 

# Assessment of Uncertainty in flood forecasting using downscaled rainfall data

Mohammad Karamouz[1]

Sara Nazif[2]

Mahdis Fallahi[3]

Sanaz Imen[4]

**Abstract:**

Flood is one of the most important natural disasters which causes extensive loss of life and properties every year around the world. By forecasting the hydrograph of probable floods in each year, their damages could be reduced by implementing plans, decisions, and actions. Before flood forecasting, it is necessary to forecast rainfalls. By forecasting rainfall, the amount of runoff can be estimated. So it can be determined whether a severe flood would happen or not. The simulated flood hydrograph is affected by uncertainties in rainfall forecasting that must be considered when flood prevention plans are developed.

In this study, a long lead flood forecasting model is developed, considering the uncertainties in forecasting. This model uses long lead predicted rainfalls on a daily basis. The predicted rainfalls must be disaggregated to smaller time and space spans (say 6 hours) in order to be used in rainfall-runoff models. So, predicted rainfalls by a GCM (General Circulation Model) have been downscaled using a regression based statistical method.

Then HEC-HMS software has been used to develop a rainfall-runoff model of the basin. Using downscaled predicted rainfalls and the rainfall-runoff model of the basin, the predicted flood hydrograph is determined. The effect of uncertainties in rainfall forecasting has been considered in flood forecasting. In daily precipitation, uncertainties have been assessed by comparing means and variances of precipitation, monthly mean dry and wet spells lengths and their confidence intervals and cumulative frequency distributions (CDFs) of predicted floods. The proposed model for flood forecasting has been applied to the Kajoo River located in South Baloochestan region, in the south-eastern Iran.

**Keywords**: uncertainty, runoff forecast, downscaling, rainfall-runoff model

---

1 Professor, School of Civil Engineering, University of Tehran, Tehran, Iran, Karamouz@ ut.ac.ir

2.Ph.D. Student, School of Civil Engineering, University of Tehran, Tehran, Iran, snazif@ut.ac.ir

3 M.Sc. Student, School of Civil Eng., Amirkabir Univ. of Tech., Tehran, Iran, mahdis@aut.ac.ir

4.M.Sc. Student, School of Civil Engineering, University of Tehran, Tehran, Iran, sanazimen@ut.ac.ir

 

## 1. Introduction

Floods cause a lot of damages and deaths around the world each year. Using appropriate strategies to face them can decrease these losses considerably, and this needs forecasting the amount of flood and explicated flood peak as soon as possible. Long-lead forecasting of rainfall, in at least a daily scale, is necessary for this purpose. Day et al. (1995) used the ESP model for runoff prediction. In their approach, a hydrological model using current instream flow, rain and temperature time series is developed which calculates the sum of probable flow hydrographs. Then a statistical model determines statistical distribution of rainfall in future time periods. Therefore the short-lead and long-lead flood with different probabilities is predicted. Ingram et al. (1998) investigated the advanced method for hydrological prediction in real time. The most important characteristic of this system is correcting the real time prediction by using the meteorological Predictions. Karamouz and Zahraie (2004) have investigated a method for predicting run-off in arid and semiarid areas. They have chosen the best ARIMA model for seasonal rainfall forecasting and have improved it considering large scale signal effects as well as ENSO by a Fuzzy based model. Wilby and Harris (2006) have investigated a method, in which climate variables for a specific station can be obtained from the AOGCM[1] simulated variables.

Even if global climate models (GCMs) in the future are run at high resolution there will remain the need to 'downscale' the results from such models to individual sites or localities for impact studies. Downscaling enables the construction of climate change scenarios for individual sites at daily time-scales, using grid resolution GCM output. Wilby et al. (2002) investigated SDSM model for spatial and temporal downscaling the long lead predictions meteorological variables as well as rain and temperature using statistical methods. Harpham and Wilby (2005) predicted the precipitation for different zones of England using different models such as SDSM, radial neural network and multi layers neural network. The result of this research show that all of these methods are capable of predicting precipitation; however in different zones their capabilities are different. Massah(2006) has been downscaled long-lead predictions of rain and temperature in Zayanderud River basin, using kridging and inversed distance weighting methods. Ekstorm et al. (2005) assessed the effects of selecting the probability distributions of rain and temperature variables on runoff distribution. The

[1] Atmosphere- Ocean General Circulation Model

 

results show that the different probability function for rain and temperature variables can affect the flow distribution function.

