Ensemble Simulation From Multiple Data Sources In A Spatially Distributed Hydrological Model Of The Rijnland Water System In The Netherlands

Hartanto, Isnaeni; Andel, Schalk Jan Van; Alexandridis, T.K.; Solomatine, Dmitri

Publication date
2014

Document Version
Final published version

Published in
Proceedings of the HIC 2014 - 11th international conference on hydroinformatics

Citation (APA)
Hartanto, I., Andel, S. J. V., Alexandridis, T. K., & Solomatine, D. (2014). Ensemble Simulation From Multiple Data Sources In A Spatially Distributed Hydrological Model Of The Rijnland Water System In The Netherlands. In M. Piasecki (Ed.), *Proceedings of the HIC 2014 - 11th international conference on hydroinformatics* (pp. 1189-1194)

International Conference on Hydroinformatics

8-1-2014

Ensemble Simulation From Multiple Data Sources In A Spatially Distributed Hydrological Model Of The Rijnland Water System In The Netherlands

Isnaeni Murdi Hartanto

Schalk-Jan Van Andel

Thomas K. Alexandridis

Dimitri P. Solomatine

Recommended Citation

Hartanto, Isnaeni Murdi; Van Andel, Schalk-Jan; Alexandridis, Thomas K.; and Solomatine, Dimitri P., "Ensemble Simulation From Multiple Data Sources In A Spatially Distributed Hydrological Model Of The Rijnland Water System In The Netherlands" (2014). *CUNY Academic Works.*
http://academicworks.cuny.edu/cc_conf_hic/299

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

ENSEMBLE SIMULATION FROM MULTIPLE DATA SOURCES IN A SPATIALLY DISTRIBUTED HYDROLOGICAL MODEL OF THE RIJNLAND WATER SYSTEM IN THE NETHERLANDS

ISNAENI M HARTANTO (1), SCHALK JAN VAN ANDEL (1), THOMAS K. ALEXANDRIDIS (2), DIMITRI P. SOLOMATINE (1, 3)
(1): UNESCO-IHE Institute for Water Education, Delft, Netherlands
(2): Lab of Remote Sensing and GIS, Faculty of Agriculture, Aristotle University of Thessaloniki, Greece
(3): Water resources Section, Delft University of Technology, Delft, Netherlands

Data for water management is increasingly easy to access, it has finer spatial and temporal resolution, and it is available from various sources. Precipitation data can be obtained from meteorological stations, radar, satellites and weather models. Land use data is also available from different satellite products and different providers. The various sources of data may confirm each other or give very different values in space and time. However, from these various data sources, it can often not be judged beforehand that one data is correct and others are wrong. Each source has its own value for a particular purpose. The Rijnland area in the Netherlands is one of the areas for which various data sources are available. Data sources that are researched in this paper are precipitation from rain gauges and radar, and three different land use maps. Various sources of data are used as input to the hydrological model (SIMGRO) of the water system to produce different discharge model output. Each run provides a member of the ensemble simulation which are combined to improve prediction of discharge from the catchment. It is shown that even simple averaging allows for increasing the model accuracy.

