H SAF Associated Scientist Activity

HSAF_VSA20_03

Precipitation Downscaling of H SAF Product in different Hydrological Scenario using a Cellular Automata Approach

VSA: Annalina Lombardi (CETEMPS – University of L'Aquila)

VSA supervisors: Paolo Sanò (CNR ISAC); Daniele Casella (CNR ISAC); Luca Brocca (CNR IRPI)

H SAF Science Coordinator: Giulia Panegrossi (CNR ISAC)

CONTENTS

1. Introduction

The precipitation is one of the most important components of the hydrological cycle and the driving force for severe flood events: consequently, the accuracy of flood forecasting is linked to the accuracy of the precipitation estimation. In hydrological models, the rainfall patterns strongly impact the runoff calculation because the relationship between rainfall distribution and computed discharge is nonlinear (e.g., Goodrich et al.; 1997, Singh, 1997; Cristiano et al, 2017).

For hydrological operational activity, hydrological models are usually forced with observed and predicted rainfall data, and the uncertainty of hydrological forecasts is strongly related to the uncertainty of the rain field estimates of atmospheric forecasting models. It is very important to force hydrological models with realistic observed precipitation data, which are fundamental to the spin-up process, to reduce this uncertainty. The rain gauge data are used as the main information (Nikolopoulos et al., 2010), even if rainfall spatial pattern from rain gauges is affected by errors, depending on data scarcity, sparse sensor network, associated to the lack of a robust or redundant infrastructure, able to guarantee data transmission and functionality during a severe weather event. Indeed, rain gauges can be considered as the most accurate sensors for measurements over a limited area (nearly a point), but their small coverage (especially over complex terrain and tropical regions) limits the adequacy in representing the spatial structure of highly variable rainfall fields over large spatial scales: therefore, rain gauges data could be considered a local information, discontinuous over time and especially in space. Moreover, the global decline of rainfall networks over time has proved to be disadvantageous; this has led researchers to consider the use of satellite-derived rainfall estimates (SRFE). During the time, there has been significant development in space-based precipitation estimation very important in hydrological applications. Satellite sensors offer unique advantages compared to gauges and weather radars because they provide a global coverage and observations in regions where in situ data are inexistent or sparse. Because of this uniqueness, the use of satellite data for hydrologic applications has gained growing interest (Guetter et al., 1996; Tsintikidis et al. 1999; Wilk et al., 2006; Hughes, 2006; Su et al., 2008; Collischonn et al.;2008; Thieming et al. 2013; Jiang and Wang, 2019; Darko et al. 2021).

Dembélé et al. (2020) highlight that although satellite products are characterized by uncertainties, their most reliable key feature is the representation of spatial patterns, which is a unique and relevant source of information for distributed hydrological models. Their results demonstrate that there are benefits in using satellite data sets, when suitably integrated in a robust model parametrization scheme. Shi et al. (2020) suggest that simulation results of the hydrological model using an appropriate method for merging rainfall data can provide valuable spatially distributed precipitation. Several techniques have been proposed to merge different sets of data and reduce uncertainties in rainfall estimation. Main used techniques are based on a physical approach or statistical algorithms (e.g., French & Krajewski, 1994, Todini, 2001).

Anyway, the assimilation and downscaling of precipitation data from different sources involves a deep understanding of the source of the observations, its characteristics and its limits.

The precipitation data are characterized by complex patterns and by a high spatial variability, which increases in a complex orography: consequently, the information provided by the different observations can help to

adequately reproduce this complexity. The main aim of this work is to present a method of generating a synthesis of a discrete set of point rain gauge observations and a satellite derived rainfall product, focusing on small-medium scale river basins.

Another goal of this work is also to assess if using the combination of rain gauges data and MW/IR H03 product and MW-only precipitation product H68 could overcome the specific lack of information provided by in situ measurements, useful for defining the local rainfall amounts and intensity. These data sources could be used to obtain a mutual correction of the implicit error characteristic of the different data.

2. Preliminary Analysis

Cetemps Hydrological Operational activity is divided in 7 sub-domains, each domain runs at different spatial resolution based on the area characteristics and computational time. Daily operational activity over Italy is available at link: http://cetemps.aquila.infn.it/chym/newoper/

Three different basins have been selected for this study (Fig.1):

a) The Po River Basin (upper drainage network) is an international watershed and the largest Italian: its surface extends for about 74 000 km^2, of which about 71 000 km^2 across the Italian territory, which means a quarter of the entire national territory; it is the main Italian river also for length, 652 km. The Po, whose headwaters are on the northern slope of Monviso in Piedmont, is fed by 141 tributaries along its course. The complexity of Po River basin is highlighted by significantly different hydrological behaviors and ecosystems that coexist and coevolve. The Po River Basin Authority has identified 12 different fluvial regimes which depend on the spatial distribution of rainfall over the catchment. In the northwestern part of Italy's Piedmont region, the Tanaro River is the most significant right-side tributary to the Po in terms of length, 276km, size of drainage basin, 8.324 km2 (partly Alpine, partly Apennine), and an average flow discharge of 123 m3/s. It is a river that flows eastward across northern Italy starting from Monte Saccarello (2201 m), Ligurian Alps, near the border with France.

b) The Tevere River Basin is the second longest Italian river after the Po, rising on the slope of Monte Fumaiolo, a major summit of the Tosco-Emiliano Apennines. It is 405 km long. Twisting in a generally southerly direction through a series of scenic gorges and broad valleys, the Tiber is the fourth river for outflows and the third for length and contributes, for about 20%, to the fluvial inputs in the Tyrrhenian Sea. Its catchment area covers an area of 17375 km^2, equal to about 5% of the national territory

c) Volturno River Basin is in the south-central Italy. It rises in the Abruzzese Apennines near Alfedena and flows southeast as far as its junction with the Calore River near Caiazzo. It then turns southwest, past Capua, to enter the Tyrrhenian Sea at Castel Volturno, northwest of Naples. The river is 175 km long and has a drainage basin of 5450 km^2.

Their main settings and characteristics are reported in the Table. 1.

Figure 1. Selected River basins over Italian Peninsula. @Google Earth

Table 1. Main hydrological model simulation settings and characteristics: NLON and NLAT are the grid point numbers; SLON and SLAT are the first latitude and latitude, located on the south-est

Domain	Basin	Station	Basin Drainage Area (km²)	Estimated Drainage Area (km²)	SLAT; SLON	NLONXNLAT	Spatial resolution (m)
01	Po	Isola S. Antonio	25640	25641	✓ 44.10 ✓ 6.50	330X280	860
		Alba -Tanaro	8175	8121			
02	Tevere	Pierantonio	1953	1913	✓ 41.71 ✓ 11.59	740X470	291
		Ponte Felcino	2087	2123			
03	Volturno	Benevento - Calore			✓ 40.50 ✓ 13.60	400X400	441
		Solopaca - Calore	2966	3130			

In this work, the hourly rain gauge record data are provided from Dewetra Platform (Italian Civil Protection Department and CIMA Research Foundation, 2014; https://www.cimafoundation.org/fondazioni/ricerca-sviluppo/my-dewetra.html).

The precipitation (frequency, intensity, type, and quantity) is a key variable for specifying the state of the climate system. It varies considerably in space and time and requires a high-density network to observe its variability and extremes. A network of rain gauges represents a finite number of sample point of the two-dimensional pattern of precipitation depths.

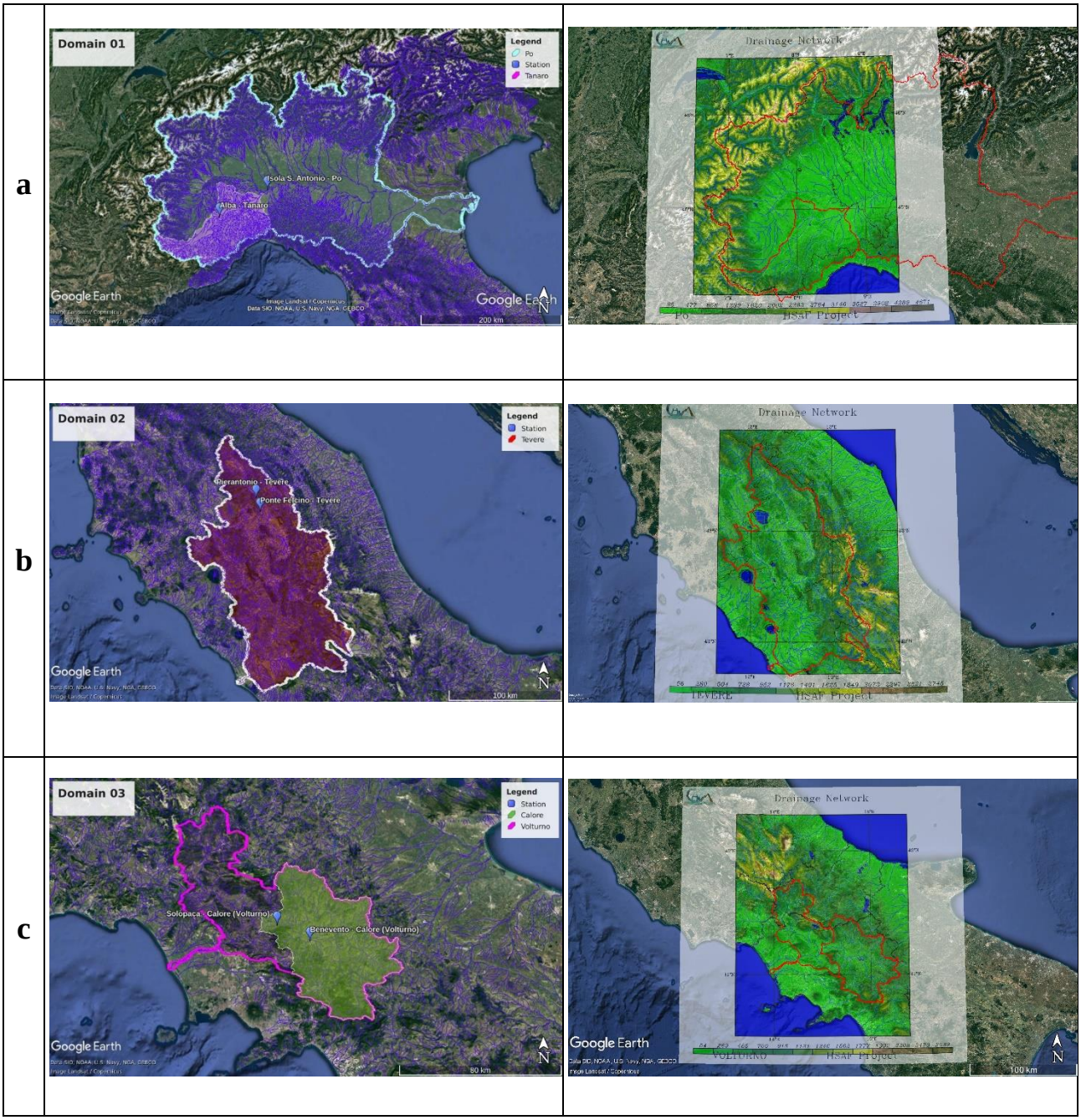


Figure 2. Three different selected river basins: a) Upper Po River Basin; b) Tevere River Basin; c) Volturno River Basin. The left panels represent the considered basins and the station localizations; the right panels show the rebuilt domains by the model. @Google Earth.

The accuracy of areal estimation depends on both the total number of gauges and their spatial distribution and a minimum density standard for constructing precipitation gauge networks is established (Sevruk, 1992; WMO, 1994).

The minimum density of a rain gauge network depends on a series of factors including the temporal resolution of the precipitation measurements and the morphological characteristics of the region considered. For these

reasons, WMO has established a minimum density standard for the construction of precipitation measurement networks as reported in Liang et al., 2012. The optimal spatial density of the monitoring stations can vary from one station every 100 km^2 for the lowland areas to one station every 50 km^2 for the hill and mountain areas. However, these considerations are not general; they change with the morphological complexity of the territory, the local environmental conditions, and the micro-climatic variability. The size and duration of the different characteristics of the precipitation areas go up to the very small convection cell in which the rain field is considered nominally uniform. A relationship between the size of the precipitation phenomena and the spacing between them has been observed and these regularities characterize the horizontal structure of the precipitation. Any precipitation event has an effective radius of influence for any considered area, which reflects the influence of a rain bandwidth (e. g. Duque-Gardeazábal et al., 2018).

Generally, an area of influence is assigned to each rain gauge in the network assuming that the precipitation field does not vary within the defined radius. The rain gauge is in the center of its circular area of influence of radius, R, defined as the radius of influence of the rain gauge.

In Shi et al. (2020) is reported that this radius of influence is not larger than a certain distance since the rain bandwidth in most regions are within 10–50 km. Based on this assumption, the mean distance between available station has been computed as:

$$L = \sqrt{\frac{S}{N}} \tag{1}$$

where L is the mean distance between stations; S is the considered area and N is the number of rain gauges.

All the defined domains are divided into three territorial sectors: mountain (h> 700 m), hilly (300 m < h < 700 m) and flat (h< 300 m) area and it is possible to calculate the average rain gauge distance for each sector (see Table 2). Obviously, it should be considered that having carried out 5-year long hydrological simulations, the rain gauge numbers change over time on the same domain.

Table 2. Density of a rain gauge network over selected domain.

Domain	Basin	Domain Number raingauge	RG N° on Basin	Mean Distance	Mountain area		Hilly area		Flat Area	
				L (km)	N	L (km)	N	L (km)	N	L (km)
01	Po	Up to 460	Up to 280	9.7	162	8.5	78	8.9	46	12.8
02	Tevere	Up to 393	Up to 207	9.5	36	11	75	9.8	96	7.7
03	Volturno	Up to 229	Up to 56	11	10	13	10	10	21	9.9

3. Precipitation Data Assimilation

Precipitation data are characterized by large spatial and temporal variability: the rainfall patterns play an important role in runoff estimation due to the strong nonlinear relationship between rainfall distribution and river discharge.

The accuracy of results of many hydrological computations depends on the accuracy of Areal Precipitation Estimates (APE, see Nemec,1986). The methods of APE are subject to various errors which considerably restrict our ability to model hydrological processes. However, obtaining the true amount of precipitation, P, over an area using conventional measurement techniques is almost impossible. For elements showing a large spatial variability in each time interval, for example hourly or daily values of P, the crucial aspect of the conventional technique is the design of the network and the resulting sampling error. This implies the solutions to such important issues as the number and location of gauges and the representativeness of a point value of P under the given physical conditions of an area. The general theory of this problem has been presented by Gandin (1970). In this context, it is important to recognize that regional, national, and private networks of gauges are generally not distributed to satisfy the specific requirements of the hydrologist or conditions of a particular basin.

