristics," *Water Resources Research*, vol. 24, no. 5, pp. 755–769, 1988. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR024i005p00755

[29] W. Green, "Ampt. GA: Studies on soil physics. 1. Flow of air and water through soils," *J. Agr. Sci*, vol. 4, pp. 1–24, 1911.

[30] S. Cucchiaro, D. J. Fallu, H. Zhang, K. Walsh, K. Van Oost, A. G. Brown, and P. Tarolli, "Multiplatform-SfM and TLS data fusion for monitoring agricultural terraces in complex topographic and landcover conditions," *Remote Sensing*, vol. 12, no. 12, p. 1946, 2020.

[31] A. Pijl, L. E. Reuter, E. Quarella, T. A. Vogel, and P. Tarolli, "GIS-based soil erosion modelling under various steep-slope vineyard practices," *Catena*, vol. 193, p. 104604, 2020.

[32] L. A. Richards, "Capillary Conduction of Liquids through Porous Mediums," *Physics*, vol. 1, no. 5, pp. 318–333, 1931.