



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공 학 석 사 학 위 논 문

Copula 함수 기반의 베이지안
네트워크와 가뭄예보를 결합한
안동댐의 계절별 유입량 예측



2018년 2월

부 경 대 학 교 대 학 원

토 목 공 학 과

김 광 훈

공 학 석 사 학 위 논 문

Copula 함수 기반의 베이지안
네트워크와 가뭄예보를 결합한
안동댐의 계절별 유입량 예측

지도교수 이 상 호

이 논문을 공학석사 학위논문으로 제출함.

2018년 2월

부 경 대 학 교 대 학 원

토 목 공 학 과

김 광 훈

김광훈의 공학석사 학위논문을 인준함.

2018년 2월 23일



주 심 공학박사 김 수 한 (인)

위 원 공학박사 이 종 섭 (인)

위 원 공학박사 이 상 호 (인)

목 차

	page
List of Tables	ix
List of Figures	x
Abstract	xiii
제1장 서론	1
1.1 연구 배경 및 목적	1
1.2 연구 동향	3
1.3 연구 내용 및 범위	4
제2장 확률분포형 결정	6
2.1 확률분포형 이론	6
2.1.1 Lognormal 분포	6
2.1.2 Gamma 분포	7
2.1.3 Gumbel 분포	8
2.1.4 Weibull 분포	8
2.1.5 Gaussian 분포	9
2.2 매개변수 추정	9
2.3 적합성 검정	11

목 차(계속)

	page
제3장 베이지안 네트워크	14
제4장 Copula	19
4.1 Copula의 정의	19
4.2 Sklar의 정리	21
4.3 Copula 함수의 종류	22
4.4 Copula 함수 기반의 베이지안 네트워크	27
제5장 가뭄지수를 이용한 가뭄예보	28
5.1 가뭄의 정의	28
5.2 가뭄지수	29
5.2.1 가뭄지수의 선정과 정의	29
5.2.2 가뭄지수 산정	32
5.3 가뭄 예보	34
제6장 Copula 함수 기반의 베이지안 네트워크를 이용한 유입량 예측	36

목 차(계속)

	page
6.1 연구 대상 댐	36
6.1.1 연구 대상 댐 선정	36
6.1.2 안동댐의 제원 및 현황	37
6.2 Copula 함수 기반의 베이지안 네트워크 구축 및 적용	40
6.2.1 Copula 함수 기반의 베이지안 네트워크 구축	40
6.2.2 Copula 함수 기반의 베이지안 네트워크를 이용한 유입량 예측 결과 및 평가	48
제7장 SSI를 이용한 계절별 가뭄 예보	56
7.1 안동댐의 계절별 SSI 산정	56
7.2 SSI를 이용한 안동댐의 계절별 가뭄 예보 결과	58
제8장 Copula 함수 기반의 베이지안 네트워크와 가뭄 예보를 결합한 계절별 유입량 예측	61
8.1 안동댐의 가뭄 단계별 계절 유입량 예측 곡선	61
8.2 베이지안 네트워크만 사용한 계절별 유입량 예측과 결합예측의 결과 비교	67

목 차(계속)

	page
제9장 결론 및 향후 연구	74
9.1 결론	74
9.2 향후 연구	76
참고문헌	77



List of Tables

	page
Table 2.1 The threshold values of the K-S test according to significance level and sample size	13
Table 4.1 The copula functions	24
Table 5.1 Classification of drought stages for the range of SSI	31
Table 6.1 Project data of Andong Dam	38
Table 6.2 The D_n values of the K-S test	47
Table 6.3 The correlation coefficient of each season' s inflow to that of prior season	47
Table 6.4 The errors of seasonal inflow forecasting without drought forecast	55
Table 7.1 The correlation coefficient of each season' s SSI to that of prior season	58
Table 8.1 Seasonal inflow forecasting errors	72

List of Figures

	page
Fig. 1.1 Schematic diagram of seasonal inflow prediction applying a Bayesian network with drought forecasting	5
Fig. 3.1 A conceptual diagram of Bayesian network model	16
Fig. 4.1 Gaussian and Student' s - t Copula	23
Fig. 4.2 Clayton, Frank, and Gumbel Copula	23
Fig. 5.1 Probability density function of summer SSI conditioned on a spring SSI of -1	35
Fig. 6.1 Monthly water supply planning of Andong Dam	39
Fig. 6.2 Seasonal inflow of Andong Dam	41
Fig. 6.3 Histogram against the five distributions of the spring inflow volumes during the training period of 1977-2010	42
Fig. 6.4 Histogram against the five distributions of the summer inflow volumes during the training period of 1977-2010	43
Fig. 6.5 Histogram against the five distributions of the fall inflow volumes during the training period of 1977-2010	44

List of Figures (Cont.)