This shows that the applicability of conventional APE methods for hydrological purposes depends on the systematic and random errors of punctual measurements of precipitation, P, and on the representativeness of the punctual values of P. The solution to these problems is mandatory, especially in small mountainous basins. Furthermore, it often happens that the spatial sampling of the rain gauge monitoring network does not respect a minimum density standard established by the WMO, as in the case of Italy. In the areas where rain gauges are very limited or even absent, an improvement using modern remote-sensing techniques such as radar, satellite imagery is essential. Without the usage of satellite rainfall estimation, the uncertainty of rainfall estimation could approach infinite in a sparse area. Instead, the auxiliary information of the remote-sensing data is used to bound the uncertainty of the blended rainfall estimation in such areas, although it is generally considered large. Given a selected domain, CHYM application rebuilds hourly precipitation areal at hydrological scale for hydrological simulation (Coppola et al, 2007), which depends on the spatial distribution of the available data. The model estimates a preliminary rainfall field, that we define as the Precipitation Background Field (PBF), to speed-up the numerical processing. All cells of the lattice are initialized with a value calculated with Cressman algorithm (Cressman, 1959), a nonparametric kernel smoothing interpolation method. Because of its simplicity, the Cressman method can be a useful starting tool (Bouttier and Courtier, 1999). The method does not rely on the stationarity assumption and can be applied to the cases where geo-statistical assumptions are violated. In the literature, the merging methods based on Cressman objective analysis scheme is considered a direct merging (e.g., Pereira Filho et al., 1998; Goudenhoofdt and Delobbe, 2008). However, the direct merging produces a significant bias around the areal boundary (i.e., the boundary between two consecutive grids of data) caused by the discontinuity of background field (i.e., gridded data). According to Li and Shao (2010), the accuracy of the merged field is depending upon the used kernel function, *K*, as in the case of the Cressman algorithm:

$$K(u) = c\frac{1-u^2}{1+u^2}I\{|u| \leq 1\} \quad (2)$$

where *c* is a constant chosen to satisfy the requirement $\int K(u)du = 1$, the kernel.

This is an example of a nonparametric regression model which uses a kernel function to weight surrounding observations by distance. In addition to defining a kernel function, it is also necessary to select a rain radius of influence that reflects the rain bandwidth defined by the climatology of the area and the number of local rain data available. The radius of influence, R, rescales the spread of kernel function and determines the smoothness of the estimated field.

It is intuitive that a small radius leads to a rough estimate field (small bias, large variance), and a large Radius creates a smooth surface (large bias, small variance).

Based on these considerations, given a discontinuous background field, the rainfall for each grid point of the selected domain can be calculated as follows

$$P_i = \sum_j \frac{1-(r_{ij}/R)^2}{1+(r_{ij}/R)^2}P_j \quad (3)$$

where P_i is the estimated rain value, P_j is the rainfall measurement available within a radius, R, and r_{ij} are the distances between for example, rain gauge locations and the cell. Obviously, the first difficulty lies in selecting the reasonable value of R.

Indeed, even if observations are available at any points of the grid and no error is assumed, the rain field rebuilt by the direct merging method is going to produce large boundary bias: smaller value of R means the bias occurs in a smaller area around the boundary. However, R selection is not a remedy to the boundary bias because the rain bandwidth is likely to be large when observed points are distributed irregularly. The problem of boundary bias is caused by the fact that the background field is discontinuous due to discretization of the field, but the nonparametric merging method is only able to generate continuous surfaces.

A smoothed merging idea to reduce the boundary bias arising from merging gridded and point data, based on the double smoothing estimation: pseudo-observations are produced by coarse interpolation; the final estimate is carried out using both original and pseudo-observations (Li and Shao, 2010; Duque-Gardeazábal et al., 2018). The basic idea used in this work is the same, although the technique changes. To improve performance of the assimilation of Satellite data and possible quantitative efficiency in the presence of sparse gauge stations we implemented, in the hydrological model Cellular Automata algorithm (Coppola et al, 2007), a technique able to interpolate the value of a random field. According to Cellular Automata theory, a CHyM grid is considered an aggregate of cellular automata and the status of a cell corresponds to the value of a CHyM precipitation field. The status s_i of the generic cell *i* of the lattice is updated according to following rule:

$$s_i = s_i + \alpha\left(\sum_{i=1}^{8}\beta_j(s_j - s_i)\right) \quad (4)$$

where s_i is carried out over all 8 surrounding cells. The coefficients β_j allow to consider the different distances between the cells, as an example for a regular equally spaced lattice, they assume the value 1 for the cells in North, East, South and West location, and the value $1/\sqrt{2}$ the cells located in the North-East, North-West, South-East and South-West direction respect to the cell i-th. The coefficient α assumes a small value (typically from 0.1 to 0.9) to ensure a slight smoothing of the original matrix: this smoothing rule is performed several times until stable status is reached, and all cells are updated synchronously. The grid point associated with the rainfall value available in the considered database is not modified by the algorithm.

In details, a regular lattice, with sites where each of them takes N possible values, is updated in discrete time steps according to the previous rule that depends on the status of the site and on the eight neighbouring cells.

The cellular automata method allows to perform a downscaling of the rain field, useful for the high resolutions needed by hydrological simulations and allows to use different sources of precipitation data.

In this preliminary analysis, the modular approach has been used: the different data sets are assimilated using a hierarchical sequence of modules; in this way it is possible to consider the different nature of the data. Therefore, the lattice of the considered domain can be divided into as many subdomains as the available data sources. Each subdomain can be defined as a set of grid points that have at least one rainfall value in a selected R, where it has a typical value of a few kilometres, depending on the density of the available data: tree different radius of influence has been selected for this work, with different coverage of the rain gauge data compared to the satellite data. Using the cellular automata technique, this study aims to identify how input data settings affect model performance, if merging rain gauge and satellite data improves hydrological outputs.

4. Software development

To this study, CETEMPS bought a new server machine for a better computational performances and reliability (Server HPE Proliant DL560; 4 processors Intel Xeon-Gold 6252N @2.3GHz (24-core); 512 GB RAM DDR4 2933 MHz). All the data were converted from GRIB to NetCDF format, using a the ***pygrib2nc*** python 2.0 script, developed in this project by CNR ISAC. A new script has been developed from scratch in python 3.8 which is the python version installed on the new server for operative purposes, because available tools are developed in python 2.0 and therefore are incompatible with 3.8, as also suggested by the python community. Then, the NetCDF output files have been furtherly converted into a proprietary binary format compatible with the hydrological model, using another ad-hoc tool, developed in Fortran, able to generate the binary data. Once all the tools have been debugged and finalized, annual file data at hourly resolution (hsafrainYEAR.dat) have been generated and ported into our database.

It must be stressed out, along all the 5 years of data processed, about 5% of the hourly matrices are not available in the database for H03B data, nevertheless more are missing in H68. Because of this, a new subroutine has been developed into the CHyM libraries (*hsafrain*) able to recall the file from the archive and read the correct hourly matrix (Fig. 3, 4).

```
c------------------------------------------------- Interface to HSAF archive ---
      subroutine hsafrain(ora,giorno,mese,anno,pi,la,lo,n)
      implicit none
      integer ora,giorno,mese,anno,n ; real pi(1),la(1),lo(1)
      integer npt
      real , allocatable , dimension(:) :: rain,lat,lon ; save rain,lat,lon
      integer lun,year,odom ; data lun,year,odom /3*-1/ ; save lun,year,odom
      integer nlon,nlat,i,j,orec,last,xflag,nrec,index1h ; save orec,last
      if (year.ne.anno) then
         call openmuseodb(lun,'hsafrain',anno)
         read (lun) ; last=0
         do i=1,index1h(23,31,12,anno)
            read (lun) nlon,nlat,xflag ; if (xflag.eq.1) last=i
         enddo
         rewind(lun)
         read (lun) year,nlon,nlat
         if (allocated(rain)) deallocate(rain) ; allocate(rain(nlon*nlat))
         if (allocated(lat)) deallocate(lat) ; allocate(lat(nlon*nlat))
         if (allocated(lon)) deallocate(lon) ; allocate(lon(nlon*nlat))
         rewind(lun)
         read (lun) year,nlon,nlat,(lat(i),i=1,nlon*nlat),
     2                             (lon(i),i=1,nlon*nlat)
```

Figure 3. Source code of the fortran subroutine added to the CHyM libraries to recall the file from the archive and read the correct hourly matrix.

```
         lfile=lfile(7:10)//'/'//lfile
      else if (file(1:15).eq.'MonteMidia.grid') then
         lfile='radar/'//lfile
      else if (file(1:10).eq.'MonteMidia') then
         lfile=lfile(11:14)//'/'//lfile
      else if (file(1:5).eq.'Micra') then
         lfile=lfile(6:9)//'/'//lfile
      else if (file(1:6).eq.'wrfrai') then
         lfile=lfile(7:10)//'/'//lfile
      else if (file(1:3).eq.'wrf') then
         lfile=lfile(4:7)//'/'//lfile
      else if (file(1:8).eq.'hsafrain') then
         lfile=lfile(9:12)//'/'//lfile
      else if (file(1:7).eq.'dewrain') then
         lfile=lfile(8:11)//'/'//lfile
      else if (file(1:7).eq.'dewtemp') then
         lfile=lfile(8:11)//'/'//lfile
      else if (file(1:7).eq.'dewidro') then
         lfile=lfile(8:11)//'/'//lfile
      else if (file(1:7).eq.'dewdisc') then
         lfile=lfile(8:11)//'/'//lfile
      else if (file(1:9).eq.'worldtemp') then
```

Figure 4. List of assimilated data into CHyM Code

Moreover, into the model main script a safety check is performed on the availability of the H SAF data before to run the cell-automata procedures. This check is needed for the merging of the GAUGE-SAT data when a different radius of influence, R, is forced to the model, since a coverage of the grid point of the matrix must be guaranteed even if a satellite data would not be available (Fig. 4).

```
rain=0.0
rsrm=0.0
ca=8
do i=1,nlon ; do j=1,nlat
   if (luse(i,j).eq.mare) ca(i,j)=-2
end do ; end do
rainsrc=' '
write(6,'(a10)') 'ceck data'
call gmafrommm5index(time,ora,giorno,mese,anno)
call hsafrain(ora,giorno,mese,anno,pi,la,lo,n)
if (n.eq.0) then
        rchym(10)=35.0
        print*,'nosat data'
else
        rchym(10)=3.0
endif
print*,rchym(10),'RADIUS GAUGE'
if (srcflag(01)) call buildrainfield(1,ca)      ! raingauges
if (srcflag(08)) call buildrainfield(8,ca)      ! hsaf
if (srcflag(02)) call buildrainfield(2,ca)      ! radar
if (srcflag(05)) call buildrainfield(5,ca)      ! WRF Domain 2
if (srcflag(04)) call buildrainfield(4,ca)      ! WRF Domain 1
```

Figure 5. Source code of the check routine for the merging of the GAUGE-SAT data when a different Radius of Influence is force to the model.

5. Hydrological Evaluation

One of the effective strategies for the validation of satellite data is an indirect method through a hydrological assessment. The main critical issue in this method is the presence of the anthropic impact on the considered area, such as the highly regulated basins that increase the difficulty in hydrological modelling because simulations reproduce the natural flow discharge of rivers and data on artificial water management are not available. A preliminary analysis of both natural and anthropogenic characteristics of the basin and a selection of the hydrological period to be simulated should be a mandatory approach for the evaluation of the Satellite Precipitation Product. The indirect validation method is based on the ability to reproduce the observed streamflow, where the hydrological model has been forced with satellite product. the idea is to find the optimal source of observed data to rebuild a realistic rain areal and to do this a comparison and sensitivity study is carried out where, in addition to forcing the hydrological model with only rain gauge or satellite data, it is also tested a possible combination of them.

5.1. CETEMPS Hydrological Model (CHyM)

The Cetemps Hydrological Model (Coppola et al., 2007, Verdecchia et al., 2008b) has been applied for climatological studies to investigate the effects of Climate Changes on the hydrological cycle (Coppola et al., 2014, Sangelantoni et al., 2019). The model mainly has been developed and used as an operational tool for early warning system (Tomassetti et al., 2005; Ferretti et al., 2019; Colaiuda et al., 2020; Lombardi et al., 2021).

CHyM is a fully distributed, physical-based hydrological model, where hydrological processes (surface runoff, infiltration, evapotranspiration, percolation, melting and return flow) are explicitly simulated by a physical based numerical scheme. Cellular Automata (CA) based algorithm has been developed and implemented in the CHyM code. This native algorithm allows the model to simulate the hydrological cycle on any geographic domain, with any spatial resolution up to DEM resolution (90 meters in the current version). The lower limit in choosing the spatial resolution deals with the validity of the numerical schemes used to simulate the hydrological processes (e.g.: the kinematic wave of shallow water, which used to solve the continuity equation, is considered a good approximation with a horizontal resolution of few hundreds of meters). Furthermore, the CA algorithm can acquire different data sources or rebuilt the spatial distribution of precipitation on a hydrological scale.

5.2. Simulated Flow Discharge

In this study, the hydrological simulation is stands from 1st January 2015 to 31st December 2019: 1826 days, 43824 hours. For the simulations, no calibration was done and neither melting component nor the anthropic impact are considered. This condition could affect the estimate of the baseflow discharge. Obviously, this is to the detriment of the analysis as the scores are influenced by a systematic error due to a baseflow not well reproduced. However, this method does not affect the conclusions as the purpose of the work is not to validate

the performance of the hydrological model but to establish whether the performance improves based on the rain scenarios with which the model is forced. In detail, this report shows the workflow and results relating to three different domains (Fig. 2) selected on the Italian territory (Table. 4). For each domain, hydrographic basins have been identified for which the observed data are available, useful for the analyzes. Long time series of observed flow discharge data have been made available by the CIMA Foundation[1].

Furthermore, a hydrographic basin has been identified, the Tanaro river, a right tributary of Po River, for which different flood periods have been selected. Four case studies were selected, linked to four different flow peaks (Table. 7); once the time of maximum peak was identified, a 120-hour time series containing it was selected, trying to avoid time series with NO DATA (-9999.0). Anyway, if some value of -9999.0 is present, this is adjusted thanks to Cellular Automata technique, using the previous and next value.

In this report, the results about four different hydrological simulations were reported: 5 different used rain input settings for DOMAIN01 and 4 for DOMAIN02 and DOMAIN03 (Table. 3), based on data assimilation of:

- ✓ Rain gauge data: Dewetra Platform (every hour).
- ✓ H SAF Product: MW/IR H03B product (mm/h at 5 km spatial resolution and 15 min temporal resolution).
- ✓ H SAF Product: H SAF MW-only precipitation products H68 (mm/h at 27.7 km spatial resolution and 30 min temporal resolution).