INTRODUCTION

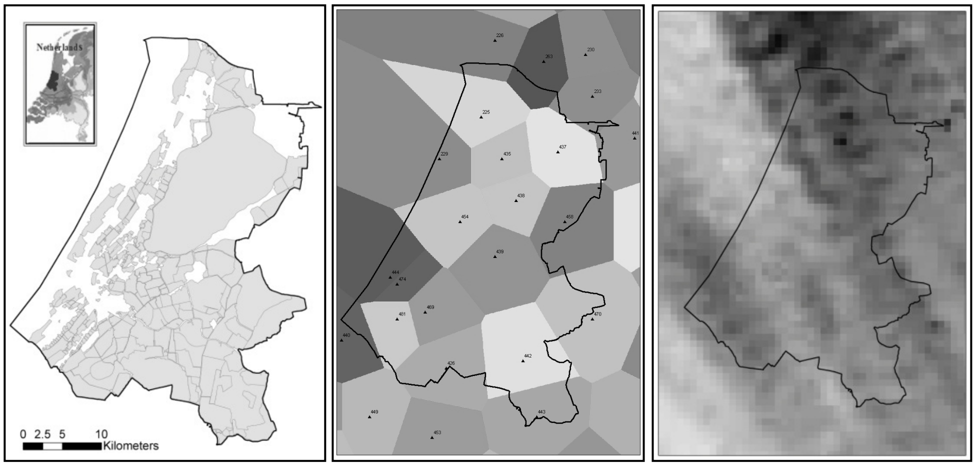
The hydro-meteorological data from various sources allows for building an ensemble of models. The Ensemble modeling method has long been utilized in weather prediction and hydro-meteorological prediction [1, 2, 3]. Ensemble members can be generated on the basis of perturbed initial conditions, differences in parameterization, different model structure, or, as is the case in this paper, on the basis of different input data. Each of the ensemble members has a different combination of input data, e.g. precipitation or land use, with their own subsequent estimation of hydrological variables. Although one data source might be considered as the best reliable source, it is often advisable to combine several sources. For example, precipitation quantity from rain gauges is considered more reliable than from radar, but the spatial variability provided by the latter is considered more accurate in return. The using of several data sources can capture events that might be overlooked by using only one [4]. This study presents a method of combining several model outputs based on a variety of data sources.

CASE STUDY OF RIJNLAND, THE NETHERLANDS

The Rijnland area is located in the western part of the Netherlands with an approximate area of 1000 km^2 (Figure 1.a). The area mostly consists of flat low-lying reclamation land (72%). The reclamation land has a ground elevation below the sea level, which makes the area prone to flooding and highly dependent on dikes for flood protection. The area is characterized by clay and peat soils with a shallow ground water level. Rijnland is divided into hundreds of polders (small irrigation and drainage sections), where the water in each polder is stored in canals and discharged to a main storage basin. The water level in each polder is maintained around a certain water level by pumps and/or weirs. The main storage basin covers an area of 45 km^2 in which the water level maintained between -0.60 and -0.65 NAP (Normal Amsterdam level~mean sea level) [5], which is higher than the water level in most of polders.

The Rijnland area is a highly controlled water system, involving a high influence of human interference. A different target water level is applied in each polder. Furthermore, the target level for summer time (dry season) is different from winter time (wet season), with higher summer target levels to maintain the same ground water level. The water managers control hundreds of pumping stations, weirs and sluices to meet the strict water level requirements. The human influences make the catchment behave in a non-natural way, with natural discharge and levels being adjusted on the basis of predictions, early or late pumping, custom small gate openings from local farmers etc. This intervention occasionally makes that the measured discharge appears disconnected from actual rainfall events.

Located in the Netherlands, Rijnland area has a good coverage of rainfall stations, although the stations are maintained by several owners. In this paper, open access data from KNMI has been used [6], where archived recordings of 21 rain gauges have been processed on the basis of Thiessen polygons for Rijnland area (Figure 1.b). Radar rainfall data with one kilometer resolution is also available for the area (Figure 1.c). Both of the data are in daily time step and are already calibrated, and the missing values are also already filled.



(a) (b) (c)

Figure 1. Rijnland area and polder map (a), Thiessen polygon for rainfall (b), one kilometer radar rainfall grid (c)

Besides the two sources of precipitation data, three land use-land cover (LULC) maps are used to build the ensemble. The first LULC map is the LGN map [7], which is the main land use product for the Netherlands, as used by the Rijnland Water Board. The second land use map is the LULC map from two Landsat 5 TM images (June and September) with a 30 m resolution

[8]. This was created using supervised spectral classification with training samples defined during a field survey, populated with extra points identified with photo-interpretation of very high resolution satellite images. The resulting map had an overall accuracy of 88%. The third map is a 300 m resolution LULC map product from GLOBCOVER [9] using MERIS sensor onboard ENVISAT satellite. The nominal accuracy of this product is 79% on a global scale [9].

