GIS-BASED DATA PREPARATION AND SURFACE FLOW SIMULATIONS OF THE NILÜFER RIVER BASIN IN TURKEY

Serdar Korkmaz[1]

[1]Department of Civil Engineering, Uludağ University, Turkey, 16059 Bursa
E-mail: skorkmaz06@gmail.com

Abstract

The surface hydrology of the Nilüfer river basin in northwestern Turkey is studied. The basin forms the essential water resource of the city of Bursa which is the fourth largest in Turkey. Due to the Mount Uludağ with a peak of 2531 m, highly variable weather conditions are observed inside the basin. In current study, a numerical hydrological model is formed in order to simulate the surface flows of the basin. The inputs are prepared in GIS environment. The digital elevation model (DEM) of the basin is obtained from SRTM database. Through the analysis of the DEM, the flow directions and drainage areas are obtained. The river network is compared to the ones obtained from international databases as well as to a printed map. In the end of the analysis, total basin area is found as 1986 km^2. The missing precipitation data is estimated using two methods, namely, inverse distance weighting method and correlation coefficient weighting method. The potential evapotranspiration data is acquired from CGIAR-CSI Global Potential Evapo-Transpiration and Aridity Index database and 45 zones covering the basin area are obtained. The land cover data is obtained from Turkish Statistical Institute in the same format as CORINE. The numerical model applied to the Nilüfer river basin is MODCOU developed at Ecole des Mines de Paris. Initially precipitation is partitioned into infiltration, surface runoff, evapotranspiration, and soil stock, and then, the simulations of surface flows are made. A good fit between filtered direct flow and MODCOU simulations is sought.

Introduction

The availability of high resolution data and easy access have contributed to the performance of hydrological models. Today, the input data of most hydrological models are being prepared in Geographic Information Systems (GIS) environments using digital databases available in standard formats with continuously increasing resolutions. Through the analysis of a Digital Elevation Model (DEM) the hydrological structure of a basin is easily constructed. Meteorological data is another important input for hydrological models. Unfortunately, in many populated regions of the world, ground-based measurement networks are either sparse in both time and space or nonexistent (Behrangi et al. 2011). Nevertheless, there are many methods for the estimation of missing meteorological data, two of which are used in this study.

The surface hydrology of the Nilüfer river basin is studied. For the flow simulations, a numerical distributed model named MODCOU (Ledoux 1980; Ledoux et al. 1984), developed at Ecole des Mines de Paris, is used. The daily values of surface runoff and the flow transfer along the river network is computed by the model and the groundwater flow is not considered. However, it is not possible to compare the results with the observed discharges. This is because the streamgage observations represent the total runoff which is the summation of baseflow and direct flow. For this reason, a two-parameter filter equation presented by Eckhardt (2005) is used to separate the baseflow from the total runoff hydrograph. The direct flow obtained from MODCOU simulations are compared to the filtered observation.

Nilüfer River Basin

Nilüfer river is the essential stream of the city of Bursa in northwestern Turkey and forms the main water resource of the region. The city is emphasized with a population of 2.5 million with large industrial zones that consume and contaminate high volumes of water. The length of the river is about 135 km (Fig. 1). Its sources are located in the southern slopes of Mount Uludağ (the ancient Mysian Olympus). The exact location falls to the north of the district Keles at altitudes slightly higher than 1000 m. Its outlet is the Susurluk (Simav) river in the west.

Nilüfer has many tributaries, some of which are Değirmen, Deliçay, Ismetiye, Ayvalı and Hasanağa. The basin has an area of 1986 km^2, and the mean annual discharge is 16.77 m^3/s. There are two streamgage stations on the river, namely EIE321 and DSI344 with catchment areas of 1297 and 373 km^2 respectively (Fig. 1).

The climate in the basin possesses the general characteristics of the Marmara region. The weather is mostly cloudy and humid in the winter and dry in the summer. There is generally a mild climate in the basin;

however, regional variations are also observed. While the Marmara Sea in the north provides a mild and soft climate, the Mount Uludağ in the south with a crest elevation of 2531 m causes hard weather conditions, particularly, in the winter. The annual rainfall is about 716 mm. There is also a considerable amount of snowfall around Uludağ. Although highly variable among years, the annual snow depth is 728 mm.

