World Academy of Science, Engineering and Technology
International Journal of Geological and Environmental Engineering
Vol:1, No:4, 2007

# Development of Integrated GIS Interface for Characteristics of Regional Daily Flow

Ju Young Lee, Jung-Seok Yang, and Jaeyoung Choi



***Abstract*****—**The purpose of this paper primarily intends to develop GIS interface for estimating sequences of stream-flows at ungauged stations based on known flows at gauged stations. The integrated GIS interface is composed of three major steps. The first, precipitation characteristics using statistical analysis is the procedure for making multiple linear regression equation to get the long term mean daily flow at ungauged stations. The independent variables in regression equation are mean daily flow and drainage area. Traditionally, mean flow data are generated by using Thissen polygon method. However, method for obtaining mean flow data can be selected by user such as Kriging, IDW (Inverse Distance Weighted), Spline methods as well as other traditional methods. At the second, flow duration curve (FDC) is computing at unguaged station by FDCs in gauged stations. Finally, the mean annual daily flow is computed by spatial interpolation algorithm. The third step is to obtain watershed/topographic characteristics. They are the most important factors which govern stream-flows. In summary, the simulated daily flow time series are compared with observed times series. The results using integrated GIS interface are closely similar and are well fitted each other. Also, the relationship between the topographic/watershed characteristics and stream flow time series is highly correlated. .

***Keywords*****—**Integrated GIS interface, spatial interpolation algorithm, FDC.

## I. INTRODUCTION

THE hydrologic study affecting stream daily flow are addressed by numerous journals and text books. The information of stream daily flow is essential for the development of water management and planning, water quality management and environmental conservation. The accuracies at ungauged stations are very important in the view of these objectives. Hence, the hydrological estimation and getting daily flow information at ungauged stations have been addressed by [7], [8], [10], [16] as well as numerous journals. Also, the hydrologic processes affecting stream flow have been mentioned by [1], [3], [4], [9], [11], [12], [14] and [15]. However, these procedures spend a lot of time to obtain daily flow information. Moreover, stream flow analysis must be highly affected by non-homogenous watershed and other characteristics. Then, the Geographic Information System (GIS) is very important tool. The more GIS are developed, the better precise data must be obtained. Especially, integrated GIS interface is to be developed for simulation of continuous daily flow time series. Its tasks include major factors of uncertainty affecting the accuracy of flow estimates including hydrologic processes such as transforming rainfall to stream flow, infiltration, evapotranspiration, surface runoff and other. After all, it is for estimation of regional characteristics.

Jaeyoung Choi is with the KIST. He is now with the Senior Research Scientist, Gangneung, South Korea 210-340 (Corresponding author to provide phone: 82-33-650-7301; e-mail: jchoi@kist.re.kr).

Ju Young Lee is with the KIST (Korea Institute of Science and Technology), Gangneung, 210-340 South Korea (phone: 82-33-650-7302; fax: 82-33-650-7302; e-mail: jyl7318@ kist.re.kr).

Jung-Seok Yang is with the KIST (phone: 82-33-650-7303, e-mail: yjseok@kist.re.kr).

## II. PROCEDURE

### A. Study Area

Yeongwol-kun (or Young Wall Kun) is a region of fidelity with beautiful natural resources in southern Ganwon-do, the middle inland of the Korean peninsula. Fig. 1 shows that it is situated in between 128 degrees 6 minute and 55 minutes of east longitude, and between 37 degrees 2 minute to 24 minute of north latitude. The total area of Yeongwol-kun is 1,127.12 $km^2$. It includes farmland of 100.95 $km^2$, forest land of 962.56 $km^2$, and others take 63.61 $km^2$.

Fig. 2 shows that Yeongwol-kun includes two major rivers which is Han-river and Peng-Chang River. Yeongwol-kun is the source of the Han-river. It creates the headwaters that provide clean water to Seoul, capital of South Korea. About 90% of the 2 major rivers originate as snow or rain in Tae-bak mountain range near northeastern Yeongwol-kun. Recently, South Korea government interests in drought protection plan, waterpower project and allowing the appropriation of water system-wide within this region to meet water needs of customers like farmers and local civilians in the future.