Table. 3. Rain input settings

	DATA	NAME SIMULATION	GAUGE RADIUS (km)	SAT RADIUS (km)
1	Rain Gauge	GAUGE	35	X
2	MW/IR H03B	SATH03B	X	5
3	Gauge + MW/IR H03B	G3SATH03B	3	5
4	Gauge + MW/IR H03B	G5SATH03B	5	5
5	MW-only precipitation products H68	SATH68	X	28

In any case, even when the hydrological model is forced with satellite data only, if these are not available, the model uses rain gauge data with a default radius of influence.

Table 4 shows the results related to the size of the surface covered by the rain gauge data according to the use of the two selected radiuses of influence: 5 km and 3 km. The use of two radiuses is aimed at a sensitivity analysis. For example, in the case of the mountain area of the Po basin, this area covers 44% of the entire Po

[1] https://www.cimafoundation.org/

basin. Regarding this area, 71% coverage of the rain gauge data is expected if the used radius of influence is equal to 5 km and 19% if the radius of influence is equal to 3 km.

The distribution and density of the rain gauge data, also as a function of the radius of influence, are shown in fig. 6, for DOMAIN01.

Table. 4. Percentage of Covered Surface size by the rain gauge data.

Domain	Basin	Covered Basin Area %		Montain area			Hilly area			Flat area		
		R = 5 km	R = 3 km	Area%	R = 5 km	R = 3 km	Area%	R = 5 km	R = 3 km	Area%	R = 5 km	R = 3 km
01	Po	0.56	0.15	0.44	0.71	0.19	0.23	0.77	0.19	0.33	0.39	0.09
02	Tevere	0.57	0.24	0.30	0.45	0.14	0.39	0.64	0.27	0.31	0.80	0.33
03	Volturno	0.40	0.19	0.27	0.29	0.15	0.18	0.42	0.27	0.16	0.52	0.19

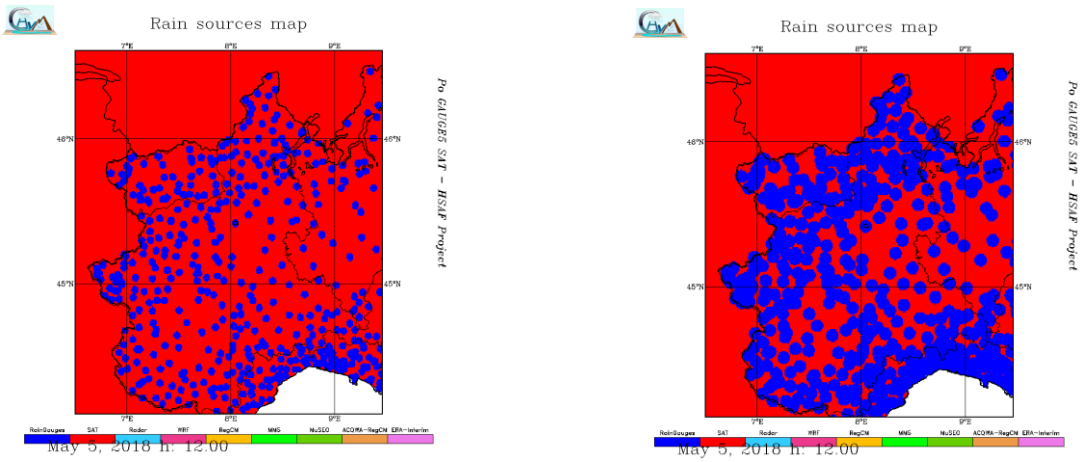


Figure 6. DOMAIN01: distribution and density of the rain gauge. The coverage rain gauge data (blue area) is a function of the radius of influence: left panel, RADIUS = 3 km; right panel, RADIUS = 5 km. Red area represents the Satellite data coverage.

5.3. Objective Quality Scores

To evaluate the fit between observed and simulated stream flow time series, the objective functions are selected (Table. 5). The traditional performance indicators have been used, such as the Nash–Sutcliffe Efficiency (Nash and Sutcliffe, 1970), percentage bias (PBIAS) measuring the average tendency of the simulated values to be larger or smaller than the observed ones. The optimal value of PBIAS is 0.0, with low-magnitude values indicating accurate model simulation. Furthermore, the following scores were considered: Root Mean Square Error (RMSE); Mean Absolute Relative Error (MARE) is sensitive to extreme values (i.e., outliers) and to low values; Original Kling-Gupta Efficiency (KGE, Gupta et al., 2009), Modified Kling-Gupta Efficiency

(kgeprime, Kling et al., 2012), Non-Parametric Kling-Gupta Efficiency (kgenp, Pool et al., 2012). According Mathevet et al. (2006), KGE and NSE can be calculated in a bounded version: Bounded Nash-Sutcliffe Efficiency (nse_c2m), Bounded Original Kling-Gupta Efficiency (kge_c2m), Bounded Modified Kling-Gupta Efficiency (kgeprime_c2m), Bounded Non-Parametric Kling-Gupta Efficiency (kgenp_c2m). The analysis is carried out using an open-source evaluator for streamflow time series in Python (Hallouin, 2019). In addition to the conventional scores, other indicators were selected to obtain a more objective analysis, independent of the limits of the scores commonly used for hydrological analyses, for a total of 17 quality scores (Table. 5).

Table. 5. Quality scores

QUALITY SCORES		
1	Nash-Sutcliffe Efficiency	nse
2	Original Kling-Gupta Efficiency	kge
3	Modified Kling-Gupta Efficiency	kgeprime
4	Non-Parametric Kling-Gupta Efficiency	kgenp
5	Root Mean square Error	RMSE
6	Mean Absolute Relative Error	MARE
7	Percent Bias	PBIAS
8	Bounded Nash-Sutcliffe Efficiency	nse_c2m
9	Bounded Original Kling-Gupta Efficiency	kge_c2m
10	Bounded Modified Kling-Gupta Efficiency	kgeprime_c2m
11	Bounded Non-Parametric Kling-Gupta Efficiency	kgenp_c2m
12	Match Correlation	MC
13	Correlation Time Delay	CT_D
14	Time Peak Delay	TP_D
15	ERROR %	ER%
16	Dynamic Time Warping	DTW
17	Derivative Dynamic Time Warping	DDTW

The idea is to consider the river flow discharge profile as a signal and for this reason, indicators, commonly used in generic signal studies, have been used.

The Match Correlation (MC) is the relationship between the Auto-correlation curve and the Cross-Correlation curve (Observe VS Simulate) and allows to understand the two curves overlap: the best value obtained will be close to 1.

$$MC = \frac{\int \text{AutoCorrelation\_of\_Observed\_values}}{\int \text{CrossCorrelation\_of\_Index\_values\_VS\_Observed\_values}} \quad (5)$$

The cross correlation (CC) is typically used in the signal theory (Rabiner and Gold, 1975; Rabiner and Schafer, 1978; Benesty et al., 2004), for the assessment of similarity between two signals.

The Correlation Time Delay (CT_D, Lombardi et al., 2021) represents the value of time lag L that maximizes the product obtained in CC estimation. CT_D represents an estimation of time shift between two series; therefore, the score is to be suitable to measure the effectiveness of the signal given by the hydrological simulations.

The Time Peak Delay (TP_D) is a timing score and represents the hourly delay of the estimated maximum peak flow discharge compared to the observed one.

The percentage Error (E%) at the peak value of the flow discharge was calculated as follows:

$$\frac{\max D_{Sim} - \max D_{Obs}}{\max D_{Obs}} \tag{6}$$

where D_{sim} indicates the simulated flow discharge and D_{Obs} represents the observed flow discharge peak.

The Dynamic Time Warping (DTW, Berndt and Clifford, 1994; Keogh and Ratanamahatana, 2005; Maier-Gerber et al., 2019 and Di Muzio et al., 2019) find the similarity between two sequences based on the "warping" the time axis of one (or both) sequences, to achieve a better alignment. Given two discrete series *x(i)* and *y(j)* of *N* and *M* components respectively, an N-by-M matrix is built. Each element *V(i,j)* represents the Euclidean distance between the i-*th* element of the first sequence and j-*th* element of the second sequence (Fig. 7). For this matrix, a "warping" path *W* is defined as a contiguous set of *L* matrix elements, and the measure of misalignment *d* for the path *W* is given by:

$$d(W) = \frac{\sum_{i,j} V(i,j)}{\frac{1}{2}L(L-1)} \tag{7}$$

where the sum in the numerator is carried out over all the elements belonging to the warping path *W*. The denominator is used to normalize different length sequences. The DTW index is then calculated as the minimum value of *d(W)*, considering all the possible path *W*.

$$DTW = \min_{W} \; d(W) \tag{8}$$

For instance, if the two considered sequences are aligned and have the same number of components (*N=M*), the optimal path will be the *N* diagonal elements of matrix *V*.

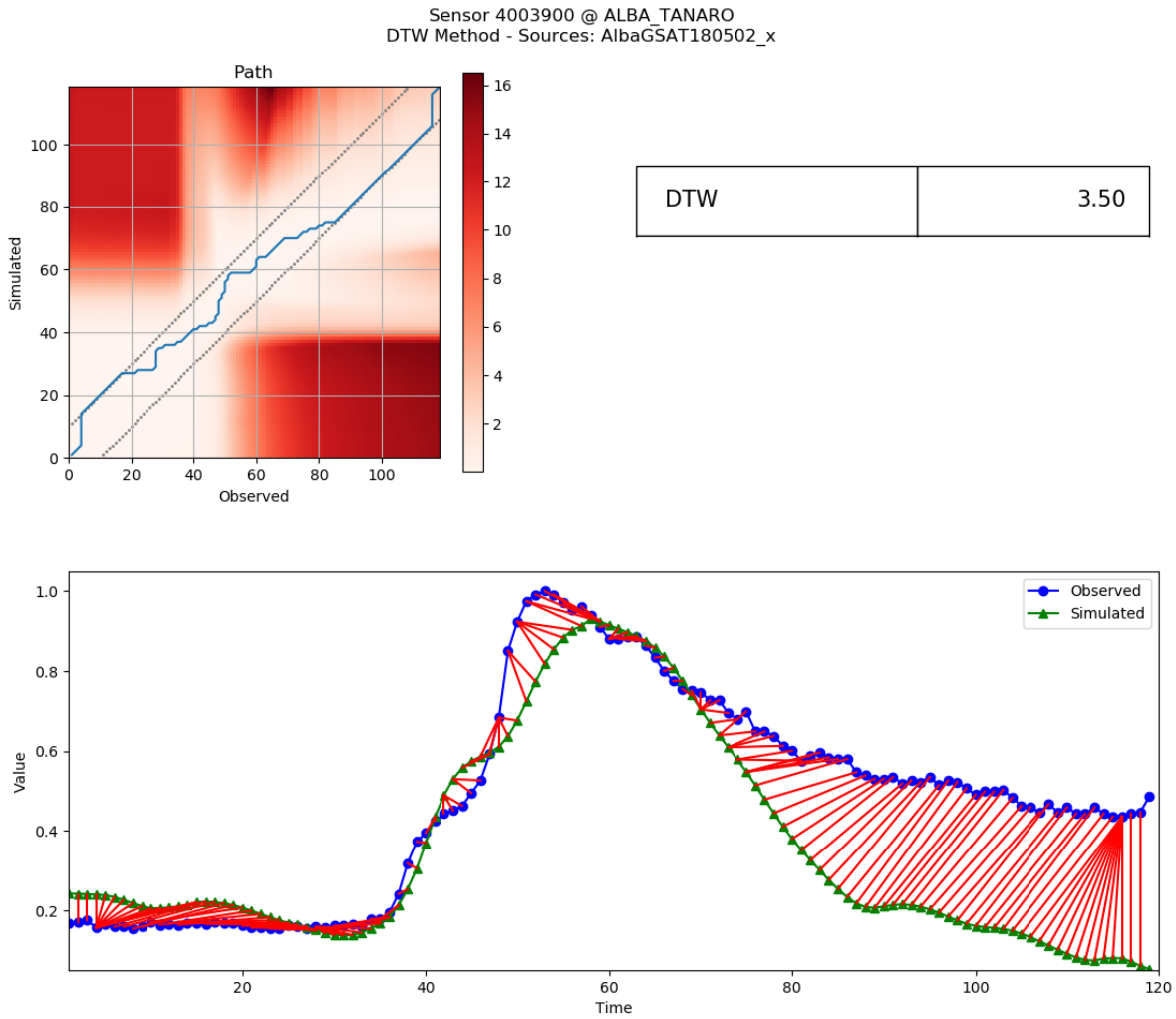


Figure. 7. DTW score related to CS02 (02/05/18 00 - 06/05/18 23) G5SATH03B hydrological simulation at Alba Station, Tanaro River.

The DTW has been successfully used, anyway it may lead to wrong results; as an example, the technique may fail in finding the optimal alignment because a feature (i.e., peak, or local minimum) in one sequence is higher or lower than its corresponding feature in the other sequence.

To overcome this problem, Keogh and Pazzani (2001) proposed the computation of warping using the local derivative of the time series to be compared and called this algorithm "Derivative Dynamic Time Warping" (DDTW).

The DDTW algorithm implementation replaces the data time series with their first derivative and the Euclidean distance is measured on them (Fig. 8). The first derivative has been calculated for each time series as follows

$$D(x[i]) = \frac{(x[i]-x[i-1])+((x[i+1]-x[i-1])/2}{2} \quad (9)$$

The main limitation linked to both analyses are defined singularities (Sakoe, & Chiba 1978; Keogh and Pazzani, 2001), i.e., the algorithm may try to explain variability in the Y-axis by warping the X-axis. This can

lead to unintuitive alignments where a single point on one time series maps onto a large subsection of another time series.

To overcome those limits, we used the Windowing method (Berndt and Clifford, 1994). Allowable elements of the matrix can be restricted to those that fall into a warping window according to the follow equation:

$$|i - (n/(m/j))| < Ra \tag{10}$$

where Ra is a positive integer window width. In this work, Ra is equal to 10 and this allows us to mitigate the effects of the baseflow discharge that the model does not accurately reproduce, due to anthropic impact or for example melting effect.