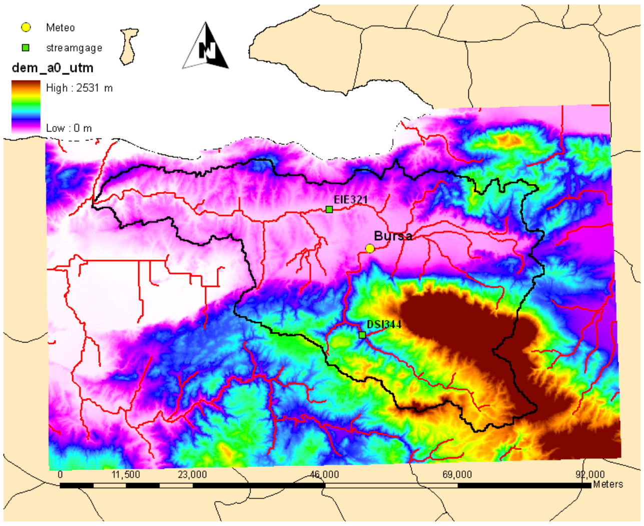

Figure 1: Nilüfer river basin (black), river network (red) and the DEM

Digital Elevation Model Analysis

The NASA Shuttle Radar Topographic Mission (SRTM) has provided digital elevation models (DEMs) for over 80% of the globe (Farr et al. 2007). The SRTM data is available as 3 arc second of latitude and longitude which corresponds to approximately 90 m resolution at the equator. The DEMs are provided in 5 deg x 5 deg tiles in the geographic coordinate system WGS84 datum. The two DEM tiles containing the Nilüfer river basin were chosen. As the tiles covered a larger area than needed, an approximate rectangular mask was drawn and used to cut out the necessary portions from these tiles. Afterwards, the resultant DEMs were merged and projected to WGS84 UTM Zone 35N coordinate system. As there was no initial opinion about the river network of the DEM, analyses were carried out without mapping any river network. The sinks that hinder the continuity of flow were filled using a D4 algorithm (Korkmaz et al. 2009). Afterwards, the flow directions are assigned. Using flow directions, the drainage area at each cell was calculated. By giving a threshold value to drainage areas, a river network was obtained. The threshold used is 2600 cells (approx. 21 km^2). However, this river network was obtained using merely the information contained in the DEM. The actual river network could be in a different form. For this reason the network was compared to a printed map and to several digital river networks acquired from internet databases. The printed 1:250 000 scaled map was provided by the State Hydraulic Works of Turkey (DSI) and contained a detailed hydrography of the region. On the other hand, the first digital database is called GeoCommunity (2011). The scale of the data is 1:1million. Natural Earth (2011) is another database that provides vector and raster map data at different scales in the geographic projection WGS84 datum. The final database used in current study is the River and Catchment Database version 2.1 (July 2008) of Catchment Characterisation and Modelling (CCM2) activity of Joint Research Center (Vogt et al. 2007). CCM2 has a pan-European database of river networks and catchments on the WGS84 coordinate system.

The mapped river was burned onto the initial DEM using the AGREE method (Maidment 2002) and the analyses were resumed. A total of 45 sub-basins were obtained. Their areas ranged from 0.28 to 152.37 km^2. Finally the outlet of the basin was marked and the basin area was delineated. The total area of the basin was computed as 1986 km^2 (Fig. 1).

Meteorological Data

The meteorological data was obtained from the global surface summary of day product of the National Climatic Data Center of the US (NCDC 2011). The data of the stations located around the basin were downloaded. The station of Bursa falls inside the basin, and all the rest are outside. Their observation periods and distances to the basin are presented in Table 1.

When the Thiessen polygons are drawn, it was observed that among the outside stations, Bandırma and Yenişehir cover very small portions over the basin area and the others does not reach the basin at all. For this reason, the station of Bursa was assumed to represent the entire basin area. The available data was supplied for the period between Jan 1, 1973 and Dec 31, 2010 inside which a total of 78 days were missing. The other stations were used to estimate the missing data.