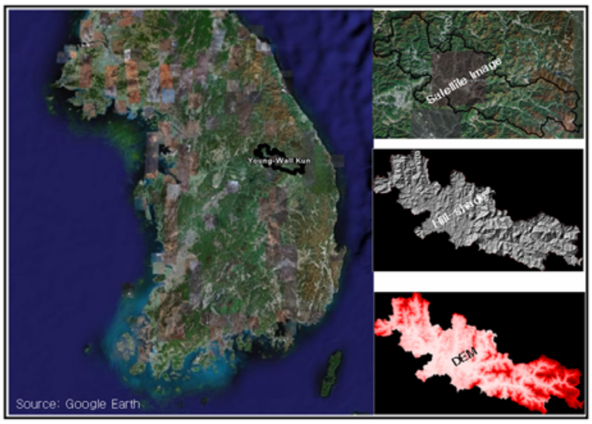



Fig. 1 Yeoungwol gun (Young Wall Kun)

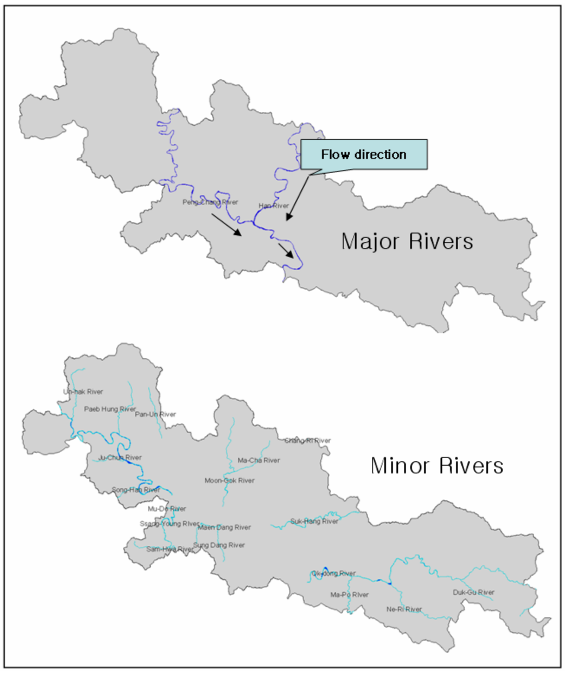


Fig. 2 Major and minor river network

### B. Computation of the Parameter Values of Multiple Linear Regression Using Integrated GIS Interface

Integrated GIS interface is the most convenient tool for simulation of regional daily flow at ungauged stations. It has a very user friendly extension, and because of this it is accessible even to users who are not familiar with GIS and Spatial Interpolation technique. Watershed characteristics part is operated within integrated GIS interface and computes the hydrologic parameter values from digital elevation model (DEM). Moreover it generates geological region files which contain topological information and geo-hydrological parameters to make sure simulation of regional daily flow.

Fig. 3 shows that extracting hydrologic parameter values from digital elevation model (30m) has been accomplished by GIS-SI Hydro Part. The computed hydrologic parameters are independent variables for multiple regression models within Spatial Interpolation algorithm. Projected 30m Digital Elevation (DEM) is the first input for running GIS-SI Hydro Part. The flow direction grid which stores either of the 8 numbers (e.g., 1,2,4,8,16, 32, 64 or 128) is calculated based on 8 point-pour algorithm and Filling Sink/Burn-In process. This creates a raster of flow direction from each cell to its steepest down slope neighbor. The flow accumulation creates a raster which stores in each cell the number of cell flowing to it based on the flow direction raster. Based on the outlet raster and flow direction raster, delineated watershed raster is made. It will have a field Grid-Code which associates it with outlet and stream link. The final step makes geo-hydrological parameter files.