	page
Fig. 6.6 Histogram against the five distributions of the winter inflow volumes during the training period of 1977–2010	45
Fig. 6.7 Probability density function of summer inflow volume conditioned on a spring inflow volume of $100 \times 10^6 \text{ m}^3$	49
Fig. 6.8 Spring inflow forecasting curve	50
Fig. 6.9 Summer inflow forecasting curve	51
Fig. 6.10 Fall inflow forecasting curve	52
Fig. 6.11 Winter inflow forecasting curve	53
Fig. 6.12 The result of seasonal inflow forecasting without drought forecast	55
Fig. 7.1 Seasonal SSI for Andong Dam	57
Fig. 7.2 The result of seasonal drought forecast	60
Fig. 8.1 Example of estimating representative inflow corresponding to lower bounds of SSI ranges of drought stages	62
Fig. 8.2 Spring inflow forecasting curves conforming to drought stages	63

List of Figures (Cont.)

	page
Fig. 8.3 Summer inflow forecasting curves conforming to drought stages	64
Fig. 8.4 Fall inflow forecasting curves conforming to drought stages	65
Fig. 8.5 Winter inflow forecasting curves conforming to drought stages	66
Fig. 8.6 The results of seasonal inflow forecasting without or with drought forecast under the threshold probability of drought occurrence to 50 %	69
Fig. 8.7 The results of seasonal inflow forecasting without or with drought forecast under the threshold probability of drought occurrence range 50 % to 60 %	71

Seasonal inflow prediction for Andong Dam applying copula based Bayesian network and drought forecasting

Kwanghoon Kim

Department of Civil Engineering, The Graduate School,

Pukyong National University

Abstract

The higher the accuracy of reservoir inflow forecasting, the more reliable the water supply from a dam is especially for drought periods. The article aims at probabilistic forecasting of seasonal inflow to reservoir and deciding estimates from the probabilistic seasonal inflow according to the drought forecast results. The probabilistic seasonal inflow was forecasted by a copula based Bayesian network employing Gaussian copula function. Drought forecasting was performed by calculation of the standardized streamflow index value. The calendar year is divided by four seasons and the total inflow volume of water to a reservoir for a season means the seasonal inflow. Seasonal inflow forecasting curves conforming to drought stages produce estimates of probabilistic seasonal inflow according to the drought forecast results. The forecasted estimates of seasonal inflow were calculated by using the inflow records of Andong dam in the Republic of Korea. Under the threshold probability of drought occurrence ranging 50 to 60 %, the forecasted seasonal inflows reasonably matched critical drought records. Combining the drought forecasting with the seasonal inflow forecasting may produce reasonable estimates of drought inflow from the probabilistic forecasting of seasonal inflow to a reservoir.

제1장 서론

1.1 연구 배경 및 목적

우리나라의 기후는 계절풍의 영향을 받는 몬순기후로 규정할 수 있다. 겨울철에는 춥고 건조한 대륙성 기단인 시베리아 기단에 영향을 받아 비가 적고 매우 건조하다. 반면에 여름철에는 남쪽에서 발생한 덥고 습한 북태평양 기단과 북쪽에서 발생한 춥고 습한 오호츠크해 기단이 만나 장마전선이 형성되고, 이로 인해 6월 말부터 집중호우가 내린다. 이러한 기후 특성상 우리나라의 다목적 댐들은 비가 많이 내리는 홍수 시 댐 내에 물을 저류하여 그 이후에 사용하고 있다. 만약 홍수기에 강수량이 적어 댐 내에 용수를 충분히 저류하지 못하게 되면, 건조한 겨울 및 봄철에 충분한 용수 공급이 이뤄지지 않을 수도 있다. 또한 여러 해 가뭄이 연속되는 경우, 어느 해 홍수기에 용수를 충분히 저류하지 못하면, 이듬해 홍수기에 마른장마가 지속되거나 장마전선이 지나치게 빨리 북상하는 등의 이유로 강수량이 적어 용수를 충분히 확보하지 못할 경우 용수 공급에 큰 애로가 발생할 수도 있다(진영규 등, 2017).

댐의 용수 공급 실패는 그 댐의 용수를 사용하는 주민들과 기업들에게 큰 피해를 입힌다. 이에 K-water에서는 가뭄에 의한 피해를 저감시키기 위해 ‘댐 용수부족 대비 용수공급 조정기준’을 만들어 전국 15개의 다목적 댐을 운영하고 있다. 이 기준은 가뭄을 4 단계(관심, 주의, 경계, 심각)로 나누고 있으며, 각 단계별 필요 용수 감축량을 제시하고 있다(김광훈 등, 2017a).

더욱 안정적으로 물을 공급하기 위해서는 적정 수준의 정확성을 가진 유입량 예측이 필요하다. 하지만 ‘댐 용수부족 대비 용수공급 조정기준’은 물을 공급할 때 미래의 댐 유입량으로서 빈도분석에 의한 저수지 유입량을 사용하고 있다. 이러한 방법은 예보강수와 가뭄 전망을 제대로 고려하지 못한다는 한계점이 있다. 이를 보완하기 위해 앙상블 하천유량 예측 기법을 이용해 댐의 유입량을 예측하는 연구가 진행되어 왔다. 앙상블 하천유량 예측 기법은 강수의 기상 예보와 유역 토양의 수분상태를 감안해 유출량을 예측한다. 하지만 이 기법은 과거 기록만을 이용해 유출량을 예측하는 특성상 과거 기록에 존재하는 가뭄보다 더 큰 가뭄이 발생하면 이를 예측하는데 한계점이 있다. 그러므로 현재의 가뭄상태와 미래의 가뭄상태를 고려하여 댐 유입량을 예측할 필요가 있다.