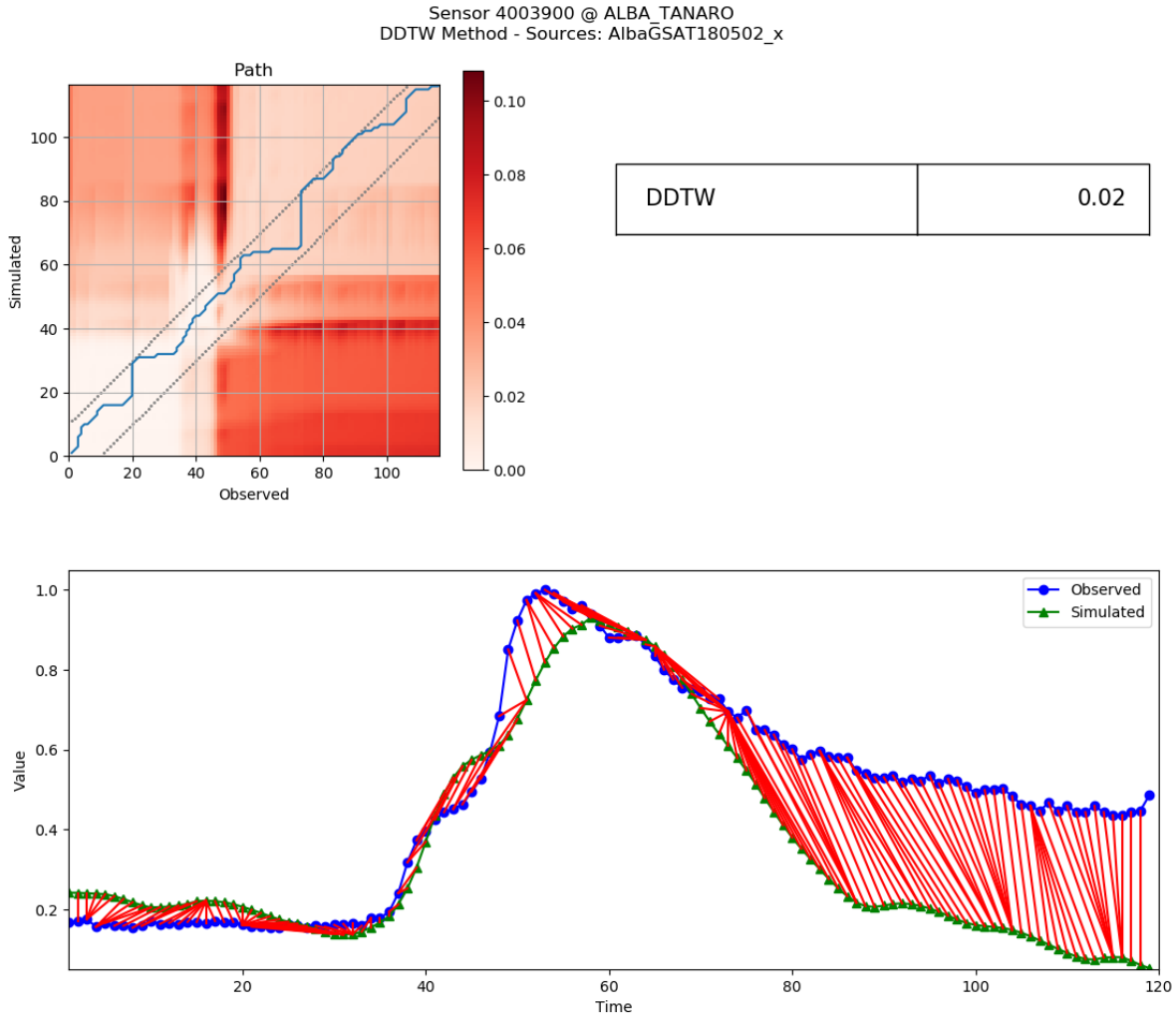


Figure. 8. DDTW score related to CS02 (02/05/18 00 - 06/05/18 23) G5SAT hydrological simulation at Alba Station, Tanaro River.

6. Results

According to the project supervisors, it was decided to use five-year long hydrological simulations for each domain and for each identified scenario. The analysis is carried out on 4 years of data, since the first year serves as a spin up of the model. The main limits in the hydrological assessment are also linked to the quality of the flow discharge observed data. Figure 9 shows the time series of two consecutive stations in the basin of the Calore river, a tributary of the Volturno river. The time series released by the Benevento station on the Calore river shows several significant peaks not evident in the next Solopaca station. In this circumstance, it seems reasonable to consider those peak values as a measurement error and therefore they have been removed. This circumstance was also highlighted in the other used series, but in other cases it is not so evident: insignificant peaks recorded at the previous sensor not found in the next one. In this case, it was decided not to filter the data, but to overcome these limitations, two long observed time series have been used on the same basin to reduce the uncertainty associated with the observed data (Table. 1).

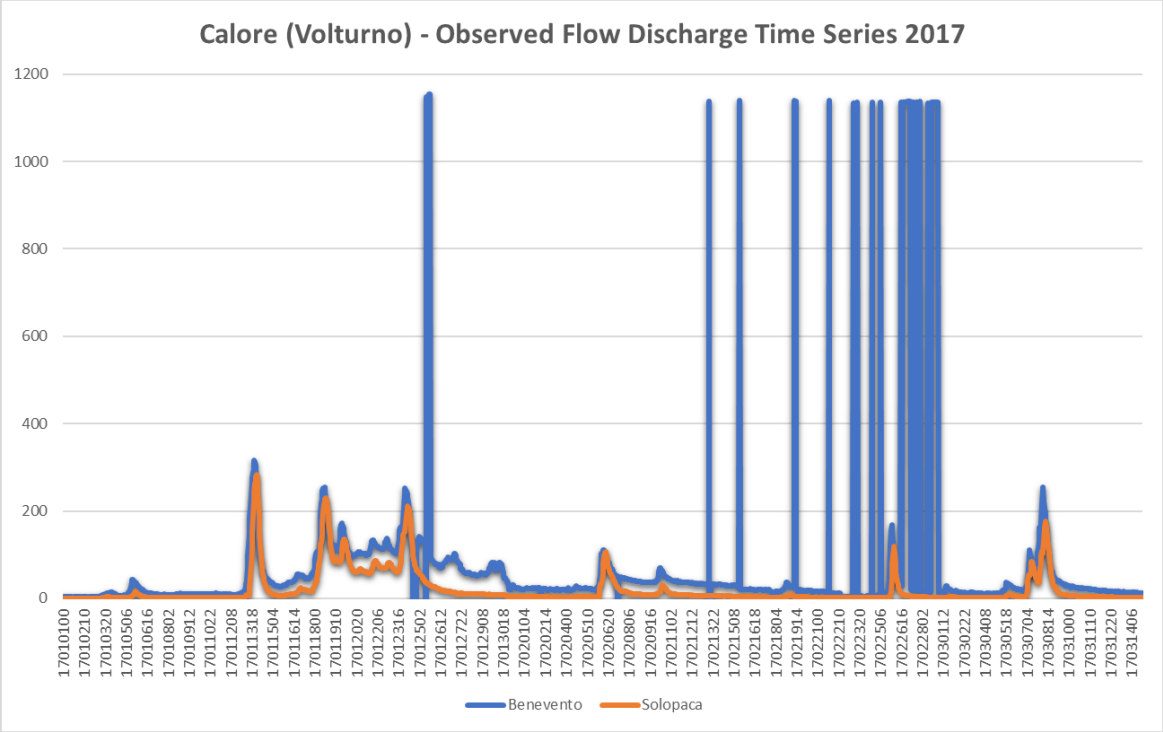


Figure 9. Observed flow discharge time series in the Calore basin (Volturno) relating to two consecutive measurement stations, Benevento and Solopaca, from January 1, 2017, to March 15, 2017.

A second strategy was to use different quality scores, up to 17 which should give more objective information on the results of the analyzes. A traditional metric used in hydrology to summarize model performance is, for example, the Nash-Sutcliffe Efficiency (NSE) or the Kling-Gupta Efficiency (KGE). A score value = 0 corresponds to using the mean flow as a benchmark predictor: negative values are often viewed in the literature as bad model performance and positive values are seen as good model performance. However, the

interpretation of the quality score is closely related to the purpose of the study. For example, in the case of the KGE, in the literature, Knoben et al., (2019) suggest that the mean flow as a predictor does not result in KGE = 0, but instead KGE = $1-\sqrt{2} \approx -0.41$. Thus, KGE values greater than -0.41 indicate that a model improves upon the mean flow benchmark, even if the model's KGE value is negative. Furthermore, the different values of the different scores cannot be directly compared because their relationship is not unique. As in the case of the PBIAS which measures the average tendency of the simulated values to be greater or less than those observed with respect to RMSE, and MARE is sensitive to extreme values (outliers) and to low values. Furthermore, for the analysis it was decided to also select specific case studies, relating to domain 1. For these comparisons, an unconventional metric was used considering the river flow profile as a signal and for this reason the indicators of generic signal studies have been used.

6.1. Annual Hydrological Analisys

The hydrological model has been forced with different precipitation scenarios, according to the Table 3: GAUGE, SATH03B, G3SATH03B, G5SATH03B and for DOMAIN01, the simulation has been carried out also forcing the model with the satellite data SATH68.

6.1.1. DOMAIN01 – Po River Basin

The comparison between the different time series for the year 2019 are showed in Fig. 10. For this type of comparison, an intuitive analysis is not immediate, for this reason the graphical representations are shown. Figure 11 represents the graphs relating to the results of some of the selected quality scores, which are shown as an example.

From a hydrological point of view, the results are not always satisfactory, for example negative KGE values are not always accepted by the scientific community (Knoben et al., 2019). It is necessary, however, not to forget that the hydrological model was not voluntarily calibrated ad hoc for this study, as it was used as a pure analysis tool. Furthermore, given that the other components of the hydrological balance, such as melting and anthropogenic presence, have not been considered, an analysis over a long time series can distort the result. This mainly depends on the base flow of the river which is certainly not well simulated. For this reason, an analysis of the case studies representing periods of flood waves have also been selected and described in the following section. The purpose of this study is to demonstrate that the use of the EUMETSAT satellite data improves the performance of the hydrological model compared to the use of the rain gauge data alone, which has always been considered the most reliable product.

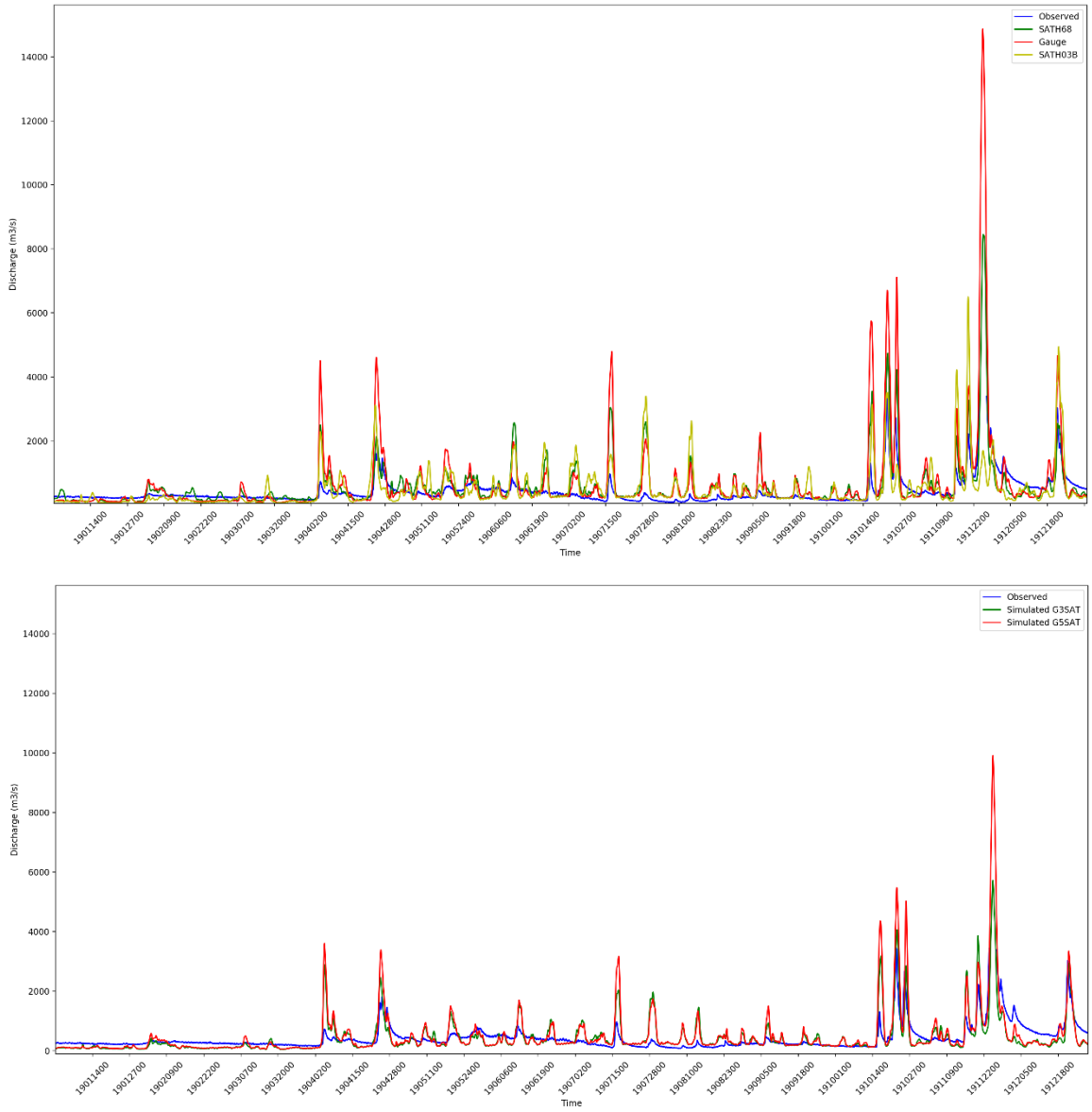
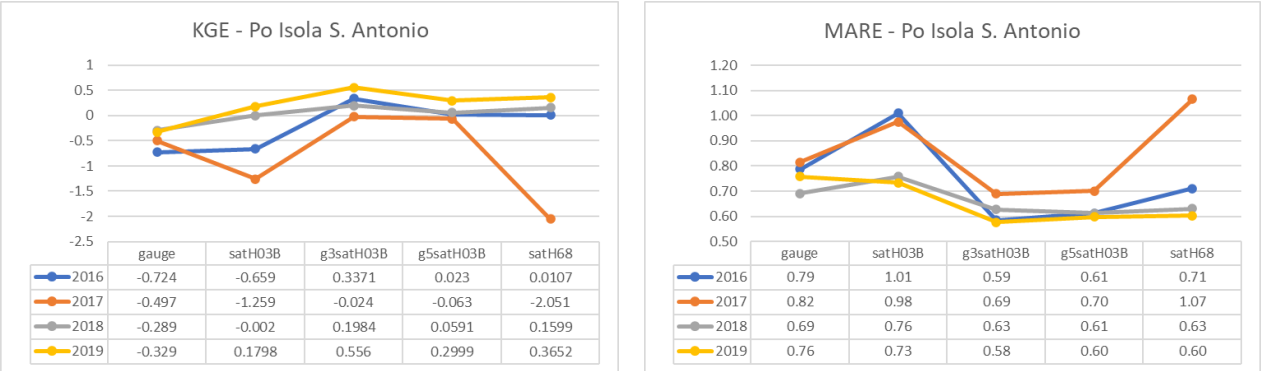


Figure 10. Flow discharge time series comparison at Isola S. Antonio Station on Po River Basin for the year 2019.



	gauge	satH03B	g3satH03B	g5satH03B	satH68
2016	-0.724	-0.659	0.3371	0.023	0.0107
2017	-0.497	-1.259	-0.024	-0.063	-2.051
2018	-0.289	-0.002	0.1984	0.0591	0.1599
2019	-0.329	0.1798	0.556	0.2999	0.3652

	gauge	satH03B	g3satH03B	g5satH03B	satH68
2016	0.79	1.01	0.59	0.61	0.71
2017	0.82	0.98	0.69	0.70	1.07
2018	0.69	0.76	0.63	0.61	0.63
2019	0.76	0.73	0.58	0.60	0.60

Figure 11. Quality Scores (KGE and MARE) related to Po – Isola S. Antonio Station.