SIMGRO MODEL OF RIJNLAND

SIMGRO is a spatially distributed hydrological model build for shallow ground water level areas and able to simulate controlled water systems [10]. The software is suitable to model the hydrological processes in the Rijnland area. SIMGRO modeling system is a hub to three different models, which are Meta-SWAP for modeling soil-vegetation-atmosphere interactions, MODFLOW for three dimensional ground water flows, and SurfW a surface-water model. These three models communicate through shared states [11].

The Rijnland SIMGRO model has a grid of 50x50 m resolution, covering an area of 125 km^2, resulting in over 500,000 of grid cells. The 50x50 m resolution is chosen due the fact that in the polder area there are canals with a typical distance of 50 meters, and the model tries to capture it. The surface-water model is a simplified approach of linked reservoir model which is acceptable for low-gradient, low flow velocity water of Rijnland's canals. The area is divided into more than 700 sub-polders with their own target water level and pump/weir operational properties.

RESULTS AND DISCUSSION

Using a combination of two rainfall data sources and three LULC map, an ensemble model result of six members is built and run (Table 1), and a simple average of the six members is also presented. Each of the members gives already good results during calibration and validation period of 2011 and 2012 respectively [12]. However, there are certain differences in discharge estimations. We show that an ensemble of discharge simulation results from 2013 allows for generation of a better estimate.

Table 1. Ensemble member descriptions

Ensemble member	Precipitation data source	LULC map
Q_Th_LGN	Rain gauges, Thiessen polygon	LGN
Q_Th_Landsat5	Rain gauges, Thiessen polygon	Landsat 5 TM
Q_Th_Meris	Rain gauges, Thiessen polygon	GLOBCOVER
Q_Rd_LGN	Radar data	LGN
Q_Rd_Landsat5	Radar data	Landsat 5 TM
Q_Rd_Meris	Radar data	GLOBCOVER

Due to the unpredicted nature of controlled water system, the daily discharge values are smoothed by taking a three-day moving average.

The model performance is evaluated with Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and ratio of the root mean square error to the standard deviation of measured data (RSR). The three evaluation criteria are recommended as evaluation for watershed simulation proposed by Moriasi et. al. [13].

The model performance for the period from January to October 2013 can be seen in Table 2. The additional *average_ensemble* in the table is a simple averaging method using equal weight.

However, there are many averaging methods available such as those based on performance-based weighting and dynamic weighting [14], regression-based, neural network and Bayesian model averaging [4], etc. In this preliminary study, a simple weighting is used, and more sophisticated weighting schemes are planning to be applied for further research.

The Table 2 shows that even using simple averaging method of the ensemble members gave better performance than the individual members. The average ensemble gave better performance for all three evaluation criteria. The relevance of considering both ground station and radar data in this case is shown by the negative biases of all three models with ground stations while the models that use radar all have a positive bias.

Table 2. Model performance for each ensemble member and the average

Ensemble member	Jan-Oct 2013		
	NSE	PBIAS	RSR
Q_Th_LGN	0.81	-11.6%	0.43
Q_Th_Landsat5	0.82	-3.9%	0.42
Q_Th_Meris	0.83	-10.3%	0.42
Q_Rd_LGN	0.83	7.1%	0.41
Q_Rd_Landsat5	0.83	14.3%	0.41
Q_Rd_Meris	0.83	7.1%	0.41
average_ensemble	0.84	0.4%	0.40