Table 1: Meteorological stations

Station no.	Location	Distance from basin (km)	Observation period
17115	Bandırma	41.3	1950-2011
17116	Bursa	0.0	1973-2011
17118	Yenişehir	25.1	2007
17119	Yalova	34.0	2008-2011
17120	Bilecik	51.0	1995-2011
17150	Balıkesir	90.1	1950-2011

Methodology for precipitation estimation

In literature there are many studies for the estimation of missing precipitation data. For example, Teegavarapu and

Chandramouli (2005) introduce and compare 8 different methods. In current study, the most commonly used one, the inverse distance weighting method (IDWM) (ASCE 1996) and one of the methods suggested by Teegavarapu and Chandramouli (2005), the correlation coefficient weighting method (CCWM), were used. The IDWM uses the closeness of surrounding stations to the base station as a criterion to determine the weight, whereas the CCWM uses the spatial auto-correlation as the weight. In both methods the missing value is estimated using the following expression:

$$P_b = \sum_{i=1}^{n} w_i P_i \tag{1}$$

where P_b is the missing value at the base station; P_i is the observation at station i; n is the number of surrounding stations; and w is the weight. In IDWM, the weight function has the following form:

$$w_i = \frac{d_{bi}^{-k}}{\sum_{j=1}^{n} d_{bj}^{-k}} \tag{2}$$

where d_{bi} is the distance between the base station and station i; k is the power parameter and usually k=2. For CCWM the following weight function is used:

$$w_i = \frac{\rho_{bi}}{\sum_{j=1}^{n} \rho_{bj}} \tag{3}$$

where ρ_{bi} is the correlation coefficient computed using the historical values of the base station and the station i. In order to compare the effectiveness of the two methods, all the historical rainfall and snow data at the Bursa station were regenerated using the stations at Balıkesir and Bandırma. The distances and the weights used are given in Table 2. It is clearly seen that in IDWM the weight of Bandırma station is higher than that of Balıkesir due to the given distances. On the other hand, when the spatial auto-correlation is considered, it is observed that the rainfall and snow depth observations at Balıkesir station are slightly more correlated to those at Bursa than Bandırma.

Table 2: The distances to Bursa station and weights used in IDWM and CCWM

		Weights (w)		
Name of station	Distance to Bursa (km)	IDWM (both)	CCWM (rainfall)	CCWM (snow)
Balıkesir	115.2	0.38	0.52	0.55
Bandırma	89.8	0.62	0.48	0.45

In order to have a decision on which method to choose, the error criteria are considered. In Table 3, the root mean squared error (RMSE), mean absolute error (MAE) and R^2 values are presented for both rainfall and snow depth estimations. It is observed that in all the criteria CCWM yields a better performance. The results can be improved by installing additional meteorological stations inside the basin. As a result, the missing values at Bursa station were estimated using the CCWM method.

Table 3: The error criteria for comparing IDWM and CCWM

	Rainfall		**Snow**	
	IDWM	CCWM	IDWM	CCWM
RMSE (mm)	7.31	7.12	20.16	19.54
MAE (mm)	1.83	1.80	1.92	1.88
R^2	-0.10	-0.04	0.09	0.14

Climatology

For the period between 1973 and 2010, the mean annual rainfall depth is 716 mm. The smallest value is observed in 2004 as 373 mm and the largest in 2010 as 1181 mm. When the monthly values are considered it is observed that the rainfall depth is highly variable throughout the year. Values are near and above 80 mm from October to January and lower than 20 mm in August. The mean annual snow depth is 728 mm and shows large variations among years. The highest value is observed in 1992 as 3856 mm whereas there is no observed snow in some years. When the mean monthly snow depths are considered, the highest value is in February with 423 mm and there is no snow between April and September.

The mean annual temperature is 14.5°C and the mean monthly temperature varies between 5.1°C and 24.4°C. The maximum temperature recorded in the given period is 43.8°C in July 2000 and the minimum is -16.4°C in February 1985.