In order not to divert attention from the purpose of this work, it was decided to report the results of the quality scores as a percentage value with respect to the reference simulation, where the model has been forced with only the rain gauge data (fig. 12), according to the following equation:

$$\frac{score_{GAUGE} - score_{SCEN}}{|score_{GAUGE}|} \tag{11}$$

where $score_{SCEN}$ refers to the scores of each single scenario. The trend of the percentages depends on the type of score. For example, in the case of KGE, where values greater than 0 indicate better performance of the model, the positive percentages indicate that the rain field is better represented than the one rebuilt with only the rain gauges. The results indicate a marked improvement for the year 2019, up to 269% for the g3satH03B simulation, 155% for the satH03B, 191% for the g5satH03B and 211% for the satH68 compared to KGE = -0.33 of the GAUGE simulation (Table 6).

In the case of 2017, the good results obtained from the merging of rain gauge and satellite data were verified, although an improvement in performance using only satellite data is not confirmed, as in the case of 2019. This evidence has been carried out in the analyzed data relating to the other domains. In the case of the Mean Absolute Relative Error (MARE), the best performances are associated with values close to zero. This means that an improvement in performance is associated with negative percentages.

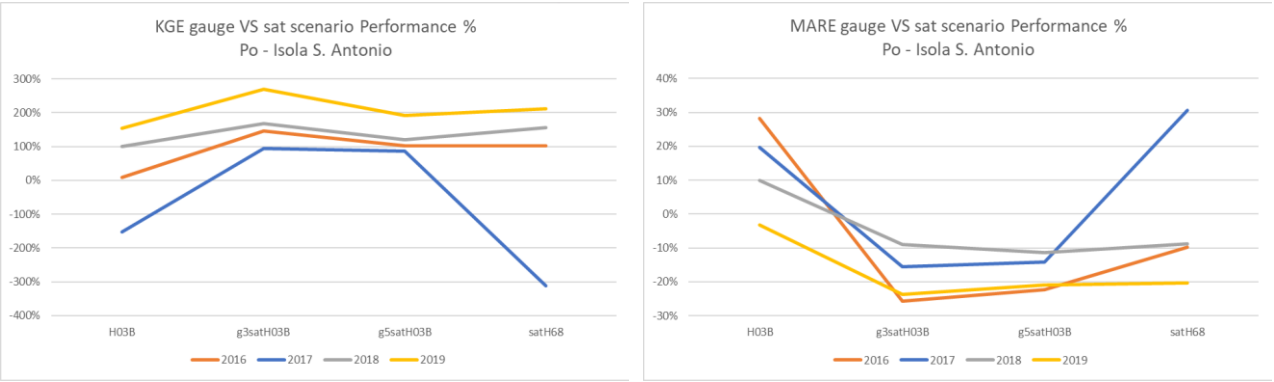


Figure 12. Quality Score percentage value with respect to the GAUGE reference simulation.

Although MARE is a score that describes different characteristics of the time series comparison, the performance results are comparable with those of the KGE. The same result is also obtained by observing the other quality scores with the same approach. Performance is improved by using the NSE and the RMSE (Figure 13). The complete table, with all the quality scores is reported.

Figure 13. Quality Score percentage value with respect to the GAUGE reference simulation.

Table 6. Quality Score relate to Po – Isola S. Antonio.

Year	NAME SIM	KGE	NSE	RMSE	KGE prime	KGE enp	MARE	PBIAS	NSE cm2	KGE c2m	KGE prime_c2m	KGE enp_c2m
2016	GAUGE	-0.72	-3.42	620.75	-0.04	0.34	0.79	-38.28	-0.63	-0.27	-0.02	0.20
	SATH03B	-0.66	-4.51	693.38	-0.01	0.08	1.01	-72.29	-0.69	-0.25	0.00	0.04
	G3SATH03B	0.34	-0.43	353.69	0.47	0.47	0.59	-12.83	-0.18	0.20	0.31	0.31
	G5SATH03B	0.02	-0.99	416.66	0.27	0.45	0.61	-16.31	-0.33	0.01	0.16	0.29
	SATH68	0.01	-1.22	439.83	0.38	0.35	0.71	-43.15	-0.38	0.01	0.24	0.21
2017	GAUGE	-0.50	-3.65	372.71	0.01	0.26	0.82	-34.24	-0.65	-0.20	0.01	0.15
	SATH03B	-1.26	-8.36	528.88	-0.28	0.04	0.98	-60.72	-0.81	-0.39	-0.12	0.02
	G3SATH03B	-0.02	-1.90	294.35	0.14	0.28	0.69	-13.09	-0.49	-0.01	0.07	0.16
	G5SATH03B	-0.06	-1.94	296.68	0.14	0.30	0.70	-15.09	-0.49	-0.03	0.07	0.18
	SATH68	-2.05	-14.98	691.22	-0.64	-0.06	1.07	-75.52	-0.88	-0.51	-0.24	-0.03
2018	GAUGE	-0.29	-2.13	825.20	0.12	0.42	0.69	-25.42	-0.52	-0.13	0.06	0.27
	SATH03B	0.00	-1.75	773.58	0.03	0.33	0.76	-2.36	-0.47	0.00	0.02	0.20
	G3SATH03B	0.20	-0.97	654.99	0.02	0.40	0.63	11.46	-0.33	0.11	0.01	0.25
	G5SATH03B	0.06	-1.21	692.85	0.08	0.45	0.61	-1.51	-0.38	0.03	0.04	0.29
	SATH68	0.16	-0.94	649.08	0.40	0.43	0.63	-23.66	-0.32	0.09	0.25	0.27
2019	GAUGE	-0.33	-1.89	657.43	0.25	0.37	0.76	-43.33	-0.49	-0.14	0.14	0.23
	SATH03B	0.18	-0.95	539.81	0.37	0.38	0.73	-17.27	-0.32	0.10	0.22	0.24
	G3SATH03B	0.56	0.06	375.97	0.54	0.47	0.58	1.85	0.03	0.39	0.37	0.30
	G5SATH03B	0.30	-0.34	448.02	0.45	0.49	0.60	-11.90	-0.15	0.18	0.29	0.33
	SATH68	0.37	-0.25	432.34	0.55	0.47	0.59	-28.72	-0.11	0.22	0.38	0.31

To confirm what was observed, the same analysis was made for the data relating to the Tanaro station in Alba. Also, for these comparisons it is evident that an improvement in the performance of the model using the merged gauge-sat data, in particular the simulation G3SATH03B.

Figure 14. Quality Score percentage value with respect to the GAUGE reference simulation.

Table 7. Quality Score relate to Tanaro - Alba.

Year	NAME SIM	KGE	NSE	RMSE	KGE prime	KGE enp	MARE	PBIAS	NSE cm2	KGE c2m	KGE prime_c2m	KGE enp_c2m
2016	GAUGE	-0.157	-0.903	172.3	0.2592	0.275	1.1037	-26.37	-0.311	-0.073	0.1489	0.1594
	SATH03B	0.55	0.1432	115.61	0.5334	0.2559	1.0469	-17.32	0.0771	0.3793	0.3637	0.1467
	G3SATH03B	0.7315	0.5784	81.097	0.6368	0.3278	0.7682	17.482	0.4069	0.5767	0.4672	0.196
	G5SATH03B	0.4346	0.1641	114.19	0.4036	0.3514	0.8743	2.0777	0.0894	0.2776	0.2528	0.2132
	SATH68	0.5273	0.3586	100.02	0.6143	0.3417	0.832	-7.048	0.2185	0.3581	0.4434	0.2061
2017	GAUGE	-1.143	-7.173	68.633	-1.196	0.3238	1.0888	1.8043	-0.781	-0.363	-0.374	0.1931
	SATH03B	-2.089	-14.35	94.079	-1.714	0.1979	1.1725	-10.89	-0.877	-0.51	-0.461	0.1098
	G3SATH03B	-0.444	-3.865	52.956	-1.208	0.184	0.9237	27.786	-0.659	-0.181	-0.376	0.1013
	G5SATH03B	-0.597	-4.442	56.005	-1.182	0.2548	0.9696	20.682	-0.689	-0.23	-0.371	0.146
	SATH68	-2.755	-20.47	111.25	-1.699	0.213	1.2203	-31.19	-0.911	-0.579	-0.459	0.1192
2018	GAUGE	0.0824	-1.106	138.62	-0.184	0.3835	0.7903	13.776	-0.356	0.043	-0.084	0.2372
	SATH03B	0.0538	-1.488	150.65	-0.699	0.1489	0.9458	40.265	-0.426	0.0276	-0.259	0.0804
	G3SATH03B	0.2541	-0.538	118.46	-0.528	0.1956	0.8014	48.731	-0.212	0.1455	-0.208	0.1084
	G5SATH03B	0.3079	-0.491	116.62	-0.316	0.2947	0.737	35.181	-0.197	0.1819	-0.136	0.1728
	SATH68	0.3246	-0.595	120.64	0.0301	0.3521	0.7474	22.113	-0.229	0.1938	0.0153	0.2136
2019	GAUGE	-0.584	-3	122.14	-0.52	0.3945	0.9376	-2.614	-0.6	-0.226	-0.206	0.2457
	SATH03B	0.0798	-1.384	94.299	-0.522	0.2788	0.8451	29.611	-0.409	0.0415	-0.207	0.162
	G3SATH03B	0.3299	-0.394	72.123	-0.397	0.2649	0.7437	40.788	-0.164	0.1975	-0.165	0.1526
	G5SATH03B	0.0393	-1.094	88.379	-0.513	0.3464	0.796	24.035	-0.353	0.02	-0.204	0.2095
	SATH68	0.1609	-0.775	81.377	-0.149	0.391	0.7412	16.092	-0.279	0.0874	-0.069	0.243

6.1.2 DOMAIN02 – Tevere River Basin

The same considerations were made for DOMAIN02. Also, in this case the G3SATH03B simulation seems to be more performing than the others in most of the years simulated for the data related to the sensor in Ponte Felcino.

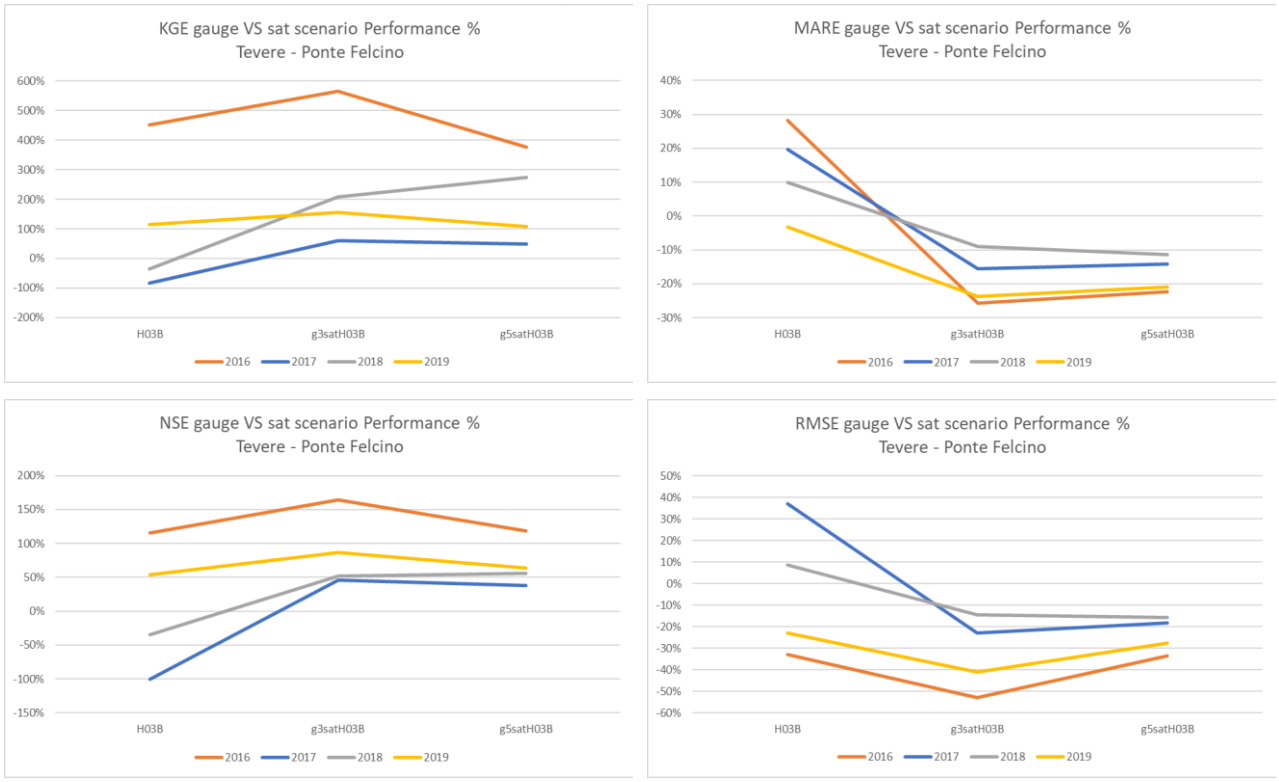


Figure 15. Quality Score percentage value with respect to the GAUGE reference simulation.

Table 8. Quality Score relate to Tevere – Ponte Felcino.