In this paper a method has been developed for long-lead flood forecasting. In this method, rainfall has been forecasted on a daily scale, by long-lead climatological signals using a statistical downscaling model, SDSM. Results of this model are used as inputs of the rainfall-runoff model that has been developed by HEC-HMS software and flood hydrograph. Uncertainties of flood forecasting are considered by fitting the probability distribution to the predicted amount of floods.

## 2. Study area and data

The study area is the Kajoo River sub-basin, with an area of 3659 $km^2$, located in the South Baloochestan region, in south-eastern Iran (Figure 1). Because of the limited carrying capacity of the Kajoo River, yearly floods cause damages to agricultural lands and rural areas. Therefore flood forecasting in this area is vital for development of schemes and contingency plans.

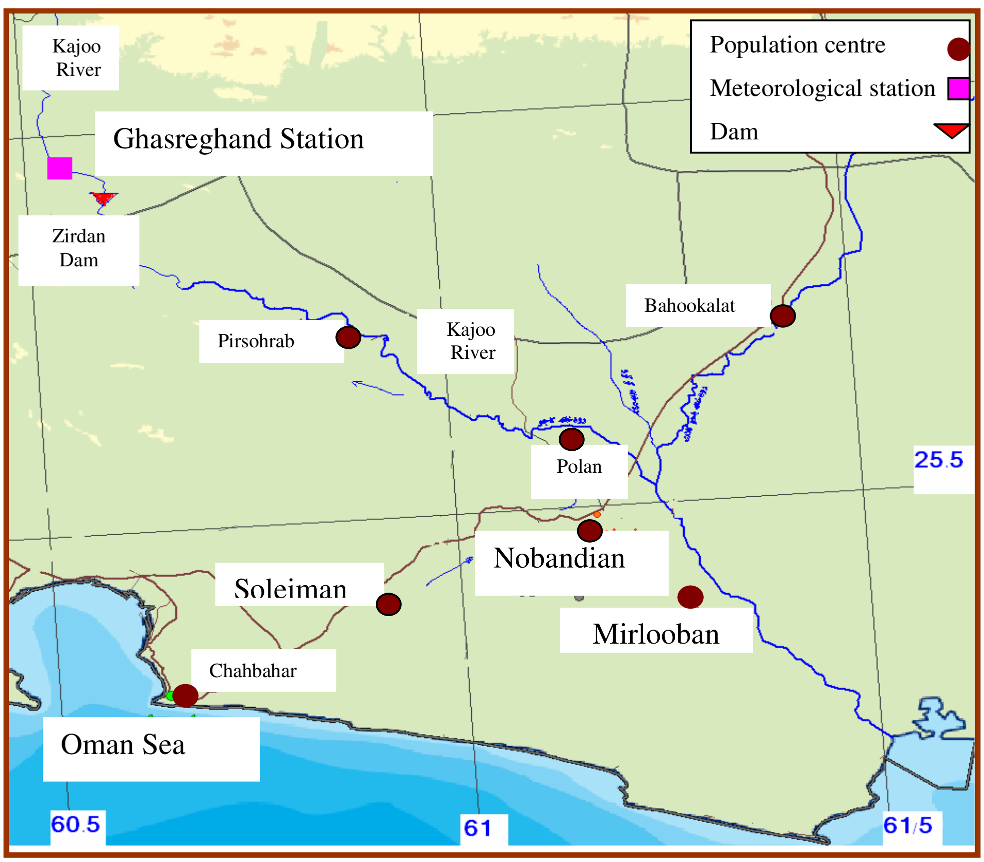


Figure1: Kajoo River in the South Baloochestan watershed

The study area includes 12 sub-basins, their hydrological characteristics, area, main river channel length and time of concentration are presented in Table 1 and discrete sub-basins of the study area are shown in Figure 2.

