# WARNING OF DROUGHT BY USING ARIMA MODEL IN THE NAKDONG RIVER BASIN IN KOREA

**Ngo Van Quan**[1]

**Abstract:** *Drought tendency is significantly increasing and highly negative influencing causing of extreme impact of the over global climate change in the future and especial at river basins that is facing severe challenge for all water resources managers in the world. The early realization of drought event can help managers to have drought mitigation strategies, which is significantly able to damage reducing in river basins. Therefore, the main objective of this study is to provide methodology to assess and warn of drought by using of the Autoregressive Integrated Moving Average (ARIMA) model. Based on the rainfall in the basin, the Standardized Precipitation Index (SPI) series were determined by using fitting a probability density function, then the model was applied to simulate drought using this index series in the river basin. The simulated results obtained by using from the best model's parameters after they were calibrated and compared with the actual data, and it showed reasonably good agreement with the actual data. So the models can be used to warn droughts with reasonably accuracy. Besides, the results of the study also was indicated occurrence drought levels based on SPI index that were shown that the years 1990, 1992 represented severely dry and double of consecutive years of 2000-2001, 2003-2004, 2008-2009 represented extremely dry conditions in basin. Through results of the study provide useful methodology for managers to have drought mitigation strategies, which are significantly able to damage reducing, and suitable policies of water resources management in region.*

**Keywords:** ARIMA model, SPI index, Assessment, Nakdong river basin.

## 1. INTRODUCTION

Drought is a normal, recurrent feature of climate, and is one of the world's costliest natural disasters, causing an average US$6–8 billion in global damages annually, and affecting more people than any other form of natural catastrophe [7]. It occurs almost everywhere, although its features vary from region to region. Defining drought is therefore difficult; it depends on differences in regions, needs, and disciplinary perspectives. In the most general sense, drought originates from a deficiency of precipitation over an extended period of time, resulting in a water shortage for some activity, group, or environmental sector. Whatever the definition, it is clear that drought cannot be viewed solely as a physical phenomenon. Korea shows a typical seasonal pattern in its climate. Four seasons are developed due to the change of its temperature and rainfall amount. The dry season includes most of spring, fall, and winter, and months of June, July, August and September are generally included in the wet season in Korea. More than 60% of the total annual precipitation is concentrated in the wet summer season, so flooding has become an annual event in the Korean Peninsula. Monsoon lasts about a month from mid-June to July, and several typhoons hit the Korean peninsula from late August to September. The recent trends of mean and extreme precipitation with significantly change in Korea that is based on studied by Jung et al [6]; the projected changes of drought is also shown in drought occurrence under future global warming from multi-model, multi-scenario [15]. Convective storms are also frequently developed locally to cause flash floods [3]. On the other hand, a long dry spell continues until the monsoon season begins.

[1] *Thuyloi University*

According to Yoo et al [17], almost every 2 years, the Korean Peninsula suffers from a shortage of rain, needed mostly for agricultural use, which sometimes turns into a very deadly on lasting more than 2 years. However, these recent frequent droughts have severely stressed water supply systems and the community that depends on them. Drought management has therefore become an important issue especially in south-eastern Korea. The forecasting aspect, which is very important from the point of view of drought preparedness and early warning, is still fraught with great difficulty [13]. Nevertheless, Beran et al [1] suggested that it may be possible to forecast the probable timings of inception and termination of droughts reasonably well over a short period such a month or a season. Others, the drought occurrence probabilities and risks of dependent hydrological processes [2]. There are several methods that have been used in the past as the drought assessment tools and many attempts have been tried to quantify the drought, most of which have been based on shortage of precipitation data analysis; the drought assessment tools such as measurement of shortage of streamflow, reduced levels of water storage, and Drought Indices. Of these, DIs are widely used for drought assessment [5,7, 12,16]. Moreover, the methods used to prediction of drought by the SPI by Paulo et al [11]; On the use of SPI for drought intensity assessment [8]. Basin-scale stream-flow forecasting using the information of large-scale atmospheric circulation phenomena [9]. However, in order to assess drought levels and to forecast drought that is not only need based on a standard index but also need using model to drought simulation in river basin. Therefore, the determination time series of SPI index for time scale and to develop valid ARIMA models to forecast and simulate SPI series, and both of these are to assessment and forecasting of drought that is the goal in this study.