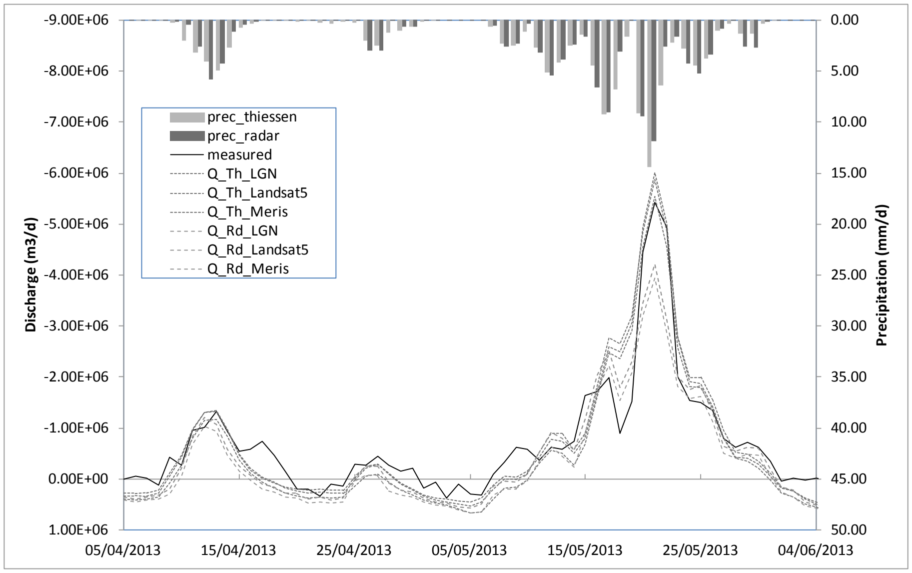


Figure 2. Discharge comparison between ensemble members and measured data for April to June.

Figure 2 shows the ensemble plot of the simulated discharge together with the measured discharge, where all discharge values are again a three-day moving average. The negative discharge values indicate that water is pumped out from the area, and positive values indicate an inflow has occurred. In April and beginning of May, with relatively long periods without rainfall, all members are most of the time underestimating the outflow and overestimating the inflow (model lines are below the measured line. Note the inverted y-axis). This bias for all the members may be because the human influence is relatively high in the summer (dry) period, which runs from April to August/September. Still, because all members show the same bias for periods without rain, this is a bias in the model that has to be reduced. On the other hand, members with radar rainfall input give good simulated values for the high peak around 20th of June. The comparison of average values of ensemble member is presented in Figure 3. The average ensemble generally is better than the individual ensemble members.

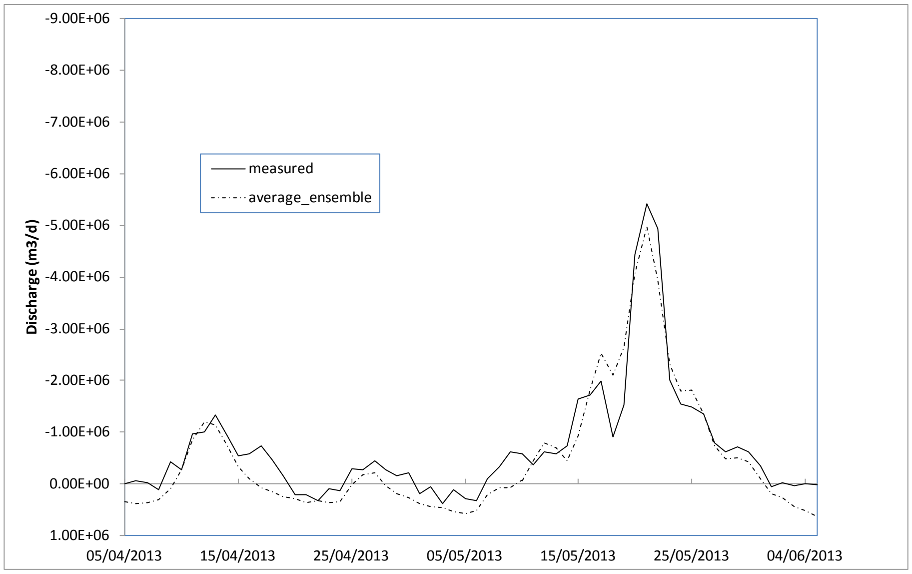


Figure 3. Discharge comparison between average of ensemble members and measured data.