In the past fifteen years there have been six considerable floods in this region, which occurred in the years of 1991,1992,1995,1997 and 2005. Although these floods did not cause any deaths, they left behind extensive damages. This shows the vulnerability of the study area to floods and the significance of flood forecasting with adequate lead time.

Table1-the characteristic of sub-basins in the study area

| Sub-basin | A ($km^2$) | L (km) | $T_C$ (hr) |
|---|---|---|---|
| 1 | 558 | 43.2 | 5.75 |
| 2 | 177 | 29.6 | 3.71 |
| 3 | 70 | 19.1 | 2.52 |
| 4 | 435 | 22.9 | 2.63 |
| 5 | 331 | 29.4 | 3.59 |
| 6 | 462 | 57.7 | 6.81 |
| 7 | 402 | 43.8 | 5.84 |
| 8 | 528 | 27 | 2.88 |
| 9 | 52 | 13.8 | 1.68 |
| 10 | 286 | 18 | 1.87 |
| 11 | 80 | 12.4 | 1.73 |
| 12 | 264 | 29.8 | 3.95 |

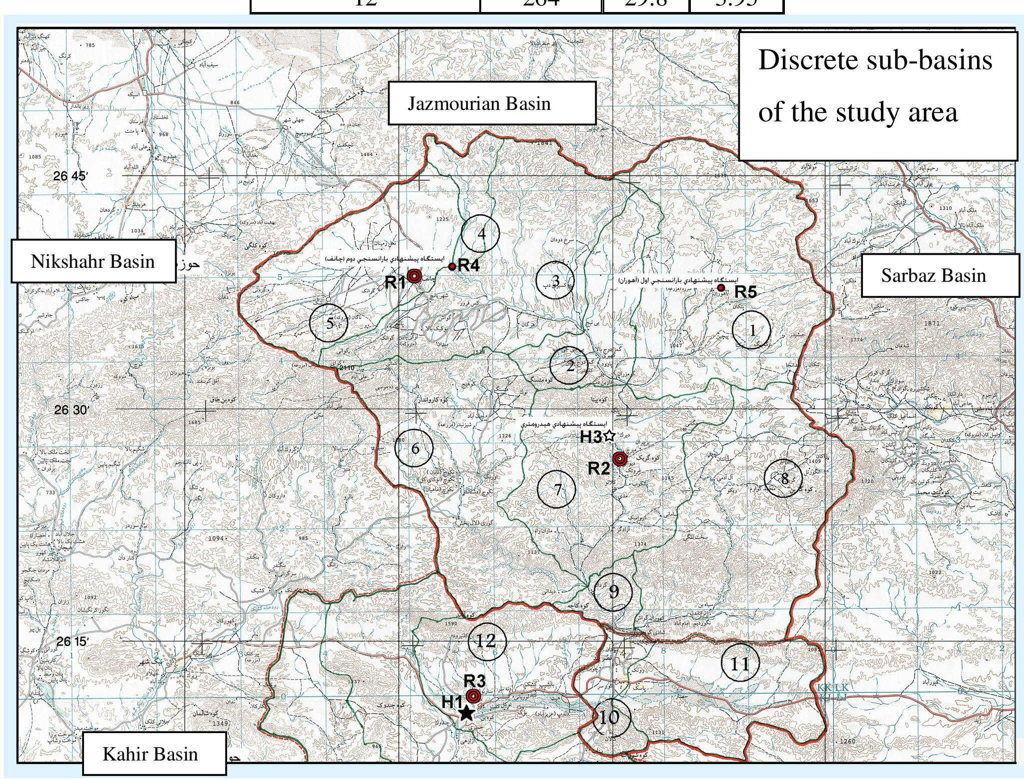


**Figure 2: Discrete sub-basins of the study area**

A meteorological station inside the basin namely Ghasre-Ghand is used for downscaling rainfall, which is the only available station in the region. Twenty four years (1976–1999) daily precipitation data has been used as predictands. Hydrographs of observed historical floods at Ghasre-Ghand station during the study period have been gathered for rainfall-runoff model testing. Observed large-scale NCEP (National Centre for Environmental Prediction) reanalysis atmospheric variables for the same period have been used as predictors. Figure 3 shows the location of the study area in the library of large-scale NCEP predictors.