## 2. STUDY AREA AND DATA

The Nakdong River basin is situated in the monsoon region of South Korea (35–37° N, 127–129° E) as shown figure 1.

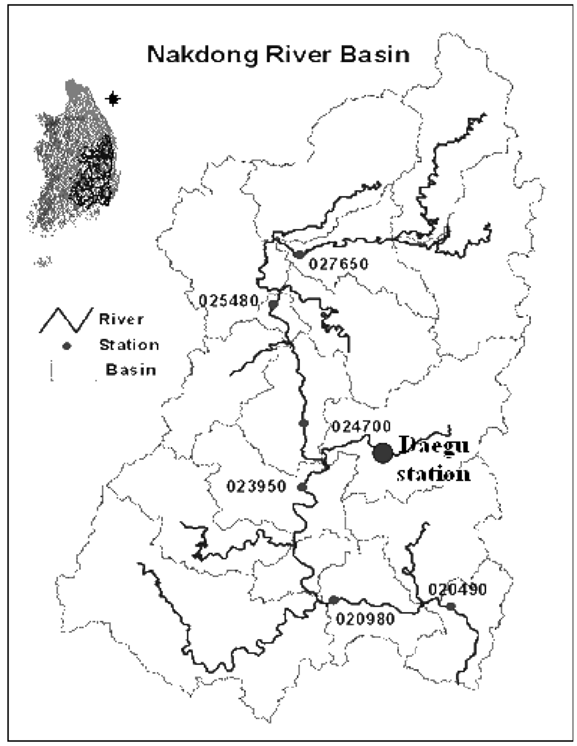


*Figure1. Location of study are*

This region is characterized by heavy rainfall in the monsoon season in early summer from June to August. This is one of four major river systems in Korea, serves as an important water resource for the southeastern area. The river drains an area of 23,817 km2 and length of the main stream is over 525 km. The annual mean precipitation across the river basin is about 1200mm, but more than 60% of the annual rainfall is concentrated during the summer season (June–August). The mean air temperature is 2.2°C during the coldest month (January) and 25.9°C in the warmest month (August). Currently, about 7 million people reside within the basin and more than 13 million people intake a drinking water from the river. Data used in this study include meteorological data during 1980-200 was used in this study which include data of precipitation at weather stations in Nakdong river basin. In this study the Daegu station data was applied to simulate for basin

## 3. METHODOLOGY

The ARIMA models presented in this paper are based on the Standardized Precipitation (SPI) as drought index. In the present study, SPI is used as drought index, due to its several advantages as following. (i) Plotting a time series of months against SPI gives a good indication of the drought history of a particular station. (ii) The primary reason is that SPI is based on rainfall alone, so that drought assessment is possible even if other hydro-meteorological measurements are not available. (iii) The SPI is also not adversely affected by topography because it do not need to input this factor, and adversely affected by topography changes is expressed in rain stations. (iii) The SPI is defined over various timescales; this allows it to describe drought conditions over a range of meteorological, hydrological and agricultural applications. (iv) The fourth advantage of SPI comes from its standardization, which ensures that the frequencies of extreme events at any location and on any time scale are consistent, which are discussed by Hayes et al [4]. The SPI is defined over various timescales; this allows it to describe drought conditions over a range of meteorological, hydrological and agricultural applications. The Standardized Precipitation Index (SPI) is an index based on the probability of precipitation for any time scale, and can be computed for different time scales, can provide early warning of drought and help assess drought severity levels. Drought classification levels based on SPI that show in table 1. SPI determination method, the Standardized Precipitation Index (SPI) is computed by fitting a probability density function to the frequency distribution of precipitation summed over the time scale of interest. The SPI was developed by McKee et al [10-11] with the purpose of identifying and monitoring local droughts. This is performed separately for each month and for each location in space. Each probability density functions in then transformed into a standardized normal distribution. The gamma distribution is defined by its probability density function as given in Equation (1):