CONCLUSION

We have setup a simulation model and used six different data sources to produce six model outputs (discharges). It is shown that by using the simple averaging of the ensemble members, it gives better results than each of the individual simulations. Each ensemble member could take into account variables and events that might be overlooked by other members, resulting in a better combined simulation. A simple direct averaging method applied in ensemble simulation gives improvement on the output performance.

In the future experiments we will be testing other ensemble averaging schemes, in particular, those based on performance-based weighting and dynamic weighting [14].

ACKNOWLEDGEMENTS

EU FP7 MyWater research project. http://www.mywater-fp7.eu/

REFERENCES

[1] *Van Andel SJ, Price RK, Lobbrecht AH, et al. "Ensemble Precipitation and Water-Level Forecasts for Anticipatory Water-System Control". Journal of Hydrometeorology vol 9 (2008), pp776–788*

[2] *Chiang Y-M, Hsu K-L, Chang F-J, et al. "Merging multiple precipitation sources for flash flood forecasting". Journal of Hydrology vol 340(2007), pp 183–196.*

[3] *Strauch M, Bernhofer C, Koide S, et al. "Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation". Journal of Hydrology vol 414–415(2012), pp 413–424.*

[4] Duan Q, Ajami NK, Gao X, Sorooshian S, "Multi-model ensemble hydrologic prediction using Bayesian model averaging". Advances in Water Resources vol 30 (2007), pp 1371–1386.

[5] Van Andel SJ, Price R, Lobbrecht A, van Kruiningen F, "Modeling Controlled Water Systems". Journal of Irrigation and Drainage Engineering vol 136 (2010), pp 392–404.

[6] KNMI (2013). https://data.knmi.nl/portal/KNMI-DataCentre.html

[7] Wageningen University (2011)http://www.wageningenur.nl/nl/Expertises-Dienstverlening/Onderzoeksinstituten/Alterra/Faciliteiten-Producten/Kaarten-en-GISbestanden/Landelijk-Grondgebruik-Nederland.htm

[8] Alexandridis, T., Cherif, I, Topaloglou, C., Monachou, S., Strati, S., Stavridou, D., Iordanidis, C., Bilas, G., Kalogeropoulos, C., Araujo, A., Nunes, N., Leitao, P.C., Best practices. Report of the project "MyWater - Merging hydrologic models and EO data for reliable information on water". Submitted to Research Executive Agency - European Commission, (2012), pp 189

[9] Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R. & Weber, J.-L. "Globcover: The most detailed portrait of Earth". European Space Agency Bulletin (2008), pp 24–31.

[10] Querner EP, "Description of a regional groundwater flow model SIMGRO and some applications". Agricultural Water Management vol 14(1988) pp 209–218.

[11] Van Walsum PE V, Veldhuizen AA, "Integration of models using shared state variables: Implementation in the regional hydrologic modelling system SIMGRO". Journal of Hydrology vol 409 (2011) pp 363–370.

[12] Hartanto IM, van Andel SJ, Lobbrecht AH, van Griensven, A., Solomatine, D. P. "Integrating earth observation data into hydrological modeling and water management". EGU Gen Assembly Conference Abstract vol 14, Vienna (2012) pp 13110

[13] Moriasi D, Arnold J, "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations". Transactions of the ASABE 50.3 (2007) 885–900

[14] Kayastha. N., Ye, J., Fenicia, F., Kuzmin, V. and Solomatine, D. P. Fuzzy committees of specialized rainfall-runoff models: further enhancements and tests. Hydrology & Earth System Sciences, 17(2013), pp 4441-4451.