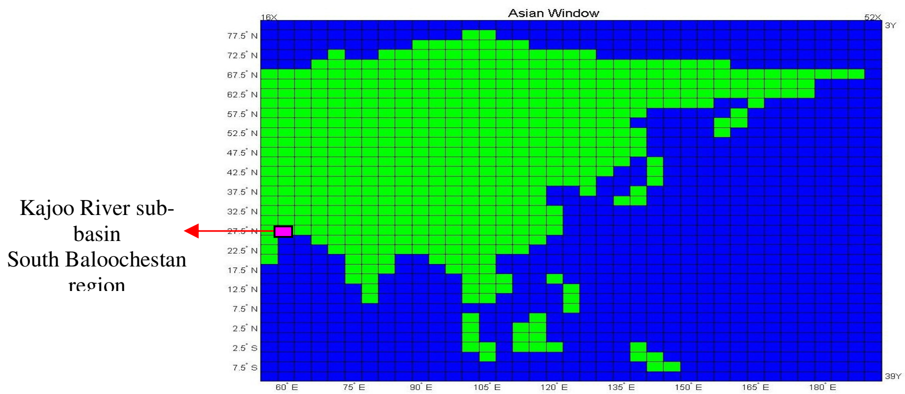


Figure 3: Location and nomenclature of the Baloochestan in the NCEP data archive

**3- SDSM model**

Even though global climate models in the future are run at a high resolution, the need will remain to "downscale" the results from such models to individual basins or localities for impact studies (Department of the Environment, 1996)**.**

The general theory, limitations and practice of downscaling, have been discussed in detail by Giorgi and Mearns, 1991; Wilby and Wigley, 1997 and Xu, 1999. Downscaling techniques are divided into two main groups: Statistical and dynamical. In situations where low–cost, rapid assessments of localized climate change impacts are required, statistical downscaling (currently) is more effective. In this study, SDSM has been used for long-lead rainfall prediction for an individual site at daily time–scales, by using grid resolution of GCM output.

During downscaling with the SDSM, a multiple linear regression model is developed using selected large-scale predictors and local scale predictands such as temperature and precipitation. The parameters of the regression equation are estimated using the efficient dual simplex algorithm. Large-scale relevant predictors are selected using correlation analysis, partial correlation analysis and scatter plots, and also considering physical sensitivity between selected predictors and predictand for the region. Ideally, predictor candidates should be physically and conceptually sensible with respect to the rainfall and accurately modeled by GCMs. For precipitation downscaling, it is also recommended that the selected predictor should contain variables describing atmospheric circulation such as thickness, stability and moisture content. In practice, the choice of predictor variables is constrained by data availability from GCM archives.

Structure and operation of SDSM includes five distinct tasks: (1) preliminary screening of potential downscaling predictors; (2) assembly and calibration of SDSM(s); (3) synthesis of ensembles of current weather data using observed predictor variables; (4) generation of ensembles of future weather data using GCM-derived predictor variables; (5) diagnostic testing/analysis of observed data and climate change scenarios.

**4- Rainfall downscaling**

At first, daily precipitation data has been transformed by the fourth root function to better fit normal distribution. The correlations between different combinations of available predictors and daily precipitation have been calculated to find the most appropriate combination. It must be noted that correlations have been calculated in the winter because of the occurring floods in this season. Finally the combination of 3 predictors was selected: 1) relative humidity at 850 hPa height, 2) near surface specific humidity and 3) near surface relative humidity. Karamouz et al. (2006) were selected 4 Predictors for this region, here one of these predictors has been omitted to accurate the model for predicting the September's rainfall.

This combination has been selected because of its maximum correlation with daily precipitation. Correlation coefficients between selected predictors and daily precipitation have been presented in Table 2. This table also reports the P-value between the predictors and precipitation that helps to identify the amount of explanatory power for each predictor. The correlation statistics and P values indicate the strength of the association between two variables. Higher correlation values imply a higher degree of association. Smaller P_values than 0.05 indicates that the result can be statistically significant but not certainly of practical significance, in the other hand the high P value would indicate that the predictor–predictand correlation is likely to be due to chance.