$$g(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad \text{for } x > 0 \qquad (1)$$

where: $\alpha>0$ is a shape factor, $\beta>0$ is a scale factor, and $x>0$ is the amount of precipitation. $\Gamma(\alpha)$ is the gamma function. The resulting parameters are then used to find the cumulative probability of an observed precipitation event for the given month or any other time scale using equation (2)

$$G(x) = \int_0^x g(x)dx \qquad (2)$$

Since the gamma function is undefined for $x = 0$ and a precipitation distribution many contain zeros, the cumulative probability becomes as given in equation (3):

$$H(x) = u + (1-u)G(x) \qquad (3)$$

where: u is the probability of zero precipitation. The cumulative probability, H(x) is than transformed to the standard normal random variable Z with mean zero and variance once, which is the value of SPI as shown this equations (4) and (5):

$$Z = SPI = -\left(k - \frac{c_0 + c_1k + c_2k^2}{1 + d_1 + d_2k^2 + d_3k^3}\right) \quad for\ 0 < H(x) \le 0.5 \qquad (4)$$

$$Z = SPI = +\left(k - \frac{c_0 + c_1k + c_2k^2}{1 + d_1 + d_2k^2 + d_3k^3}\right) \quad for\ 0 < H(x) \le 0.5 \qquad (5)$$

where:

$$k = \sqrt{\ln(\frac{1}{H(x)^2})} \quad for\ 0 < H(x) \le 0.5$$

$$k = \sqrt{\ln(\frac{1}{1-H(x)^2})} \quad for\ 0.5 < H(x) < 1$$

And, $c_0 = 2.515517$, $c_1 = 0.802853$, $c_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, $d_3 = 0.001308$

**Table 1. Drought classification levels based on SPI index**

| SPI Values | Drought levels |
|---|---|
| 2.0+ | extremely wet |
| 1.5 to 1.99 | very wet |
| 1.0 to 1.49 | moderately wet |
| -0.99 to 0.99 | near normal |
| -1.0 to -1.49 | moderately dry |
| -1.5 to -1.99 | severely dry |
| -2 and less | extremely dry |

ARIMA model is Auto-Regressive (AR) models be effectively coupled with Moving Average (MA) models to form a general and useful class of time series models called autoregressive moving average models. In these, (AR) to order p and MA to order q and operates on $d^{th}$ difference of the time series. ARIMA method, the acronym ARIMA stands for "Auto-Regressive Integrated Moving Average" Lags of the differenced series appearing in the forecasting equation are called "auto-regressive" terms, lags of the forecast errors are called "moving average" terms, and a time series which needs to be differenced to be made stationary is said to be an "integrated" version of a stationary series. ARIMA models are, in theory, the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing and logging. However, in fact that the time series data are most of non-stationary characteristic, so how can it become the stationary time series, and how to identify this time series data is stationary time series. ARIMA model is classified as an "ARIMA(p,d,q)" model, where: p is the number of autoregressive terms, d is the number of differences, and q is the number of lagged forecast errors in the prediction equation. The first (and most important) step in fitting an ARIMA model is the determination of the order of differencing needed to stationarize the series. Second, the steps of identifying the orders of AR(p) or MA(q) terms are essential, so after a time series has been stationarized by differencing, the next step in fitting an ARIMA model is to determine whether AR(p) or MA(q) terms are needed to correct any autocorrelation that remains in the differenced series. By looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the differenced series. Finally, values of time series are forecasted by model that is selected with good agreement. All the special mathematic functions of ARIMA method was made by coding source based on MATLAB language to simulate the time series data. After methods of SPI and ARIMA model were performance, the main goal of study is obtained based on the overall implementation that is combined both of SPI index and ARIMA model to drought simulating in study area. SPI index were determined based on monthly rainfall time series of 30 years with periods of 1980-2009 at Daegu station, and the SPI is applied on basis with scale of 12 months. Firstly, the standardized precipitation index (SPI) index values were determined from rainfall time series, which are then used as a drought indicator for drought forecasting due to its many advantages compared to other drought indices. Secondly, the Autoregressive Integrated Moving Average (ARIMA) model in drought will simulate of the standardized precipitation index that has been investigated using the correlation methods of Box and Jenkins and the AIC structure selection criteria. Finally, The ARIMA model is developed to forecast drought to give reasonably good results for SPI series of 12-months that is selected for suitable candidate model.