Year	NAME SIM	KGE	NSE	RMSE	KGE prime	KGE enp	MARE	PBIAS	NSE cm2	KGE c2m	KGE prime_c2m	KGE enp_c2m
2016	GAUGE	-0.157	-0.903	172.3	-0.035	0.3366	0.7877	-38.28	-0.63	-0.265	-0.017	0.2023
	SATH03B	0.55	0.1432	115.61	-0.008	0.0751	1.0104	-72.29	-0.692	-0.247	-0.004	0.039
	G3SATH03B	0.7315	0.5784	81.097	0.47	0.4677	0.5853	-12.83	-0.178	0.2027	0.3072	0.3052
	G5SATH03B	0.4346	0.1641	114.19	0.2707	0.4484	0.6121	-16.31	-0.33	0.0116	0.1565	0.289
2017	GAUGE	-1.143	-7.173	68.633	0.0146	0.2616	0.8159	-34.24	-0.645	-0.199	0.0073	0.1504
	SATH03B	-2.089	-14.35	94.079	-0.284	0.0403	0.9762	-60.72	-0.806	-0.386	-0.124	0.0205
	G3SATH03B	-0.444	-3.865	52.956	0.135	0.2815	0.689	-13.09	-0.487	-0.012	0.0724	0.1638
	G5SATH03B	-0.597	-4.442	56.005	0.1369	0.3007	0.7005	-15.09	-0.492	-0.03	0.0735	0.177
2018	GAUGE	0.0824	-1.106	138.62	0.1189	0.4194	0.6903	-25.42	-0.516	-0.126	0.0632	0.2653
	SATH03B	0.0538	-1.488	150.65	0.0318	0.3301	0.7588	-2.357	-0.467	-0.001	0.0161	0.1976
	G3SATH03B	0.2541	-0.538	118.46	0.022	0.398	0.6277	11.456	-0.327	0.1101	0.0111	0.2485
	G5SATH03B	0.3079	-0.491	116.62	0.0839	0.4509	0.612	-1.51	-0.376	0.0304	0.0438	0.2911
2019	GAUGE	-0.584	-3	122.14	0.249	0.3719	0.7573	-43.33	-0.485	-0.141	0.1422	0.2284
	SATH03B	0.0798	-1.384	94.299	0.3662	0.3845	0.7335	-17.27	-0.32	0.0988	0.2241	0.238
	G3SATH03B	0.3299	-0.394	72.123	0.5394	0.4656	0.5771	1.8464	0.0289	0.3851	0.3693	0.3034
	G5SATH03B	0.0393	-1.094	88.379	0.45	0.4949	0.5982	-11.9	-0.145	0.1764	0.2903	0.3288

An improvement in performance is also recorded for the Pierantonio sensor, although they are better for the G5SATH03B simulations. In figure 16 in the graph of the KGE relating to the curve of the year 2018 there is a marked percentage variation of the various precipitation scenarios with respect to the GAUGE simulation whose KGE value is equal to 0.0241, compared to -0.313 of the SATH03B and 0.4157 of G5SATH03B.

Figure 16. Quality Score percentage value with respect to the GAUGE reference simulation.

Table 9. Quality Score relate to Tevere – Pierantonio.

Year	NAME SIM	KGE	NSE	RMSE	KGE prime	KGE enp	MARE	PBIAS	NSE cm2	KGE c2m	KGE prime_c2m	KGE enp_c2m
2016	GAUGE	-0.586	-3.183	53.877	-0.025	0.3566	1.2331	-32.06	-0.614	-0.226	-0.012	0.217
	SATH03B	-1.988	-12.65	97.33	-0.532	-0.278	2.1423	-91.83	-0.863	-0.498	-0.21	-0.122
	G3SATH03B	-0.693	-4.929	64.139	-0.229	0.1872	1.5116	-28.32	-0.711	-0.257	-0.102	0.1032
	G5SATH03B	-0.08	-1.939	45.16	0.0409	0.3392	1.1603	-7.986	-0.492	-0.038	0.0208	0.2043
2017	GAUGE	-1.888	-10.26	45.945	-0.401	-0.064	1.9481	-88.23	-0.836	-0.485	-0.167	-0.031
	SATH03B	-5.262	-47.65	95.476	-1.476	-0.842	2.9347	-166.5	-0.959	-0.724	-0.424	-0.296
	G3SATH03B	-2.658	-19.33	61.716	-1.408	0.1497	1.7509	-41.08	-0.906	-0.57	-0.413	0.0809
	G5SATH03B	-1.349	-8.508	42.206	-0.592	0.2574	1.5426	-34.85	-0.809	-0.402	-0.228	0.1477
2018	GAUGE	0.0241	-1.348	58.06	0.2122	0.5172	1.1003	-13.25	-0.402	0.0122	0.1186	0.3488
	SATH03B	-0.313	-3.37	79.208	-0.223	0.2438	1.5064	-6.259	-0.627	-0.135	-0.1	0.1388
	G3SATH03B	0.1616	-1.042	54.147	-0.083	0.2818	1.1791	25.288	-0.342	0.0879	-0.04	0.164
	G5SATH03B	0.4157	-0.218	41.818	0.2372	0.428	0.9875	23.737	-0.098	0.2624	0.1345	0.2722
2019	GAUGE	-0.601	-2.715	74.425	0.1206	0.2172	1.3941	-49.71	-0.575	-0.231	0.0641	0.1218
	SATH03B	0.1928	-0.639	49.44	-0.041	0.2367	1.1645	33.726	-0.242	0.1067	-0.02	0.1342
	G3SATH03B	0.2885	0.0509	37.613	0.1158	0.2324	0.9602	45.739	0.0261	0.1686	0.0614	0.1314
	G5SATH03B	0.5539	0.1877	34.796	0.2888	0.3748	0.8824	22.448	0.1035	0.383	0.1687	0.2306

6.1.3 DOMAIN03 – Volturno River Basin

For DOMAIN03, the hydrological simulations cannot be considered satisfactory. Quality Score values are very low compared to those observed for the other domains. However, this could also depend on the low quality of the data observed for this catchment area, as shown in Figure 9. The aspect that remains constant, however, is that despite the limitations of the hydrological model, an improvement in the performance of the model can be observed using the different rainfall scenarios.

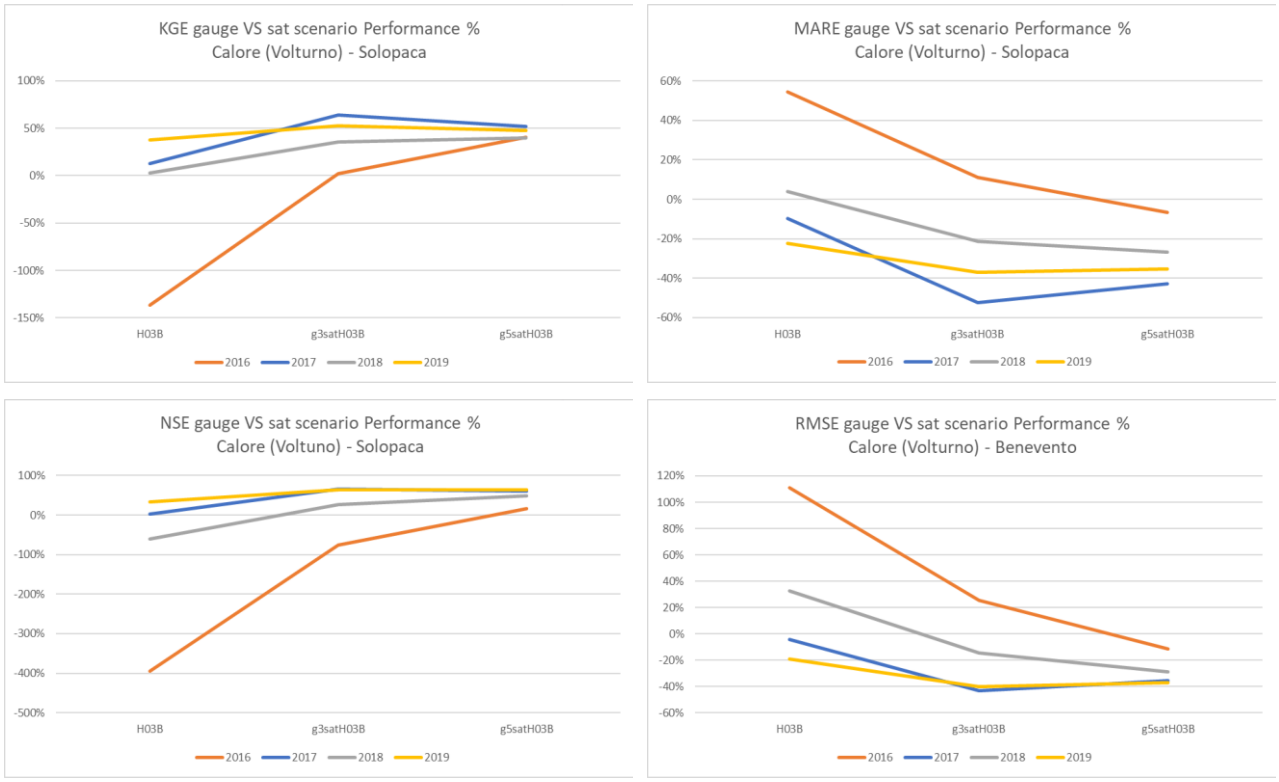


Figure 17. Quality Score percentage value with respect to the GAUGE reference simulation.

Table 9. Quality Score relate to Calore – Solopaca.

Year	NAME SIM	KGE	NSE	RMSE	KGE prime	KGE enp	MARE	PBIAS	NSE cm2	KGE c2m	KGE prime_c2m	KGE enp_c2m
2016	GAUGE	-0.782	-0.994	102.19	-0.827	-0.726	2.3935	-161.9	-0.332	-0.281	-0.292	-0.266
	SATH03B	-1.851	-4.915	175.97	-1.616	-1.599	3.6973	-240.6	-0.71	-0.48	-0.447	-0.444
	G3SATH03B	-0.768	-1.755	120.09	-0.787	-0.681	2.6593	-146.4	-0.467	-0.277	-0.282	-0.254
	G5SATH03B	-0.465	-0.833	97.961	-0.554	-0.407	2.2352	-120.4	-0.294	-0.188	-0.217	-0.169
2017	GAUGE	-6.794	-25.9	168.1	-5.597	-5.59	6.7655	-656.4	-0.928	-0.772	-0.736	-0.736
	SATH03B	-5.92	-25.39	166.49	-4.524	-4.528	6.1113	-548.5	-0.926	-0.747	-0.693	-0.693
	G3SATH03B	-2.436	-8.808	101.49	-1.704	-1.668	3.2164	-259.2	-0.814	-0.549	-0.46	-0.454
	G5SATH03B	-3.26	-10.27	108.82	-2.53	-2.514	3.8719	-345.8	-0.837	-0.619	-0.558	-0.556
2018	GAUGE	-4.421	-11.5	124.5	-3.707	-3.711	4.8521	-467.9	-0.851	-0.688	-0.649	-0.649
	SATH03B	-4.298	-18.47	155.34	-3.188	-3.181	5.0436	-410.1	-0.902	-0.682	-0.614	-0.614
	G3SATH03B	-2.837	-8.449	108.22	-2.293	-2.279	3.8223	-319.8	-0.808	-0.586	-0.534	-0.532
	G5SATH03B	-2.663	-5.979	93.006	-2.281	-2.27	3.5545	-320.9	-0.749	-0.571	-0.532	-0.531
2019	GAUGE	-6.162	-14.63	152.57	-5.435	-5.422	6.5635	-640.6	-0.879	-0.754	-0.731	-0.73
	SATH03B	-3.821	-9.842	127.04	-3.3	-3.245	5.1011	-420.9	-0.831	-0.656	-0.622	-0.618
	G3SATH03B	-2.919	-5.179	95.912	-2.655	-2.602	4.1454	-356.4	-0.721	-0.593	-0.57	-0.565
	G5SATH03B	-3.245	-5.142	95.626	-2.964	-2.93	4.2498	-389.7	-0.719	-0.618	-0.597	-0.594

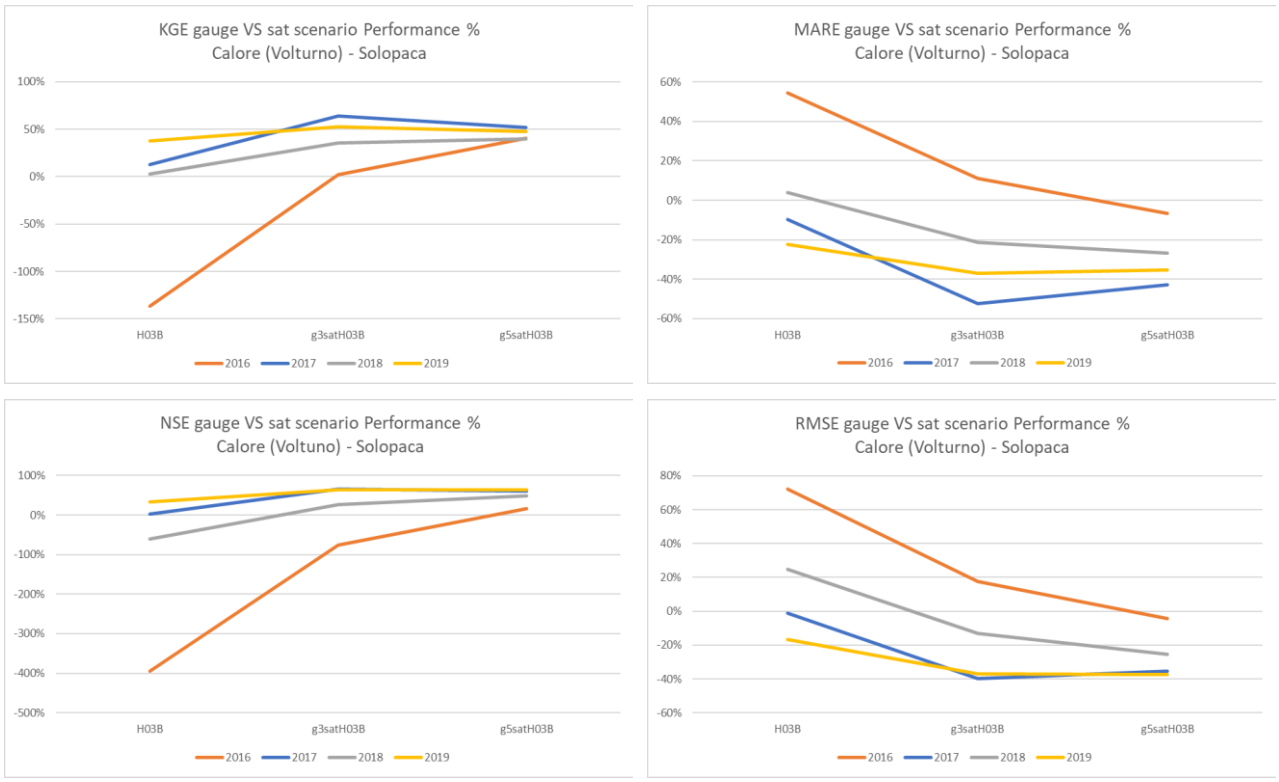


Figure 18. Quality Score percentage value with respect to the GAUGE reference simulation.

Table 10. Quality Score relate to Calore – Benevento.