The model has been calibrated with precipitation data in years 1976-1990 and validated for the remaining available data (1991-1999).

**Table 2: The Results of variables screened for the Baloochestan area**

| Predictor | Correlation coefficient between weather – variables and precipitation | P value |
|---|---|---|
| relative humidity at 850 hPa height | 0.42 | 0.002 |
| Near surface specific humidity | 0.45 | 0.000 |
| Near Surface relative humidity | 0.40 | 0.000 |

With the above procedure, the weather generator was used to downscale observed (NCEP) predictors, and generate scenarios to downscale GCM predictors representing the current climate. For this purpose 100 ensemble data of rainfall have been generated. Simulated and observed mean rainfall comparing for validation period in Figure 4, shows the ability of the model for rainfall prediction. Some deviation could be observed specially in months April and June which are not significant in this study, because only winter rainfall has been considered.

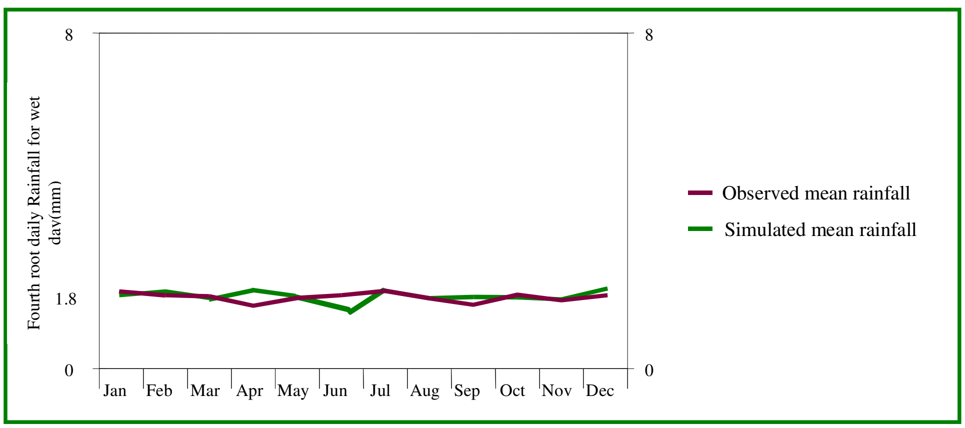


**Figure 4: Observed and simulated mean daily rainfall in each month by SDSM model during the validation period (1991-1999).**

After developing rainfall downscaling model using NCEP signals for year 1977-1999, it is necessary to test model ability for future predictions. For this purpose, one of predicted scenarios for future climate variation such as HadCM3 (Second Hadley Centre Coupled Ocean-Atmosphere GCM) is used as model input signal and rainfall is predicted. Figure 5 shows that downscaling produces almost similar mean daily rainfall in each month under selected observed signals (NCEP) and predicted signals

(HadCM3) for the study area. It must be noted that in winter months, the values obtained from NCEP and HadCM3 are almost the same. Errors of rainfall prediction as well as mean and maximum daily rainfall and wet spells, during the validation period, have been presented in Table 3 using RMSE[1] and MAE[2] error indices. Predicted wet spells defined as number of rainy days in a given month, are also important because of showing the pattern of rainfall occurrence. The accuracy of prediction of wet spells also has been considered as shown in Table 3. All of errors mentioned in Table 3 are less than 22%, which shows acceptability of results.