Potential Evapotranspiration

Another important meteorological input parameter for a hydrological model is the potential evapotranspiration (PET). In current study, PET data was obtained from CGIAR-CSI Global Potential Evapo-Transpiration and Aridity Index database (Trabucco and Zomer 2009). In order to calculate the PET data, they used the WorldClim Global Climate Data (Hijmans et al. 2005) as input parameters. The global PET data was cropped in GIS environment to cover the basin area and its resolution was reduced from 710 m to 11.5 km in order to form larger meteorological zones. In the end there were a total of 45 zones over the basin. The monthly data was converted to daily values for the hydrological model to be used.

Land Cover Data

In current study, the land cover data was obtained from Turkish Statistical Institute (Başoğlu et al. 2006). They followed CORINE methodology and nomenclature. The land cover map of Turkey was cut out by a mask and reduced to the basin area. The percentages of the land cover distribution on the basin area according to level 2 classification of CORINE are presented in Table 4. The most widespread type is Forests with 41.3% and the second is Heterogeneous Agricultural Areas with 18.1%. When level 1 classification is considered Forest and semi-natural areas cover 54.0% of the basin area.

Table 4: Land Cover Data Distribution on the Nilüfer basin

Code	Level 2	Percentage (%)
11	Urban fabric	4.9
12	Industrial, commercial and transport units	2.4
13	Mine, dump and construction sites	0.1
21	Arable land	10.8
22	Permanent crops	9.3
23	Pastures	0.2
24	Heterogeneous agricultural areas	18.1
31	Forests	41.3
32	Scrub and/or herbaceous vegetation associations	10.9
33	Open spaces with little or no vegetation	1.9
51	Inland waters	0.3

Flow Calculations

In current study, as the surface hydrology of the basin is considered, the values of direct flow are calculated. The direct flow is the portion of total runoff that does not come from the aquifer. The portion coming from the aquifer is the baseflow. Therefore, the equation defining the total runoff can be given as:

$$Q_T = Q_B + Q_D \tag{4}$$

where Q_T=total runoff, Q_B=baseflow and Q_D=direct flow. Direct flow is calculated in two ways; first, baseflow separation is made using a filter equation; and second, the numerical model MODCOU is used to simulate the direct flow.

Baseflow Separation

In literature, many studies regarding baseflow separation can be found. The most widely used methods are filtering separation methods. The name comes from the process that filters the baseflow from the total runoff hydrograph according to its long wave characteristic. Similarly, the short-duration waves and sharp peaks are associated with the direct flow.

Recently, Eckhardt (2005) stated that one-parameter filters were all special cases of two-parameter filters and proposed the following equation:

$$Q_B(t) = \frac{(1-BFI_{max})\alpha}{1-\alpha BFI_{max}} Q_B(t-1) + \frac{(1-\alpha)BFI_{max}}{1-\alpha BFI_{max}} Q_T(t) \tag{5}$$

with the condition $Q_B \leq Q_T$, where t = time level, α = recession constant, BFI_{max} = the maximum value of baseflow index (the long term ratio of baseflow to total runoff). In Eq. (5) α is easily determined by a recession analysis. On the other hand, BFI_{max} is not measurable and hence it is modeled by the algorithm. For this reason a sensitivity analysis was made in order to determine which parameter affects the result of the filter equation most. During the analysis, one parameter is modified while the other is kept constant and the effect of the parameter on the mean baseflow is observed. The result of the sensitivity analysis is shown in Fig. 2. It is observed that the filter equation is highly dependent on BFI_{max} rather than α. As a result, during baseflow separation, determination of BFI_{max} is studied thoroughly.