## 4. Results

The results determined for SPI12 as shown in Figure 2. Results of the SPI12 assessment revealed that the years of 1990, 1992 represented severely dry ($-1.5 < SPI < -2$) and double of consecutive years of 2000-2001, 2003-2004, 2008-2009 represented extremely dry conditions with ($-2 < SPI < -3$) in SPI index. Then, this time series model development consists of three stages identification, estimation, and forecasting check. The mode identified for SPI12 is ARIMA with period from 1980 to 2009. Model implementation obtained eight of models as ARIMA (p,d,q) = (0,1,1); ARIMA(p,d,q) = (0,1,2); ARIMA(p,d,q) = (1,1,0); ARIMA (p,d,q) = (1,1,1); ARIMA(p,d,q) = (1,1,2); ARIMA (p,d,q) = (2,1,0); ARIMA(p,d,q) = (2,1,1); ARIMA (p,d,q) = (2,1,2).

**Table 2.2. The result of model parameters**

| Models | AIC | RSquare | Model selected |
|---|---|---|---|
| ARIMA(p,d,q) = (0,1,1) | -0.6752 | 0.53 | |
| ARIMA(p,d,q) = (0,1,2) | -1.1573 | 0.70 | |
| ARIMA(p,d,q) = (1,1,0) | -1.1410 | 0.69 | |
| ARIMA(p,d,q) = (1,1,1) | -1.1721 | 0.71 | |
| ARIMA(p,d,q) = (1,1,2) | -1.2652 | 0.76 | Candidate models |
| ARIMA(p,d,q) = (2,1,0) | -1.1880 | 0.72 | |
| ARIMA(p,d,q) = (2,1,1) | -1.1865 | 0.72 | |
| ARIMA(p,d,q) = (2,1,2) | -1.2180 | 0.74 | |

Finally, after summary and results of ARIMA model calibration, a model was fitted to determine the best model out of these candidate models as ARIMA (p,d,q)=(2,1,2) with the minimum of Akaike Information Criterion (AIC) condition as shown in Figure 3, and alternative ARIMA models were identified by considering the ACF graph of the SPI series as shown in Figure 4, the normal probability plot and scatter of actual SPI values versus forecasted values shown in Figure 5, 6. Results showed a fairly good agreement between observations and forecasts, as it has also been confirmed by the values of some performance indices as shown in Figure 7.