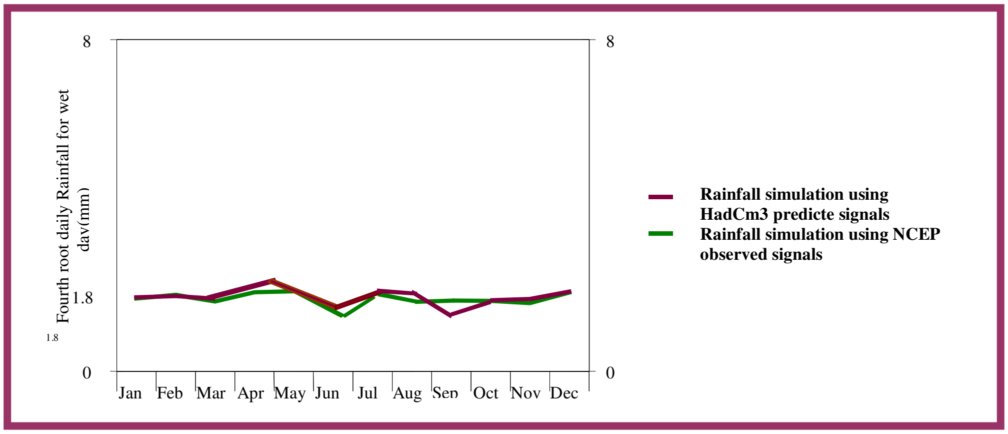


**Figure 5: Monthly mean daily rainfall totals at Baloochestan for the current climate downscaled rainfall simulated by NCEP data and GCM (HadCM3) predictors validation period (1991–1999)**

**Table3- Errors of different scenarios in predicting rainfall as mean and maximum daily rainfall and wet-spell**

| scenario | variable | MAE (%) | RMSE(%) |
|---|---|---|---|
| NCEP | Mean rainfall | 11 | 12 |
| | Maximum rainfall | 12 | 22 |
| | Wet spell | 13 | 18 |
| HadCm3 | Mean rainfall | 11 | 12 |
| | Maximum rainfall | 19 | 22 |
| | Wet spell | 13 | 18 |

In the year 2005, a major flood occurred in the study area. The daily rainfalls in that period have been forecasted to simulate the observed flood in the winter of 2005 and calibrate the rainfall runoff model. Thus, the generated scenario at previous step has been implemented for a second time using HadCM3 predictors to predict winter 2005

[1] Root Mean Square Error
[2] Mean Absolute Error

 

rainfall. The maximum predicted rainfall in each ensemble, has been used as rainfall that causes flood.

## 5-Rainfall-Runoff model

A rainfall-runoff model has been developed using HEC-HMS software. Sub-basins of the study area have been characterized and modeled in HEC- HMS software using the SCS method. Hydrograph of January 1, 2005 flood have been used for model calibration. Hyetograph of rainfall has been estimated using central pattern of SCS hyetographs because of monotonous and steady intensity during the rain time in the region. SCS curve method has been used for runoff estimation and an average CN of 85 has been estimated for the basin.

Using the central pattern of SCS hyetographs, which is almost similar to rainfall patterns of study area, the observed and simulated Hydrograph after rainfall-runoff model calibration are shown at Figure 6.