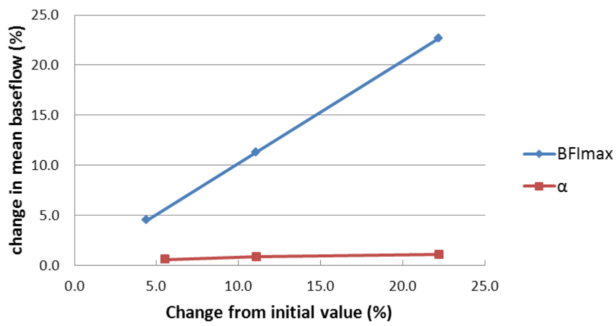

Figure 2: Sensitivity analysis for the estimation of key parameter in the filter equation

Initially, α is determined as 0.94 and 0.97 for stations EIE321 and DSI344 respectively by recession analysis in which the linear reservoir model is used. On the other hand, BFI_{max} is a non-measurable quantity and its selection has the disadvantage of being subjective (Eckhardt 2005). As the selection is subjective there are no particular criteria than visual observation. In current study, several values were tried for different years and an approximate value of 0.45 for both stations was estimated to give the best result. The results of baseflow separation on three years of data for EIE321 and on 15 months of data for DSI344 are depicted in Fig. 3.

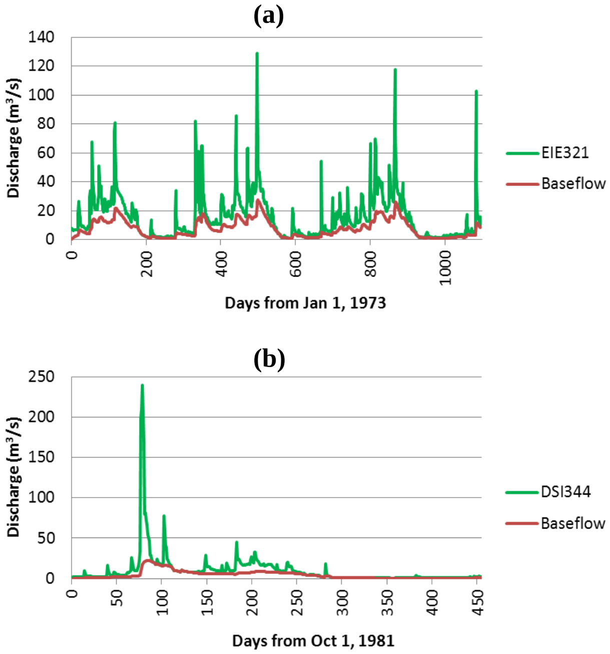

Figure 3: The gauged streamflow and filtered baseflow hydrographs at streamgage stations (a) EIE321 (1297 km^2) and (b) DSI344 (373 km^2)

Numerical Model

The numerical model applied to the Nilüfer river basin is MODCOU (Ledoux 1980; Ledoux et al. 1984) developed at Ecole des Mines de Paris. MODCOU is a spatially distributed hydrological model used to simulate the surface runoff and groundwater flow in multilayered hydrological systems. It was previously applied to several basins in France and Bulgaria (Gomez et al. 2003; Artinyan et al. 2008; Korkmaz et al. 2009).

The surface component of MODCOU, namely MODSUR, initially calculates the partition of available water among infiltration, surface runoff, evapotranspiration, and soil stock based on the type of land cover. MODCOU subsequently does the flow transfer along the river network by using the values of surface runoff. In addition, MODSUR computes the transfer of surface runoff to river cells. Each cell on the river network is considered to be the outlet of a drainage area. The drainage area of each river cell is divided into a number of isochronal zones which is equal to the number of time steps (Δt) necessary for the flow of the most upstream point to reach the river cell. In order to carry out this division, the concentration time of the basin is given as input. Therefore, the flowrate of surface runoff into a river cell, *qr*, at time level *t* is described as follows:

$$qr_i(t) = \sum_{k=0}^{N_{iz}-1} \sum_{j=1}^{N_k} QRR_j(t - k \cdot \Delta t) \tag{6}$$

where N_{iz} is the number of isochronal zones, N_k is the number of cells inside the isochronal zone *k*, QRR_j is the surface runoff produced in surface cell *j*. As can be seen in Eq. (6), the surface runoff calculated in surface cells are transferred directly to river cells without infiltrating in other cells. After calculating the daily surface runoff input to all the river cells, MODCOU is run to simulate the direct flow. The entire river network is divided into segments using isochrones and the flow of all the cells inside a segment arrives at the outlet within the same time step. The flow transfer is computed as follows:

$$Vs_i^{t+1} = (1 - Cds_i) \cdot (Vs_i^t + Qr_i^t) + \sum_{j=1}^{Nup} Cds_j \cdot \left(Vs_j^t + Qr_j^t\right) \tag{7}$$

where Vs_i^{t+1} is the total volume of water in segment *i* at time *t*+1, Cds_i is the discharge coefficient of segment *i*, Qr_i is the total volume of surface runoff arriving into segment *i*, *Nup* is the number of upstream segments which flow directly into segment *i*. Eq. (7) is composed of two parts. The first part computes the volume of water retained in the segment after discharge and the second part computes the inflowing volume of water from upstream segments. Moreover, this flow is controlled by the parameter *Cds*. In fact, every river cell has a discharge coefficient, *Cd*, however, when the transfer between segments is considered, it is the smallest of the discharge coefficients in the segment that controls the flow. Therefore *Cds* is equal to the smallest value of *Cd* in a segment. Once the transfer between the segments is completed, the total volume of water in a segment is distributed among the river cells.

For the simulations, the period between Jan 1, 1973 and Dec 31, 1982 is chosen. Initially, the surface model MODSUR is run. The main input parameters to MODSUR are rainfall, snowfall and PET, and the main output are the values of surface runoff into the river cells and the amount of infiltration into the aquifer domain. The direct flow values at the outlet and the two streamgage stations are computed by MODCOU. As the input parameters are given as daily values, the output is daily as well. In the period of simulations there is not a considerable snow input, except for the year 1982; therefore a simple density value of 10% is used in order to compute the snowmelt depth.

A two-step calibration procedure is followed. In the first step, the parameters controlling the partition of water are calibrated. Whereas, in the second step, the discharge coefficients (*Cd*) are determined by assigning different values. The effect of each value on the discharges were observed using RMSE values. As the value of *Cd* decreased the flow peaks attenuated due to the delay introduced. In majority of the years 0.99 gave the best fit and in others the value of 0.20 was more satisfactory. The change in *RMSE* values with different values of *Cd* is shown in Fig. 4.

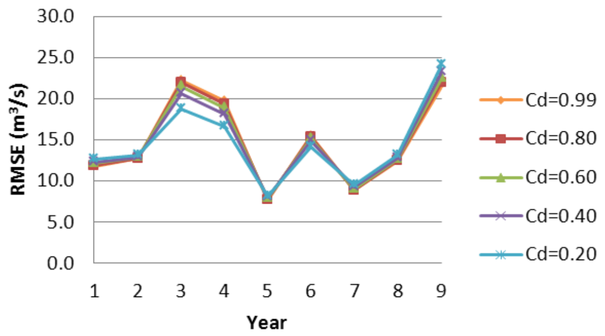

Figure 4: Change in *RMSE* with different values of *Cd*

Conclusion

In this research, the surface hydrology of the Nilüfer river basin in northwestern Turkey is studied. The basin is emphasized in hydrological aspects which result in high seasonal variations in the stream flow. There are two main reasons for these variations. The first is that both dry and wet seasons are observed throughout a year. The second is the elevation difference, which consequently cause the climatic characteristics to vary considerably inside the basin. For the hydrological model MODCOU, all the input data was prepared in GIS environment. The missing precipitation data was computed using two methods and the CCWM gave better results than IDWM.

The direct flow was computed in two ways. The first one is a baseflow separation technique using a two-parameter filter proposed by Eckhardt (2005). Through sensitivity analysis and trial-and-error procedures the two parameters were determined and used to filter the observed discharges. The second way is to simulate the direct flow using the numerical model MODCOU. With the prepared input data, the direct flow in the period between Jan 1, 1973 and Dec 31, 1982 was simulated at the outlet and the streamgage stations EIE321 and DSI344. The simulated direct flow hydrographs are fitted to the ones obtained from filtering process. It was observed that in some years a value of 0.99 gave the best fit, whereas in others 0.20 yielded better results. According to Charbonneau et al. (1977) the optimum value of *Cd* is close to unity unless there is a storage reservoir. Low *Cd* values occurred due to the inconsistencies between the precipitation and discharge data. A possible way to avoid such problems is to install more meteorological stations at different altitudes with varying climatic conditions. Moreover, increasing the number of streamgage stations will be useful in comparing the flow hydrographs during storm events. Nevertheless, on majority of the simulation years, particularly, on the 5th and 7th acceptable fits are obtained between simulated and filtered direct flows.