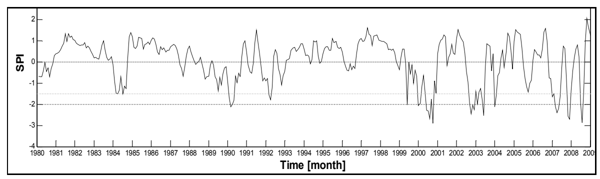


*Figure 2. SPI of time scale 12 months at Daegu station*

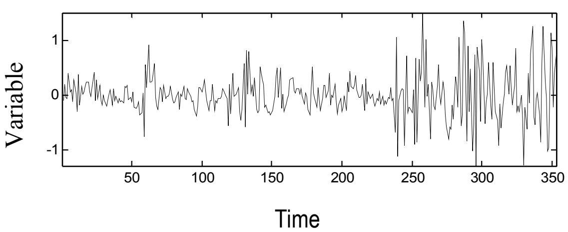


*Figure 3. Result simulation for residuals from ARIMA(2,1,2) for SPI*

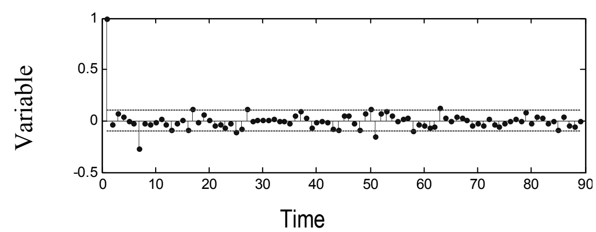


*Figure 4. Result simulation for autocorrelation of residuals (ACF) of SPI time series*

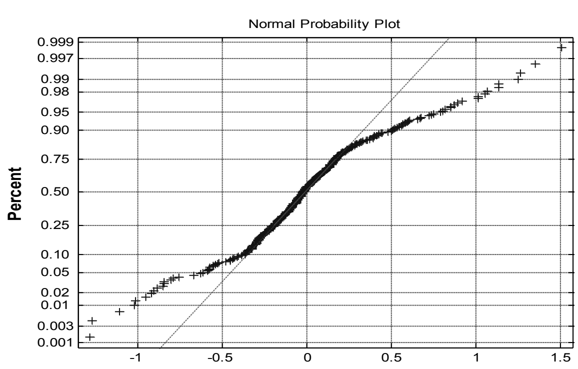


*Figure 5. Normal probability plot of the residuals SPI actual with forecasted using best ARIMA models*