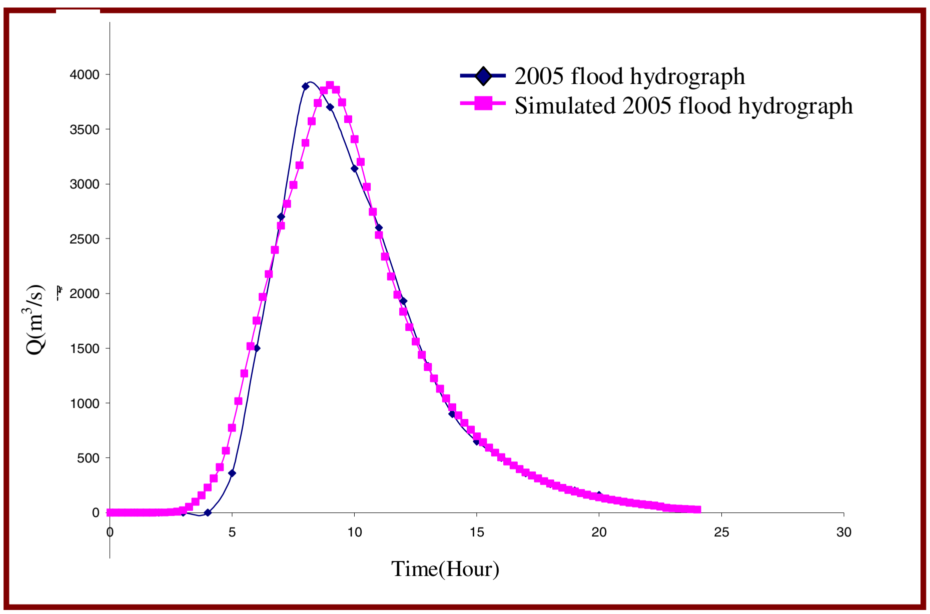


**Figure 6: Comparison between simulated and observed flood hydrograph**

The forecasted flood hyetograph has been interred in HEC- HMS model and the amount of flood is estimated. Observed and predicted flood hydrographs for January 2005 are shown in Figure 7. It must be noted that in this Figure only peak of hydrographs are important (two hydrographs are not for same time) which there is acceptable difference, about 4%, between them. For other historical floods, the same procedure has been done and differences in peak discharges of simulated and observed hydrographs are less than 10%.

**6- Uncertainty considerations**

There are uncertainties in flood prediction the main source is perhaps uncertainties in rainfall prediction. Probabilistic approach is used for dealing with uncertainties in flood prediction. Through this approach, decision makers will be able to determine the amount of flood with different probability of occurrence.

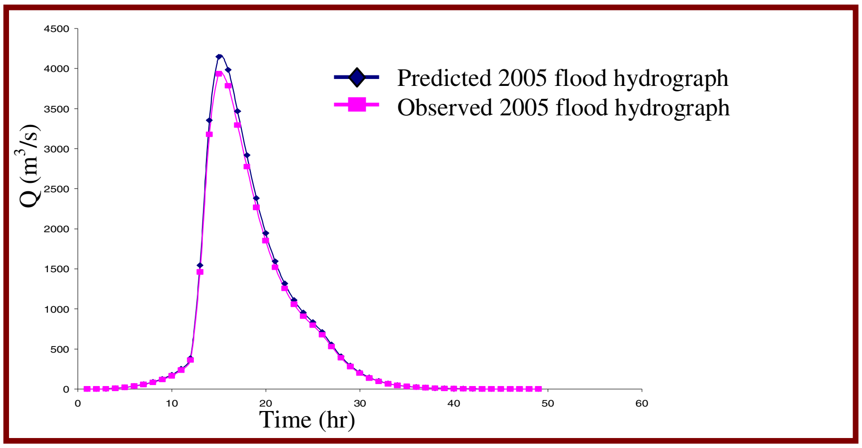


**Figure 7: Comparing of predicted and observed peak of 2005 flood**

In this study, for any observed major rainfall/flood a set of 100 ensemble rainfall are generated. Then, for each of these a cumulative probability function is fitted to their maximum rainfall (which considered as flood). Probability of occurrence of the observed rainfalls should be well within the range of predicted rainfalls. The normal distribution has been determined as the best probabilistic distribution for rainfall data at Ghasre-Ghand station. In Figure 8, as an example, the developed probability distribution of 1991 flood has been shown. For 6 major events studied here, the range is between 50 to 70%. This shows that using rainfall with 50% probability of occurrence from the derived distribution base of 100 ensembles could be an acceptable alternative for this region rainfall prediction. Using this probability distributions of predicted floods, the decision maker can setup a risk level and determine a probable value for incoming rainfall prediction. For example, if Figure 8 is used by a decision maker with 20% risk taking attitude looking for 80% probability of occurrence, a rainfall of 30 mm is predicted.

**7- Conclusion**

In this paper, different large scale predictors have been examined for winter rainfall forecasting and then surface specific humidity, near surface specific humidity, relative humidity at 850 hpa and 850 hpa vorticity predictors which have highest correlations with daily precipitation, have been selected. Winter precipitation forecasted by GCM scenarios have been downscaled for the South Baloochestan region in Iran, using large scale NCEP signals and the statistical downscaling model, SDSM.