Year	NAME SIM	KGE	NSE	RMSE	KGE prime	KGE enp	MARE	PBIAS	NSE cm2	KGE c2m	KGE prime_c2m	KGE enp_c2m
2016	GAUGE	-0.782	-0.994	102.19	-0.827	-0.726	2.3935	-161.9	-0.332	-0.281	-0.292	-0.266
	SATH03B	-1.851	-4.915	175.97	-1.616	-1.599	3.6973	-240.6	-0.71	-0.48	-0.447	-0.444
	G3SATH03B	-0.768	-1.755	120.09	-0.787	-0.681	2.6593	-146.4	-0.467	-0.277	-0.282	-0.254
	G5SATH03B	-0.465	-0.833	97.961	-0.554	-0.407	2.2352	-120.4	-0.294	-0.188	-0.217	-0.169
2017	GAUGE	-6.794	-25.9	168.1	-5.597	-5.59	6.7655	-656.4	-0.928	-0.772	-0.736	-0.736
	SATH03B	-5.92	-25.39	166.49	-4.524	-4.528	6.1113	-548.5	-0.926	-0.747	-0.693	-0.693
	G3SATH03B	-2.436	-8.808	101.49	-1.704	-1.668	3.2164	-259.2	-0.814	-0.549	-0.46	-0.454
	G5SATH03B	-3.26	-10.27	108.82	-2.53	-2.514	3.8719	-345.8	-0.837	-0.619	-0.558	-0.556
2018	GAUGE	-4.421	-11.5	124.5	-3.707	-3.711	4.8521	-467.9	-0.851	-0.688	-0.649	-0.649
	SATH03B	-4.298	-18.47	155.34	-3.188	-3.181	5.0436	-410.1	-0.902	-0.682	-0.614	-0.614
	G3SATH03B	-2.837	-8.449	108.22	-2.293	-2.279	3.8223	-319.8	-0.808	-0.586	-0.534	-0.532
	G5SATH03B	-2.663	-5.979	93.006	-2.281	-2.27	3.5545	-320.9	-0.749	-0.571	-0.532	-0.531
2019	GAUGE	-6.162	-14.63	152.57	-5.435	-5.422	6.5635	-640.6	-0.879	-0.754	-0.731	-0.73
	SATH03B	-3.821	-9.842	127.04	-3.3	-3.245	5.1011	-420.9	-0.831	-0.656	-0.622	-0.618
	G3SATH03B	-2.919	-5.179	95.912	-2.655	-2.602	4.1454	-356.4	-0.721	-0.593	-0.57	-0.565
	G5SATH03B	-3.245	-5.142	95.626	-2.964	-2.93	4.2498	-389.7	-0.719	-0.618	-0.597	-0.594

Also in this case, it has been verified the improvement of the simulations where the rain area is rebuilt using merged gauge-sat data, although a clear distinction between G3SATH03B and G5SATH03B is not evident.

6.2 Case Studies on Tanaro River Basin

The analysis of long time series is affected by systematic errors due, for example, to the components of the hydrological balance not considered (melting) or also to the quality of the observed data.

According to the author's experience, it is essential to also consider case studies of severe hydrological events, where the flood wave drivers of rivers are connected to the intensity and distribution of rainfall. This approach is important for the purpose of this work: evaluate the different rain scenarios forcing the hydrological model. The results show that the approach is valid, as quality scores improve significantly. Furthermore, the analysis confirms what has been verified in the previous sections: the use of the combined gauge-sat data improves the performance of the hydrological model. The results have been carried out for the Tanaro river. In details, they are related to the Alba station, located in the middle part of the main river path. In fig. 19 a graphic representation of the density and distribution of the rain gauge is shown and the coverage of the data as a function of the radius of influence.

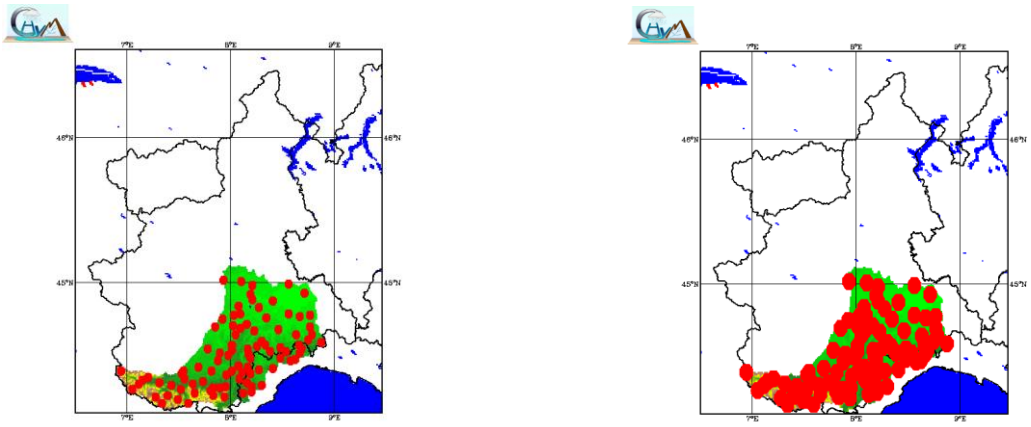


Figure 19. Tanaro Basin: distribution and density of sensors and rain gauge data coverage: left panel, RADIUS = 3 km; right panel, RADIUS = 5 km.

The Table 11 shows in more details all the information related to the considered area. For example, in the case of the mountain area of the Tanaro basin, this area covers 26% of the entire Tanaro basin. Regarding this area, 77% coverage of the rain gauge data is expected if the used radius of influence is equal to 5 km and 21% if the radius of influence is equal to 3 km.

Table 11. Percentage of covered surface size by the rain gauge data on Tanaro Basin.

Domain	Basin	Covered Basin Area %		Montain area			Hilly area			Flat area		
		R = 5 km	R = 3 km	Area%	R = 5 km	R = 3 km	Area%	R = 5 km	R = 3 km	Area%	R = 5 km	R = 3 km
01	Tanaro	0.58	0.17	0.26	0.77	0.21	0.35	0.71	0.19	0.37	0.51	0.12

Tables 13 and 14 list the 17 quality scores carried out from the comparison between observed and simulated flow discharge and they are related to the Alba station on Tanaro basin, for the four selected case studies.
The four case studies (Tab. 12) were not randomly chosen but in a strategic way, to have a continuity of the observed flow discharge data and referred to different seasonal period.
Very important evidence is that using satellite data merged with gauge data, the trend is towards improving the performance of hydrological simulations.

Tab. 12. Case Studies List

Case Study	Period
01	23/11/16 00 - 27/11/16 23
02	02/05/18 00 - 06/05/18 23
03	31/10/18 00 - 04/11/18 23
04	25/11/19 00 - 29/11/19 23

In details, referring for example to CS03, of which the related graphs to the comparison between observed and simulated data are also reported, even if the results of the NSE may not be considered good for the scientific community, the improvement of the results is highlighted in values ranging from -9.24 for simulations that use only rain gauge data to -0.24 for G3SATH03B simulation. An improvement in performance is also intuitive from the observation of the Figures 20, 21, 22 and 23. The same results are confirmed, for example, also by the Match Correlation score, which shows the best performances for the G3SATH03B simulation. The same conclusions are confirmed for the CS04, where the NSE assumes a more performing value for the G3SATH03B simulation, 0.42. The result of the simulation that uses only the satellite data is also better than the simulation with only the rain gauge data, respectively, -0.96 and -2.24, although negative results.
Also, very interesting are the results of the NSE of CS01, where the simulation with the hydrological model forced with only satellite data is associated with a higher value of NSE, 0.49, compared to -1.9 of the forced models with the gauge data. So, in this case the best performances are associated with simulations where satellite data have a greater weight: NSE for G3SATH03B is 0.50. The other scores mostly confirm the same trend. For the CS02, we have a positive NSE result associated with the simulation that uses only the rain gauge data, 0.37, which is not improved by the other simulations, but it is comparable with the G5SATH03B simulation. The performance of the model certainly improves compared to just using satellite data.
All results confirm the weight of the different data sets as a function of the area coverage due to the choice of the radius of influence. This consideration depends on the usage of the modular approach: in this way it is possible to consider the different nature of the data, assimilating the different data sets using a hierarchical sequence of modules.

Table 13. Quality Score Results

CS	NAME SIM	NSE	NSE c2m	KGE	KGE c2m	KGE prime	KGE primec2m	KGE enp	KGE enp c2m
01	GAUGE	-1.93	-0.49	-0.43	-0.18	0.07	0.04	0.09	0.05
	SATH03B	0.49	0.33	0.47	0.31	0.41	0.26	0.47	0.31
	G3SATH03B	0.50	0.33	0.72	0.57	0.66	0.49	0.79	0.66
	G5SATH03B	-0.14	-0.07	0.32	0.19	0.56	0.38	0.63	0.46
02	GAUGE	0.37	0.23	0.52	0.35	0.52	0.36	0.78	0.64
	SATH03B	-5.66	-0.74	-0.87	-0.30	-1.24	-0.38	-0.66	-0.25
	G3SATH03B	-0.96	-0.32	-0.11	-0.05	-0.02	-0.01	-0.03	-0.01
	G5SATH03B	0.37	0.23	0.68	0.51	0.51	0.35	0.56	0.38
03	GAUGE	-9.15	-0.82	-0.93	-0.32	-0.06	-0.03	0.08	0.04
	SATH03B	-0.85	-0.30	0.24	0.13	0.18	0.10	0.30	0.18
	G3SATH03B	-0.24	-0.11	0.47	0.31	0.41	0.26	0.60	0.43
	G5SATH03B	-1.92	-0.49	0.02	0.01	0.29	0.17	0.55	0.38
04	GAUGE	-2.24	-0.53	-0.70	-0.26	-0.05	-0.02	0.59	0.42
	SATH03B	-0.96	-0.32	-0.13	-0.06	-0.10	-0.05	0.21	0.12
	G3SATH03B	0.42	0.27	0.54	0.37	0.55	0.38	0.59	0.41
	G5SATH03B	0.37	0.22	0.26	0.15	0.15	0.08	0.79	0.65

Table 14. Quality Score Results

CS	NAME SIM	rmse	MARE	PBIAS	MC	CT DELAY	DTW	DDTW	TP DELAY	E%
01	GAUGE	1214.30	1.02	-90.40	1.9	-8	12.99	0.08	6	0.78
	SATH03B	505.47	0.51	44.57	0.55	-6	2.18	0.06	4	-0.39
	G3SATH03B	504.21	0.50	14.74	0.85	-9	0.4	0.07	5	-0.13
	G5SATH03B	759.20	0.63	-35.06	1.35	-9	2.36	0.07	6	0.31
02	GAUGE	180.86	0.34	-0.56	1.01	0	2.78	0.02	6	0.3
	SATH03B	590.00	1.28	29.09	0.71	44	51.58	0.27	-41	0.45
	G3SATH03B	319.73	0.69	46.61	0.75	40	13.44	0.04	3	-0.53
	G5SATH03B	181.07	0.32	25.30	0.91	2	3.5	0.02	5	-0.07
03	GAUGE	335.62	0.86	-86.13	1.86	-5	20.78	0.07	9	1.03
	SATH03B	143.31	0.42	26.08	0.74	0	3.86	0.05	51	-0.26
	G3SATH03B	117.41	0.35	18.68	0.81	-4	1.29	0.03	8	-0.18
	G5SATH03B	179.98	0.41	-26.21	1.26	-5	2.64	0.03	9	0.38
04	GAUGE	278.44	0.61	-34.26	1.34	0	4.21	0.01	0	0.92
	SATH03B	216.48	0.59	41.06	0.59	-49	2.51	0.01	56	-0.55
	G3SATH03B	117.59	0.38	38.22	0.62	0	0.67	0	0	-0.42
	G5SATH03B	123.26	0.37	6.19	0.94	0	0.56	0.01	0	0.22

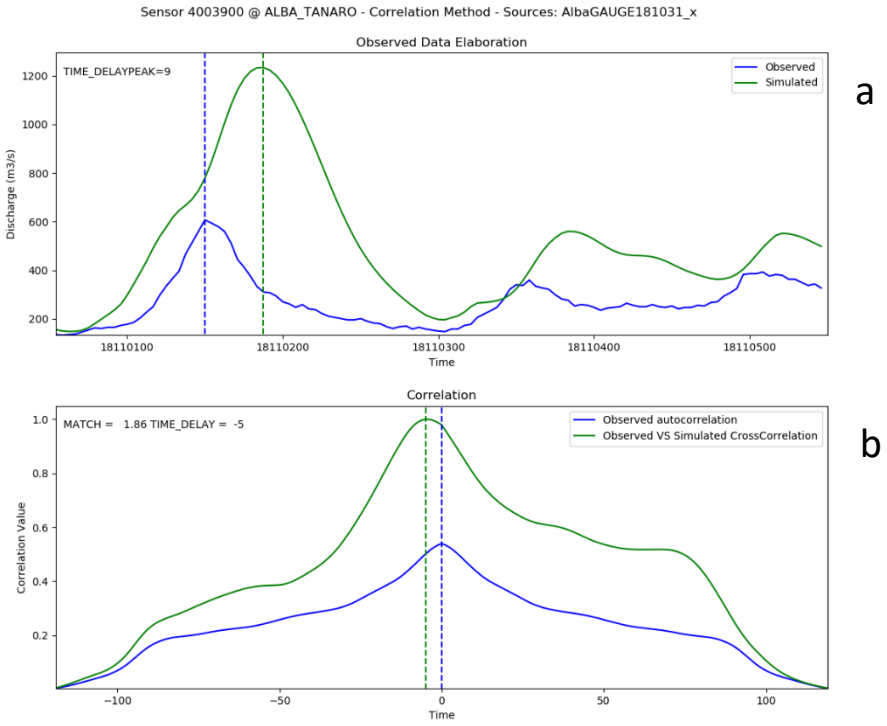


Figure 20. a) Comparison between Observed and Simulated Flow Discharge. The hydrological simulation has been carried out forcing hydrological model with rain gauge data. b) Match Correlation Quality Score: Auto Correlation and Cross Correlation are scaled respect to the max value between the two.

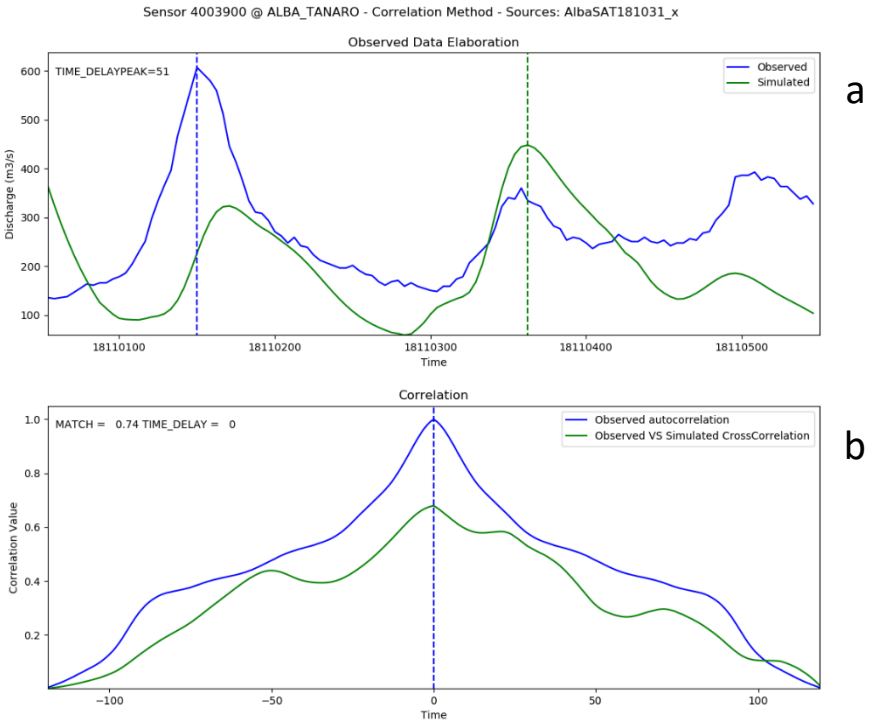


Figure. 21. a) Comparison between Observed and Simulated Flow Discharge. The hydrological simulation has been carried out forcing hydrological model with Satellite data. b) Match Correlation Quality Score: Auto Correlation and Cross Correlation are scaled respect to the max value between the two.