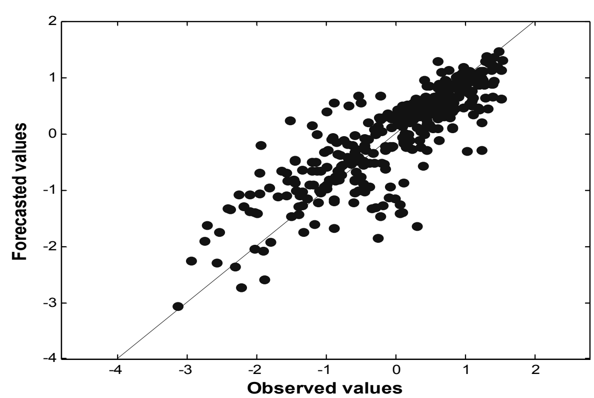


*Figure 6. Compared of SPI actual with forecasted using best ARIMA models*

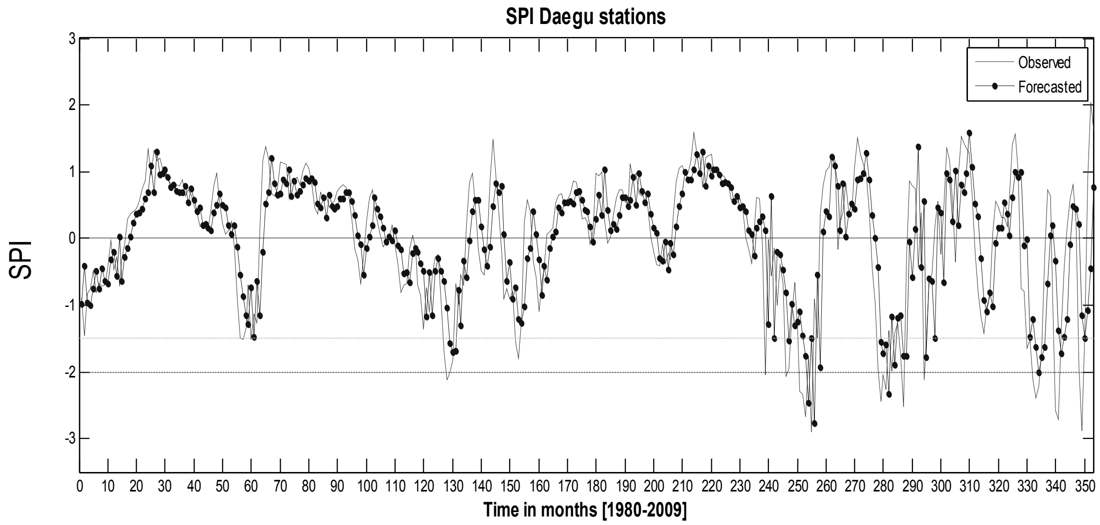


*Figure 7. SPI index simulated by using the best ARIMA models*

**5. Conclusion**

Drought is a natural hazard temporarily affecting almost every region in the world. The main target of this study is to provide method to warn and techniques development for drought characterization and for enhanced management of water resources systems during drought periods. Like many other river basins forecasting drought is very importance for water resources planning and management in Nakdong river basin. This study was focused on drought warning using SPI as a drought indicator. The SPI is used due to its lot of advantages over other drought indices and it used to quantify most type of drought. Results of the SPI12 assessment show that the years 1990, 1992 represented severely dry and double of consecutive years of 2000-2001, 2003-2004, 2008-2009 represented extremely dry conditions in basin. The ARIMA model developed to warn drought to give reasonably good results for SPI series of 12-months that is selected for suitable candidate model of ARIMA(p,d,q)=(1,1,2). Therefore, it is recommended that this model can be used in this and other hydro-meteorologically similar basin for forecasting SPI series of multiple time scales to know the drought severity in study area. Through results of the study provided useful methodology to simulate and realize drought levels, which are significantly able to damage reducing, and the sustainable water resources management and policy in region.

**REFERENCES**

[1]. Beran MA, Rodier JA (1985). Hvdrological aspects of drought. UNBSCO-WMO, Studies and Reports in Hvdrology no. 39, UNESCO, Paris, France.

[2]. Chung CH, Salas JD (2000). Drought occurrence probabilities and risks of dependent hydrological processes. Journal of Hydrologic ASCE, 5: 259-268.

[3]. Han YH, Byun HR (1994).On the existence of seasonal drought in the Korean Peninsula. Journal of Korea Meteorology Society, 30: 457-467.

[4]. Hayes MJ (2003). Drought indices, National Drought Mitigation Center, University of Nebraska-Lincoln [http://www.drought.unl.edu/whatis/indices.htm#spi], (accessed February 24, 2009)

[5]. Heim RR (2002). A Review of twentieth-century drought indices used in the United States, Bulletin of the American Meteorological Society, 8: 1149-1165.

[6]. Jung IW, Bae DH, Kim GS (2010). Recent trends of mean and extreme precipitation in Korea. International Journal of Climatology, 31: 359-367.

[7]. Keyantash J, Dracup JA (2002). The quantification of drought: An analysis of drought indices, Bull. Journal American Meteorology Society, 83 (8): 1167-1180.