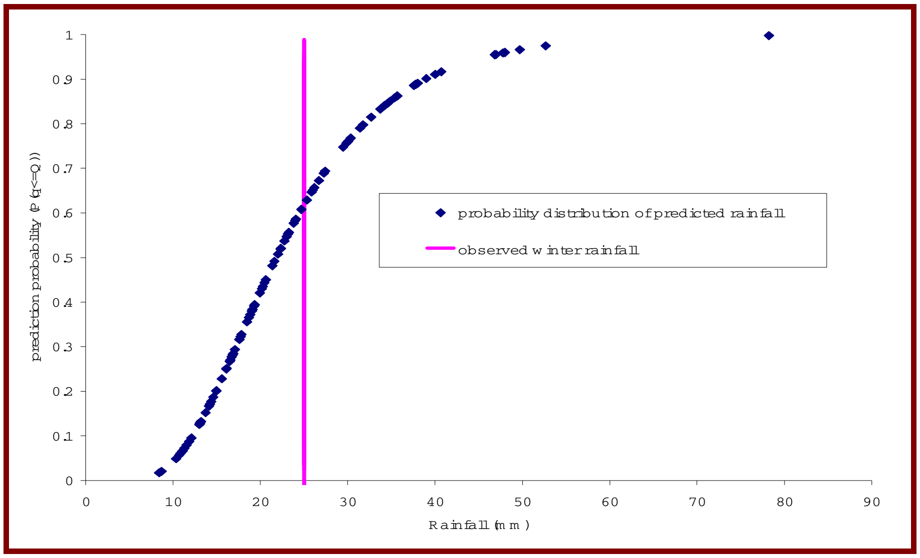


**Figure 8: Risk of rainfall predictions for year 1991 flood**

A rainfall-runoff model using HEC-HMS model has been developed to convert forecasted rainfall to runoff in the study area. Flood that occurred in the winter of 2005 has been used for model calibration. By analyzing the downscaled rainfall data in each year, the rainfall that causes the floods has been characterized and interred to the rainfall-runoff model forecasts the maximum probable peak of flood. The Results shows that the simulation hydrograph closely matches the observed hydrograph and this model can be used as an effective tool for flood forecasting. In this study uncertainties in flood prediction, have been considered by developing 100 ensemble data for rainfall prediction and fitting a probabilistic distribution to them, as real rainfalls are usually in the range of 50-70 percent of predicted rainfall. The decision makers can use this probable predictions to face with floods, with their desirable risk

**Acknowledgments**

This study was a part of flood management project entitled "Flood plane zoning downstream of Kajoo and Kariani dams and designing flood warning system".

## 8-References

Day G. N., "Extended Streamflow Forecasting Using NWSRFS', Journal of Water Resources Planning and Management, 1995.

Department of the Environment, (1996), Review of the potential effects of climate change in the United Kingdom. HMSO, London, PP. 247.

Ekstrom, M., Hingray, B., Mezghani, A. and Jones, P. D. (2005). Regional climate model data used within the SWURVE project 2: addressing uncertainty in regional climate model data for five European case study areas. Hydrology and Earth Systems Science (in press).

Giorgi, F., Mearns, L.O., (1991). Approaches to the simulation of regional climate change. A review. Review of Geophysics 29, PP. 191–216.

Harpham C., Wilby R., 2005, Multi-Site Downscaling of Heavy Daily Precipitation Occurrence and Amounts, J. Hydrology, Vol. 312, P.P. 235-255

Ingram J. J., Fread D. L. and Larson L. W.,(1998), "Improving Real Time Hydrological Service in the USA, PartI: Ensemble Generated Probabilistic Forecast", British Hydrological Society.

Karamouz, M. and B. Zahraie, (2004) "Seasonal Streamflow Forecasting Using Snow Budget and ENSO Climate Signals: Application to Salt River Basin in Arizona", Jornal of Hydrologic Engineering Vol. 9(6), PP.1312-1325.

Karamouz, M., M. Fallahi, S. Nazif ,"Flood forecasting using downscaled long-lead precipitation forecast", The Second Korea-Iran Joint Workshop on Climate Modeling, South Korea, 21-22 Nov 2006.

Massah, A., (2006), "Risk Assessment for climatological variety and it's effects on water resources, case study : Zayanderud Basin", PHD thesis, Tarbiat Modares University.

Wilby R.L. and Harris, I. (2006) framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK.Water Resources Research (in press)

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J. and Wilks,D.S., (1998). "Statistical downscaling of general circulation model output: A comparison of methods." Water Resources Research 34: 2995-3008.

Xu, C.Y. (1999). From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. Progress in Physical Geography 23, PP. 229–249.