Acknowledgments

This work was supported by the Commission of Scientific Research Projects of Uludağ University, Project number: UAP(M)-2011/85.

References

Artinyan, E., Habets, F., Noilhan, J., Ledoux, E., Dimitrov, D., Martin, E., & Le Moigne, P., (2008). Modelling the water budget and the riverflows of the Maritsa basin in Bulgaria. *Hydrol. Earth Syst. Sci*. 12, 21–37.

ASCE, (1996). *ASCE, Hydrology Handbook*. American Society of Civil Engineers, 2nd ed., New York.

Başoğlu, S. M., Uyar, Y., Ay, A., & Akova, Y., (2006). Land Cover/Use Determination of Turkey by Using Satellite Images. *Proceedings of 15th Statistics Research Symposium*, Turkish Statistical Institute, 457–473.

Behrangi, A., Khakbaz, B., Jaw, T.C., AghaKouchak, A., Hsu, K., & Sorooshian, S., (2011). Hydrologic evaluation of satellite precipitation products over a mid-size basin, *Journal of Hydrology*, 397, 225-237.

Charbonneau, R., Fortin, J-P., & Morin, G., (1977). The CEQUEAU model: Description and examples of its use in problems related to water resource management / Le modèle CEQUEAU: description et exemples d'utilisation dans le cadre de problèmes reliés à l'aménagement. *Hydrological Sciences Bulletin*, 22(1), 193–202

Eckhardt, K., (2005). How to construct recursive digital filters for baseflow separation. *Hydrol. Process*. 19, 507–515.

Farr, T. G., et al., (2007). The Shuttle Radar Topography Mission. *Rev. Geophys*., 45, RG2004, doi:10.1029/2005RG000183.

GeoCommunity, (2011). Available at: http://www.geocomm.com/

Gomez, E., Ledoux, E., Monget, J.M., & De Marsily, G., (2003). Distributed surface-groundwater coupled model applied to climate or long term water management impacts at basin scale. *Eur Water*, 1/2, 3–8

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., & Jarvis, A., (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978.

Korkmaz, S., Ledoux, E., & Önder, H., (2009). Application of the coupled model to the Somme river basin. *Journal of Hydrology*, 366, 21–34.

Ledoux, E., (1980). Modélisation intégrée des écoulements de surface et des écoulements souterrains sur un bassin hydrologique. *Ph.D. Thesis*, Ecole des Mines de Paris – Université Pierre et Marie Curie, France.

Ledoux, E., Girard, G., & Villeneuve, J.P., (1984). Proposition d'un modèle couplé pour la simulation conjointe des écoulements de surface et des écoulements souterrains sur un basin hydrologique. *La Houille Blanche*, 1/2, 101–110.

Maidment, D.R., (2002). *Arc Hydro: GIS for Water Resources*, ESRI Press, Redlands CA.

Natural Earth, (2011). Available at: http://www.naturalearthdata.com/

NCDC, (2011). Available at: http://www.ncdc.noaa.gov/cgi-bin/res40.pl?page=gsod.html

Teegavarapu, RSV., & Chandramouli, V., (2005). Improved weighting methods, deterministic and stochastic data-driven model for estimation of missing precipitation record. *Journal of Hydrology*, 312, 191–206.

Trabucco, A., & Zomer, R.J., (2009). Global Aridity Index (Global-Aridity) and Global Potential Evapo-Transpiration (Global-PET) Geospatial Database. *CGIAR Consortium for Spatial Information*. Published online, available at: http://www.csi.cgiar.org

Vogt, J.V. et al., (2007). *A pan-European River and Catchment Database*. European Commission - JRC, Luxembourg, (EUR 22920 EN) 120 pp. CCM2 data available at: http://ccm.jrc.ec.europa.eu/