[8]. Kumar MN, Murthy CS (2009). On the use of Standardized Precipitation Index (SPI) for drought intensity assessment. Journal of Meteorological Applications, 16: 381-389.
[9]. Maity R, Kumar DN (2008). Basin-scale stream-flow forecasting using the information of large-scale atmospheric circulation phenomena. Journal of Hydrological Processes 22 (5): 643-650.
[10]. McKee TB, Doesken NJ, Kleist J (1993). The relationship of drought frequency and duration to time scales, in Proceedings of the 8th Conference on Applied Climatology, Anaheim, California, USA, 179-184.
[11]. McKee TB, Doesken NJ, Kleist J (1995). Drought monitoring with multiple time scales, in Preprints, 9th Conference on Applied Climatology, 15-20 January, Dallas, TX, USA, 233-236.
[12]. Morid S, Smakhtin V, Moghaddasi M (2006). Comparison of seven meteorological indices for drought monitoring in Iran. International Journal of Climatology 26 (7): 971-985.
[13]. Panu US, Sharma TC (2002). Challenges in drought research: Some perspectives and future directions. Journal of Hydrological Sciences 47: 19-30.
[14]. Paulo AA, Pereira LS (2007). Prediction of SPI drought class transitions using Markov chains. Journal of Water Resources Management 21 (10): 1813-1827.
[15]. Sheffield J, Wood EF (2008). Projected changes in drought occurrence under future global warming from multi-model, multi-scenario. IPCC AR4 simulations, Climate Dynamics, 31: 79-105.
[16]. Smakhtin VU, Hughes DA (2007). Automated estimation and analyses of meteorological drought characteristics from monthly rainfall data, Journal of Environmental Modelling & Software, 22 (6): 880-890.
[17]. Yoo C, Ryoo S (2003). Analysis of drougth return and duration characteristics at Seoul. Journal of Korean Water Resources Association, 36: 561-573.

**Tóm tắt:**

**NGHIÊN CỨU CẢNH BÁO HẠN HÁN SỬ DỤNG MÔ HÌNH ARIMA CHO LƯU VỰC SÔNG NAKDONG CỦA HÀN QUỐC**

*Xu hướng hạn hán ngày một tăng và tác động tiêu cực của nó ngày một lớn do tác động cực đoan của sự biết đổi khí hậu trên toàn cầu trong tương lai, đặc biệt trên các lưu vực sông và điều đó đang là thách thức nghiêm trọng lớn cho tất cả các nhà quản lý tài nguyên nước trên thế giới đang phải đối mặt. Việc nhận biết và đánh giá được mức độ hạn hạn sớm hơn có thể giúp nhà quản lý có chiến lược trong việc giảm thiệt hại một cách đáng kể trong lưu vực. Vì thế, mục đích chính trong nghiên cứu này là cung cấp phương pháp đánh giá và cảnh báo sớm hạn hán bởi việc sử dụng mô hình "Tự hồi quy tích hợp trung bình trượt (ARIMA)". Trên cơ sở số liệu mưa trong lưu vực, chuỗi chỉ số dáng thủy chuẩn (SPI) được xác định, sau đó mô hình đã áp dụng chuỗi chỉ số này đánh giá và cảnh báo hạn hán tại lưu vực sông. Những kết quả cảnh báo sớm được sử dụng từ việc chọn bộ thông số cho mô hình tốt nhất sau khi mô hình đã được kiểm định và so sách với số liệu thực tế, và đã chỉ ra hợp lý với số liệu thực đo. Do đó mô hình có thể sử dụng được sử dụng để đánh giá và cảnh báo hạn hán một cách chính xác và hợp lý. Thêm vào đó, kết quả của nghiên cứu cũng đã chỉ ra được mức độ của hại hán rằng tại các năm 1990, 1992 xảy ra hạn nghiêm trọng, tại nhóm của hai năm liên tiếp như 2000-2001, 2003-2004, 2008-2009 đã xảy ra mức độ cực hạn trong lưu vực nghiên cứu. Kết quả nghiên cứu cung cấp một phương pháp hữu ích cho người quản lý có cái nhìn chiến lược để giảm thiểu hạn hán, điều đó có thể giúp giảm thiệt hại do hạn gây ra một cách đáng kể, và đưa ra các chính sách hợp lý quản lý tài nguyên nước trên các lưu vực sông.*

**Từ khóa:** Mô hình ARIMA, chỉ số SPI, đánh giá, dự báo lưu vực sông Nakdong.

---

*Người phản biện:* **TS. Hoàng Thanh Tùng** *BBT nhận bài: 12/2013*

*Phản biện xong: 15/9/2014*