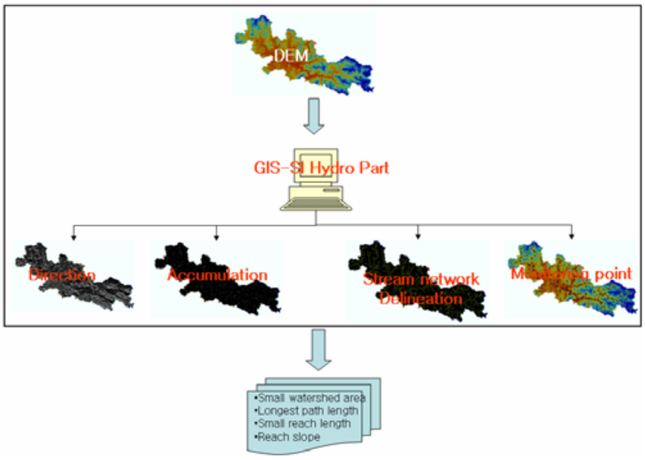


Fig. 3 Procedure for computing hydrological parameters using GIS-SI Hydro Part

### C. Simulating the Long Term Mean Daily Flow Values at Unmeasured Gauge Station Using GIS-SI Meteorological Part

[16], [18], and [17] provide an algorithm for GIS-SI meteorological Part. Spatial interpolation algorithm assumes that flows occurring simultaneously at sites, which are reasonably close to each other and hydrological similar, correspond to similar percentage points on their respective flow duration curves(FDCs) which is the curve that shows the relation between time excess probability and discharge corresponding to the probability. The long term mean daily flow values at unmeasured gage station as destination sites are generated by available FDCs at measured gage station as source sites.

### D. Generation of a Flow Duration Curves (FDCs) at Measure Station (Source Station)

Regional FDCs at source station reflect regional flow variability and is relatively close to regional flow variability at the unmeasured station (destination station). A regional FDCs and discharge at 17 point time excess probability values (0.01, 0.1, 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 99, 99.9, and 99.99 %) are computed at measured gage station.

### E. Generation of Flow Duration Curves (FDCs) and a Continuous Flow Hydrograph at Unmeasured Gauge Station (Destination Station)

The next step is to calculate the regional FDCs based on the normalization. Conversion of FDCs into generation of a continuous flow hydrograph at unmeasured gage station is computed by Spatial Interpolation algorithm. [16] assumed that source and destination flow close proximity to each other on their respective FDCs. Fig. 4 shows the procedure how to generate a continuous flow hydrograph at unmeasured station using spatial algorithm technique.

In Fig. 4, step 1 is an observed hydrograph, FDC in step 2 is calculated from observed hydrograph, and step 3 shows that source FDC find the discharge value associated with the same

% point on the destination station FDC, step 4 shows that FDC at destination station create the interpolated hydrograph in destination station.

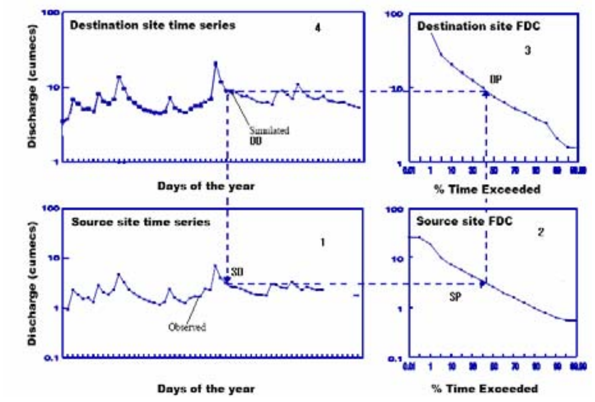



Fig. 4 Long term daily flow generation procedure at unmeasured gauge station (modified after Hughes & Smakthins, 1996)

### *F. Generation of Long Term Daily Flow Using GIS-SI*

Measured station is chosen from geologically closing unmeasured station using GIS-SI Hydro Part. GIS technique provides different weighed factors to each source sites according to geological aspects.

The generation of a continuous mean daily flow is computed by following equation 1;

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$
$$InQ_{mn} = -\beta_0 + \beta_1 InA + \beta_2 In[MDP] \quad (1)$$

where $Q_{mn}$ is continuous mean daily flow(cms), A is drainage area ($m^2$), MDP are Mean daily precipitation (mm) from 1962 to 2003, $\beta_{0,1,2}$ are parameters(weighting factors) for linear log fitting. The dataset of drainage area are obtained from GIS-SI Hydro and Mean Daily Precipitation as grid dataset are obtained from GIS-SI meteorological Part for calculating discharge-area and discharge-precipitation relationships and formula as equation 1. In the Young-Wall Kun (Yeoungwol), the mean daily flow (cms) is well fitted linear log relationship among mean daily flow ($Q_{mn}$) and area (A). Especially, Parameters ($\beta_{0,1,2}$) for linear log fitting are averaged 0.047, 0.82 and 0. 52 and $R^2$ is 97.2.