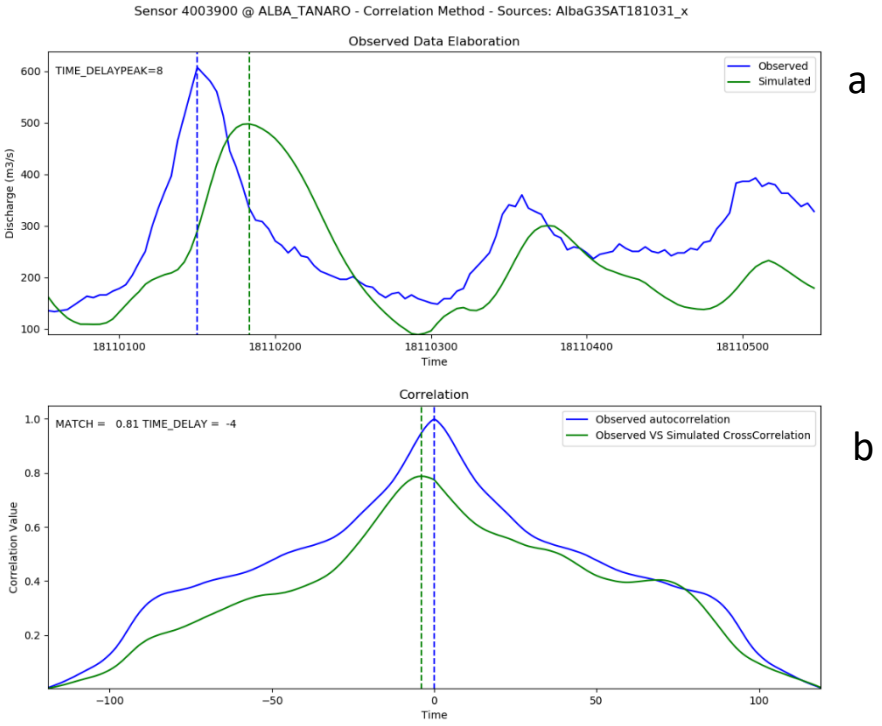


Figure 22. a) Comparison between Observed and Simulated Flow Discharge. The hydrological simulation has been carried out forcing hydrological model with merged rain gauge data (R=3 km) and Satellite data (G3SAT). b) Match Correlation Quality Score: Auto Correlation and Cross Correlation are scaled respect to the max value between the two.

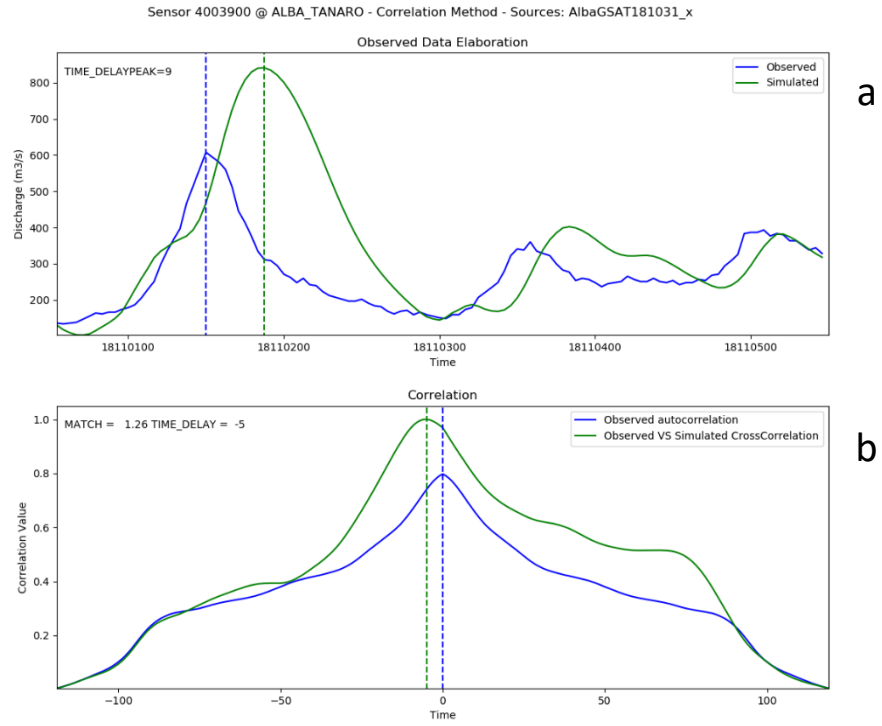


Figure 23. a) Comparison between Observed and Simulated Flow Discharge. The hydrological simulation has been carried out forcing hydrological model with merged rain gauge data (R=5 km) and Satellite data (G5SAT). b) Match Correlation Quality Score: Auto correlation and Cross Correlation are scaled respect to the max value between the two.

7 Conclusion

The assimilation of rainfall data from different data sources deals with a deep knowledge about the source of the observations, its characteristics, and its limits. The downscaling of precipitation and its applicability over different geographical domain has a particular importance for hydrological applications, such as flood prediction and monitoring, as well as water management. The precipitation data is characterized by complex patterns and high spatial variability, which is increased in complex orography. Consequently, it is characterized by the same spatial and temporal variability, and the information given by different observations can help to adequately reproduce such behaviour. The rainfall patterns strongly impact the runoff calculation in hydrological models because the relationship between rainfall distribution and computed discharge is nonlinear.

For the operational activity, hydrological models have been forced with observed and predicted rainfall data. The uncertainty of the hydrological forecasts is strongly connected to the uncertainty of the rainfall field estimates of the atmospheric forecasting models. It is very important to force hydrological models with realistic observed rainfall data, which are fundamental for the spin-up of the model, to reduce this uncertainty. The rain gauge data are considered the most accurate observed data, even if rainfall spatial pattern from rain gauges is affected by errors, depending on data scarcity, sparse sensor network, associated with the lack of a robust or redundant infrastructure, able to guarantee data transmission and functionality during a severe weather event. The preliminary analyses show that the usage of satellite data, H SAF Product, MW/IR H03B and H SAF MW-only products H68, can be a good tool to improve the performance of hydrological simulations. Moreover, the best results can be reached merging satellite data with rain gauge.

Based on the analysis of the single case studies carried out on the Tanaro river, satellite data merged with gauge data were able to improve the areal precipitation estimation, for example for the Match Correlation Score, it was improved of about 80%, while in just one of four case studies, satellite data merged with gauge were not able to improve, but at least the simulation results are comparable to the simulation with rain gauge data only. The results show that even the use of satellite data alone can improve the simulation result, thanks to the cellular automata technique. Further analysis is needed to identify the right methodology for defining the Radius of Influence. The CETEMPS experience suggests that among the assimilation data techniques, Cellular Automata are certainly a valid assimilation data tool.

Next steps are:

- ✓ Improve sensitivity analysis to better understand how to tune satellite data versus rain gauge contributions.
- ✓ Investigate on a better way to select the Radius of Influence.

References

Benesty, J., Chen, J., and Huang, Y.: Time-delay estimation via linear interpolation and cross correlation, IEEE Transactions on Speech and Audio Processing, vol. 12, no. 5, 2004.

Berndt, J. B. and Clifford, J.: Using Dynamic Time Warping to Find Patterns in Time Series, AAAIWS'94 Proceedings of the 3rd International Conference on Knowledge Discovery, 359-370, 1994.

Coppola, E., Tomassetti, B., Mariotti, L., Verdecchia, M. and Visconti, G.: Cellular automata algorithms for drainage network extraction and rainfall data assimilation, Hydrological Sciences Journal, 52:3, 579-592, DOI: 10.1623/hysj.52.3.579, 2007.

Dembélé, M., Hrachowitz, M., Savenije, H. H. G., and Mariéthoz, G.: Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite datasets, Water Resour. Res., 56, e2019WR026085, https://doi.org/10.1029/2019WR026085, 2020.

Duque-Gardeazábal, N., Zamora, D., Rodríguez, E.: Analysis of the Kernel Bandwidth Influence in the Double Smoothing Merging Algorithm to Improve Rainfall Fields in Poorly Gauged Basins; 635–626; 10.29007/2xp6; 2018.

Di Muzio, E., Riemer, M., Fink, A. H., and Maier-Gerber, M.: Assessing the predictability of Medicanes in ECMWF ensemble forecasts using an object-based approach, Quarterly Journal of the Royal Meteorological Society,145:720, 1202-1217, 2019.

Darko, S., Adjei, K.A., Gyamfi, C. et al. Evaluation of RFE Satellite Precipitation and its Use in Streamflow Simulation in Poorly Gauged Basins. Environ. Process; https://doi.org/10.1007/s40710-021-00495-2; 2021.

Gupta, H.V., Kling, H., Yilmaz, K. K., Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling; Journal of Hydrology; Volume 377, Issues 1–2; Pages 80-91, ISSN 0022-1694; https://doi.org/10.1016/j.jhydrol.2009.08.003: 2009.

Hallouin, T.: HydroEval: Streamflow Simulations Evaluator (Version 0.0.3). Zenodo. https://doi.org/10.5281/zenodo.2591217, 2019.

Hughes, D.A.: Comparison of satellite rainfall data with observations from gauging station networks; Journal of Hydrology, 327 (3–4), pp. 399-410, 2006.

Italian Civil Protection Department, CIMA Research Foundation: The Dewetra Platform: A Multi-perspective Architecture for Risk Management during Emergencies. In: Hanachi C., Bénaben F., Charoy F. (Eds.), Information Systems for Crisis Response and Management in Mediterranean Countries. ISCRAM-med 2014. Lecture Notes in Business Information Processing, Cham: Springer, pp. 165-177, 2014.

Jiang, Dejuan & Wang, Kun: The Role of Satellite-Based Remote Sensing in Improving Simulated Streamflow: A Review. Water. 11. 1615. 10.3390/w11081615, 2019.

Keogh, E. J. and Pazzani, M.: Derivative Dynamic Time Warping, Proceedings of First SIAM International Conference on Data Mining, ISBN: 978-0-89871-495-1, 2001.

Keogh, E. J. and Ratanamahatana, C. A.: Exact indexing of dynamic time warping, Knowledge and Information Systems, 7(3), 358-386, doi: 10.1007/s10115-004-0154-9, 2005.

Kling, H., Fuchs, M., Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, Volumes 424–425, Pages 264-277,ISSN 0022-1694,https://doi.org/10.1016/j.jhydrol.2012.01.011 (https://www.sciencedirect.com/science/article/pii/S0022169412000431), 2012.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Li, M., Shao, Q.: An improved statistical approach to merge satellite rainfall estimates and raingauge data, Journal of Hydrology, Volume 385, Issues 1–4, Pages 51-64, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2010.01.023; 2010.

Liang, S., Xiaowen Li, Jindi Wang, Advanced Remote Sensing, Academic Press, Pages 533-556, ISBN 9780123859549, https://doi.org/10.1016/B978-0-12-385954-9.00017-4, 2012.

Lombardi, A., Colaiuda, V., Verdecchia, M., and Tomassetti, B.: User-oriented hydrological indices for early warning systems with validation using post-event surveys: flood case studies in the Central Apennine District, Hydrol. Earth Syst. Sci., 25, 1969–1992, https://doi.org/10.5194/hess-25-1969-2021, 2021.

Maier-Gerber, M., Riemer, M., Fink, A. H., Knippertz, P., Di Muzio, E., and McTaggart-Cowan, R.: Tropical transition of Hurricane Chris (2012) over the North Atlantic Ocean: a multi-scale investigation of predictability, Monthly Weather Review, https://doi.org/10.1175/MWR-D-18-0188.1, 2019.

Mathevet, T. , Michel, C. , Andreassian, V., Perrin, C.: A bounded version of the Nash–Sutcliffe criterion for better model assessment on large sets of basins; V. Andréassian, A. Hall, N. Chahinian, J. Schaake (Eds.), Large Sample Basin Experiment for Hydrological Model Parameterization: Results of the Model Parameter Experiment – MOPEX, IAHS Publ, p. 567; 2006

Nash, J.E., Sutcliffe, J.V.: River flow forecasting through conceptual modelspart I - A discussion of principles. Journal of Hydrology 10 (3), 282–290; 1970.

Pool, S., Vis, M. and Seibert, J.: Evaluating model performance: towards a non-parametric variant of the Kling-Gupta efficiency, Hydrological Sciences Journal, 63:13-14, 1941-1953, DOI: 10.1080/02626667.2018.1552002, 2018.

Rabiner, L. R. and Gold, B.: Theory and Application of Digital Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, p. 401, ISBN 0139141014, 1975.

Rabiner, L.R. and Schafer, R.W.: Digital Processing of Speech Signals. Signal Processing Series. Upper Saddle River, NJ: Prentice Hall., 147–148, ISBN 0132136031, 1978.

Sakoe, H. & Chiba, S. (1978) Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoustics, Speech, and Signal Proc., Vol. ASSP-26, 43- 49.

Saouabe, T.; El Khalki, E.M.; Saidi, M.E.M.; Najmi, A.; Hadri, A.; Rachidi, S.; Jadoud, M.; Tramblay, Y. Evaluation of the GPM-IMERG Precipitation Product for Flood Modeling in a Semi-Arid Mountainous Basin in Morocco. Water, 12, 2516, 2020.

Shi, H, Chen, J., Li, T., Wang, G.: A new method for estimation of spatially distributed rainfall through merging satellite observations, raingauge records, and terrain digital elevation model data; Journal of Hydro-environment Research, Volume 28, Pages 1-14; ISSN 1570-6443; https://doi.org/10.1016/j.jher.2017.10.006; 2020.

Sevruk, B. `Snow cover measurements and areal assessment of precipitation and soil moisture', Operational Hydrology Report, 35, Publ. 749. World Meteorological Organisation, Geneva. 91 pp., 1992.

WMO Guide to Hydrological Practices, fifth edn. WMO no. 168, World Meteorological Organization, Geneva, Switzerland, 1994.