Equation 2 is used that $Q_{mn}$ at unmeasured station in large region is computed by weighting the area of different precipitation. Meanwhile, $Q_{mn}$ at unmeasured gage station within small region is computed by weighting the area of same precipitation;

$$Q_{mn} = \sum_g w_g \frac{Q^g_{mn}}{A_g} A_s \quad (2)$$

$Q^g_{mn}$ : gage flow
$Q_{mn}$ : flow at unmeasured gage station
$A_g$: gage drainage area
$A_s$: upstream drainage area
$w_g$: weight

To obtain weighting factors, kriging technique is more precise than other method. Kriging defines that it assigns all of weights from measured data at gage station to predict unmeasured data at unmeasured station. Kriging method depends on mathematical and statistical aspects. Its basic concept is autocorrelation and distance. Spatial unmeasured data is simulated by their distance and measured MDP data are computed as autocorrelation response to a function of distance. The equation 2 is derived from the empirical semivariance values of measured data are fitted as semivariogram as measured gage station in this study area.

In sum, Fig. 5 shows how to make the simulation of long-term mean daily flow ($Q_{mn}$) at unmeasured gage site using GIS-SI interface. The process is following like that.

First, it is preparation step for geologic hydrological GIS datasets and meteorological grid datasets using GIS-SI Hydro Part and GIS-SI meteorological Part. Especially, grided mean daily precipitation datasets are obtained from kriging Interpolation technique and mathematical and geo-statistical analysis. After that, long term mean daily flow at unmeasured station is computed and exported to ASCII files. For validation, the Nash-Sutcliffe coefficient of efficiency [5] is. NS coefficient of efficiency has been widely used to evaluate the comparison of hydrologic simulated and observed value. The coefficient of determination ($R^2$) describes that the proportion of the total variance in observed data that can be explained by the simulation.

$$Deviation = \frac{\sum Q_{sim} - \sum Q_{obs}}{\sum Q_{obs}} \times 100,$$

$$Efficiency = 1 - \frac{\sum (Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - \bar{Q}_{obs})^2} \quad (3)$$

$$R^2 = \left[ \frac{\sum_{i=1}^{N} (Q_{obs} - \bar{Q}_{obs})(Q_{sim} - \bar{Q}_{sim})}{\left[\sum_{i=1}^{N} (Q_{obs} - \bar{Q}_{obs})^2\right]^{0.5} \left[\sum_{i=1}^{N} (Q_{sim} - \bar{Q}_{sim})^{0.5}\right]} \right]^2 \quad (4)$$

$Q_{obs}$ : Observed Q
$Q_{sim}$ : Simulated Q

$\bar{Q}_{sim}$ : Simulated mean Q

$\bar{Q}_{obs}$ : Observed mean Q

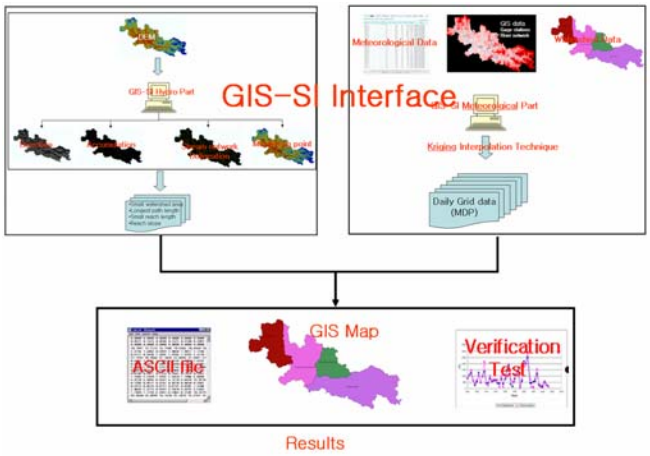



Fig. 5 The procedure for simulation regional daily flow at unmeasured gauge station using GIS-SI

## III. Results and Discussion

In Fig. 6, the Young-Wall Kun has 7 gage stations. Their stream-flow data are recorded rarely as well as are released on website from 1962 to 2003. In this paper, the Young-Wall 1 and Young-Wall 2 stations are chosen as source station and test station. That means that the Young-Wall 2 station is assumed to be unmeasured gage station for comparison between simulated and observed data. Also, it is simulated by near weighted neighbor station as Young-Wall 1.

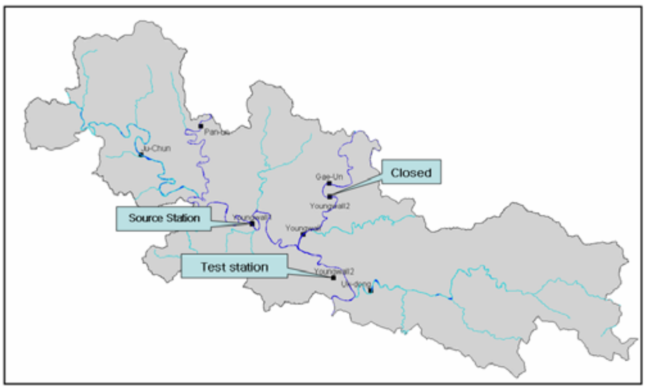


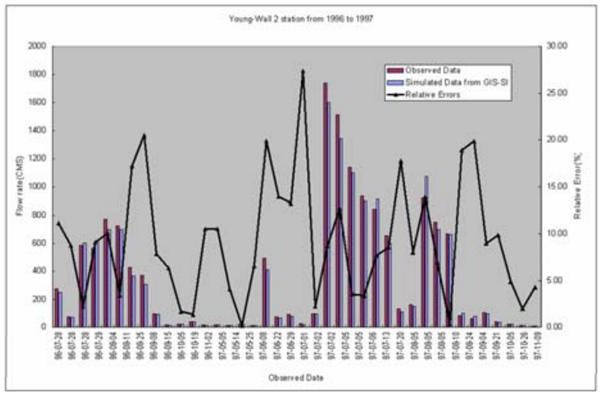

Fig. 6 The test and weighted neighbor station

Table I shows the values of independent variables (Mean annual precipitation (MAP), Drainage area) and the information of hydrological parameters. The Young-Wall Kun includes two rainfall stations. However the value of two stations can not be represent all over Young-Wall Kun rainfall data. And point data show the rainfall amount within a confined area. Thus point data should be converts to grid dataset. After all, Interpolation technique, kriging, is very useful method for expended estimation.

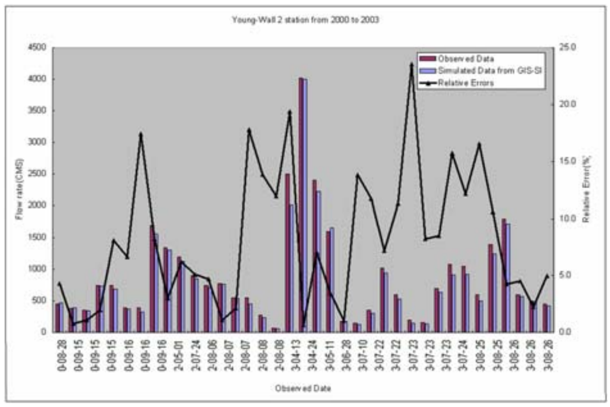
TABLE I
Gauge Stations and Hydrological Parameters

| Stations_NM | Watershed_NM | Drainage Area in GIS-SI | Mean annual rainfall(mm/yr) | Record (Streamflow) |
|---|---|---|---|---|
| Youngwall | Upper Nam Han Rr | 80 $m^2$ | 1394.5 | 1962-2003 |
| Gae-Un | Upper Nam Han Rr | | | 1965-1990 |
| Pan-Un | Peng-Chang | 321.78 $m^2$ | 1164 | 1993-2003 |
| Youngwall1 $^{1)}$ | Peng-Chang | | | 1991-2003 |
| Uk-Dong | Chung Ju Dam | 422.67 $m^2$ | 1231 | 2001-2001 |
| Youngwall2 $^{2)}$ | Chung Ju Dam | | | 1991-2003 |
| Ju-Chun | Ju- Chun | 302.67 $m^2$ | 1232 | 1965-2001 |
| 1) | Source station | | | |
| 2) | Test Station | | | |

Fig. 7 shows that High relative errors are estimated in low stream flow rate comparing to high stream flow rate. But it is not important for civil engineers to control flood and to design hydraulic structures because the purpose of stream flow simulation at unmeasured gage station is to manage a high flow rate. In addition, overall relative errors are 9.23% in Fig. 7(a) and are 8.2 % in Fig. 7(b) and it is not bad.



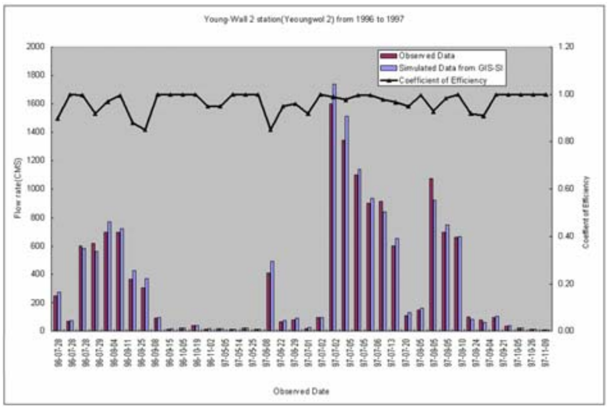

(a)



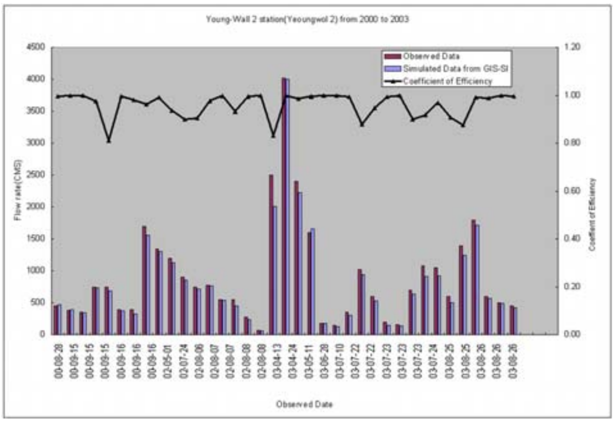

(b)

Fig. 7 Stream flow rate and relative errors for long term mean daily flow(cms) from 1996 to 1997 and from 2000 to 2003

Fig. 8 shows that observed stream flow data and simulated data are fairly similar. The reason is that NS overall average coefficients of efficiency (NSCE) are 0.97 from 1996 to 1997 and 0.95 from 2000 to 2003. [16] refer if NSCE is higher than 0.6, the simulation result can be regarded satisfactory. Moreover, Fig. 9 shows that coefficients of determination are 0.96 (1996~1997) and 0.97 (2000~2003). $R^2$ is the criteria of consonance between the observed value and the simulated value. It ranges from 0 (poor model) to 1 (perfect model). The $R^2$ shows that the simulated value is satisfactory to observed value. The NSCE and $R^2$ are criteria for validation test. The NSCE is improved over for model evaluation purposes because it is sensitive to differences in the observed and simulated means and variances as equation 2 and 3. Therefore, NSCE is sensitive to extreme values as is $R^2$.




(a)




(b)

Fig. 8 Stream flow rate and coefficient of efficiency for long term daily fliow(cms) from 1996 to 1997 and from 2000 to 2003

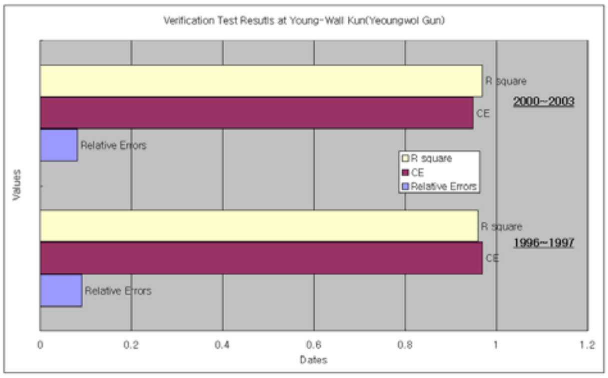



Fig. 9 Average value of validation for model verification

This purpose of the GIS-SI interface is computing conveniently and exactly the long term means daily/monthly/annual hydrographs at unmeasured gage stations. Most of all, its results are more exact than manual computational method like labor intensive work. It is composed of two categories; one is GIS-SI hydro part which is performing to get independent variables which are drainage area, hydrologic parameters, MAP/MMP/MDP, and determination of weighting factors from gage station near unmeasured gage station. The other part is technique to simulate hydrographs at unmeasured gage station using historical data at gage station. Through this paper, the simulated hydrographs are highly correlated with and fitted to the observed hydrographs and GIS-SI interface is a very convenient tool for analysis of regionalization.

## IV. Conclusion

The purpose of this paper primarily intended to provide the generated data for engineering guidelines more precise and make easily hydrological mapping tool. In the South Korea, many stream-flow gage station sites are not working, and they can not provide stream-flow datasets enough for development of the Probable Maximum Flood (PMF) for use in the evaluation of proposed and existing dams and other impounding structures. Also, many citizen and farmers interest in drought protection plan, waterpower project, and allowing the appropriation of water system-wide and water right within this region to meet water needs of customers in the future. Then, GIS-SI interface is a convenient tool for all users. Therefore it is the role of the planning and supporting document (guidelines) for water planners.

## Acknowledgment

This study was supported by the KIST-Gangneung Institute Funding (No. 2Z02920). We think P.E. Ganeshi Krishinamurshy for creating tool and data collection

## REFERENCES

[1] A.Kessler, and M.H.Diskin, "The efficiency function of detention reservoirs in urban drainage systems," Wat. Resour. Res. Vol 27(3), 1991, pp.253-258

[2] D.A. Hughes, and V.Y. Smakhtin, "Daily Flow Data Time Series Patching or Extension: a spatial interpolation approach based on flow duration curve," *Hydrol. Sci. J*. vol 41(6), 1996, pp. 851-871

[3] F.F.Snyder, "Synthetic unit graphs," *Trans*. AGU 19, 1938, pp.447-454

[4] I. Rodriguez-Iturbe and J.B.Valdes, "The geomorphologic structure of hydrologic response," *Wat. Resour.Res*. Vol 15(6), 1979, pp.1409-1420

[5] J.E. Nash, and J.V.Sutcliffe, "River flows forecasting through Conceptual Models Part 1-a Discussion of Principles", *J. Hydrol*. Vol 10, 1970, pp. 282-290

[6] K.E. Hendrickson, *Handbook of Texas Online*(Book Style), Brazos River, Texas State Historical Association, 2001

[7] M.Mimikou., and S.Kaemaki, "Regionalization of Flow Duration Characteristics," *J.Hydrol*. vol 82, 1985, pp.77-91

[8] N.Fennessey and R.M. Vogel, "Regional Flow Duration Curves for Ungaged Sites in Massachusetts, J.Wat. Resour. Plan. Mang, ASCE, Vol 116(4), 1990, pp.530-549

[9] R.A.Wooding, "A hydraulic model for the catchment-stream problem," *J. Hydrol.* Vol 3, 1965, pp.254-267

[10] R.G.Quimpo., A.A.Alejandrino., and T.A.McNally, "Regionalized Flow Duration Curves for Philippines," *J.Wat.Resour.Plan.Manag*. Vol 109(4), 1983, pp.320-330

[11] R.K.Linsley., and J.L.Paulhus, *Hydrology for Engineers*(Third Edn). McGraw-Hill, New York, USA, 1982

[12] S.K.Jenson and J.O.Domingue, "Extracting topographic structure from digital elevation data for geographical information system analysis," *Photogram.Engng.Remote Sens*. Vol 54(11), 1988, pp.1593-1600

[13] USEPA, *Ecological condition of estuaries in the Gulf of Mexico. EPA 620-R-98-004*(book style), U.S.Environmental Protection Agency, Office of Research and Development. National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, Gulf Breeze, Florida, 1999, pp. 80.

[14] V.P.Singh and J. Li, "Identification of reservoir flood wave models," *J.Hydrol. Res*. Vol 31(6), 1993, pp.811-824

[15] V.T.Chow., D.R.Maidment ., and L.W.Mays., Applied Hydrology(Book Style), McGraw-Hill, New York, USA, 1988, pp.252-257

[16] V.Y.Smakhtin, D.A. Hughes, and E. Creuse-Naudin,"Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa," *Hydrol. Sci. J*. vol 46(6), 1997, pp.919-936.

[17] V.Y.Smakhtin, and B.Mass, "Continuous daily hydrograph simulation using duration curves of precipitation index," *Hydrological Processes*, vol 14, 2000, pp.1083-1100.

[18] V.Y.Smakhtin, "Generation of natural daily flow time-series in regulated rivers using a non-linear spatial interpolation technique, regulated rivers," *Research & Management*, vol 15, 1999, pp.311-323.

[19] V.Y.Smaktin, "Estimating daily flow duration curves from nonthly stream flow data," Water SA, 2000, vol 26(1), pp.13-18.