본 연구의 목적은 직전 계절의 가뭄상태와 다음 계절에 가뭄이 발생할 확률을 고려하여 댐의 확률론적 계절 유입량을 예측하는 것이다. 확률론적 계절 유입량은 copula 함수 기반의 베이지안 네트워크를 이용해 예측되고, 미래에 발생 가능한 가뭄을 고려하기 위한 가뭄 예보는 가뭄지수의 계산을 통해 수행된다.

1.2 연구 동향

저수지 운영의 효율을 극대화하기 위해서는 첫째로 적정 수준의 정확성을 가진 예측 유입량을 활용하는 것이 반드시 필요하다 (Stedinger et al., 1984). 그리고 유입량의 월별 예측과 같은 중장기 예측은 강수, 온도 등 입력자료의 불확실성이 단기 예측에 비해 매우 크기 때문에 확정론적 예측보다는 확률론적 예측을 실시하는 것이 바람직하다(음형일, 2007). 가장 활발히 이용되고 있는 확률론적 예측 기법으로는 앙상블 하천유량 예측(ensemble streamflow prediction; ESP)이 있다. ESP의 첫 국내적용 사례로 김영오 등(2001)은 건설교통부 하천계획과와 한국건설기술연구원에서 발간한 ‘물공급전망’의 개선을 위해 금강유역에 ESP를 적용하였다. 강민석(2014)은 SWAT 모형과 ESP를 이용해 안동댐 유입량을 예측하였고, 진영규 등(2016)은 ESP와 용수 감량공급 기법을 이용해 합천댐의 가뭄 시 운영 방안을 제시하였다(김광훈 등, 2017b). 하지만 ESP는 유량 앙상블 시나리오를 생성하는 유역 유출모형의 선택과 생성된 각각의 시나리오에 대해서 가중값을 부여하는 방법에 따라 예측 결과가 지대한 영향을 받는다(진영규, 2016).

다른 확률론적 예측 기법으로는 베이지안 네트워크가 있다. 베이지안 네트워크는 수문학 분야뿐 아니라 금융, 보험 등 여러 분야에서 활발히 사용되고 있다. 김진영 등(2015)은 베이지안 네트워크를 이용해 수문학적 댐 위험도를 해석하는 기법을 개발하고 이를 이용해 댐 위험도 해석 시 수리·수문학적 변량의 관계를 효과적으로 평가하였다. 권현한 등(2012)은 미육군공병단의 HEC-1모형의 매개변수의 불확실성을 고려하기 위한 방안으로 베이지안 모형을 HEC-1모형에 연동시

켜 활용하였다. Madadgar and Moradkhani (2013)는 copula 함수 기반의 베이지안 네트워크를 이용해 Gunnison 강의 유량과 가뭄을 예측하고 그 결과를 ESP를 이용한 예측 결과와 비교하였다. Madadgar and Moradkhani (2014)는 베이지안 네트워크를 이용해 Gunnison 강 유역의 시간적, 공간적 가뭄을 예측하였다. Shin et al. (2016)은 우리나라 기상관측소를 대상으로 베이지안 네트워크 기반의 기상학적 가뭄 예보를 실시하였다. 김선호 등(2017)은 베이지안 이론과 ESP를 결합한 BAYES-ESP를 개발하여 충주댐의 유입량을 예측하였다.

1.3 연구 내용 및 범위

본 연구에서는 가뭄 예보 결과를 고려해 안동댐의 계절별 유입량을 예측하였다. 달력의 1년을 4계절로 나누었으며 각 계절은 봄(4월 ~ 6월), 여름(7월 ~ 9월), 가을(10월 ~ 12월), 겨울(1월 ~ 3월)이다. 계절별 유입량은 각 계절에 해당하는 3개월의 유입량을 합한 값이다. 댐 유입량 예측을 위해 기본적으로 copula 함수 기반의 베이지안 네트워크가 사용되었다. 이를 이용해 도출된 가뭄 단계별 계절 유입량 예측곡선은 가뭄 예보 결과를 고려하여 유입량을 예측하는데 사용되었다. 가뭄 예보는 가뭄지수를 이용해 실시되었으며 가뭄 단계는 가뭄지수를 기준으로 구분되었다. 가뭄지수를 이용해 다음 계절에 가뭄이 발생할 확률을 계산하고, 그 결과에 따라 다음 계절에 예측되는 가뭄 단계를 결정하였다. 다음 계절의 가뭄 단계가 결정되면 가뭄 단계별 계절 유입량 예측곡선 중 해당 가뭄 단계에 해당하는 곡선을 이용해 직전 계절 유입량에 대한 다음 계절 유입량을 예측하였다. 가뭄 예보를

결합한 안동댐의 계절별 유입량 예측을 위한 흐름도는 Fig. 1.1과 같다.

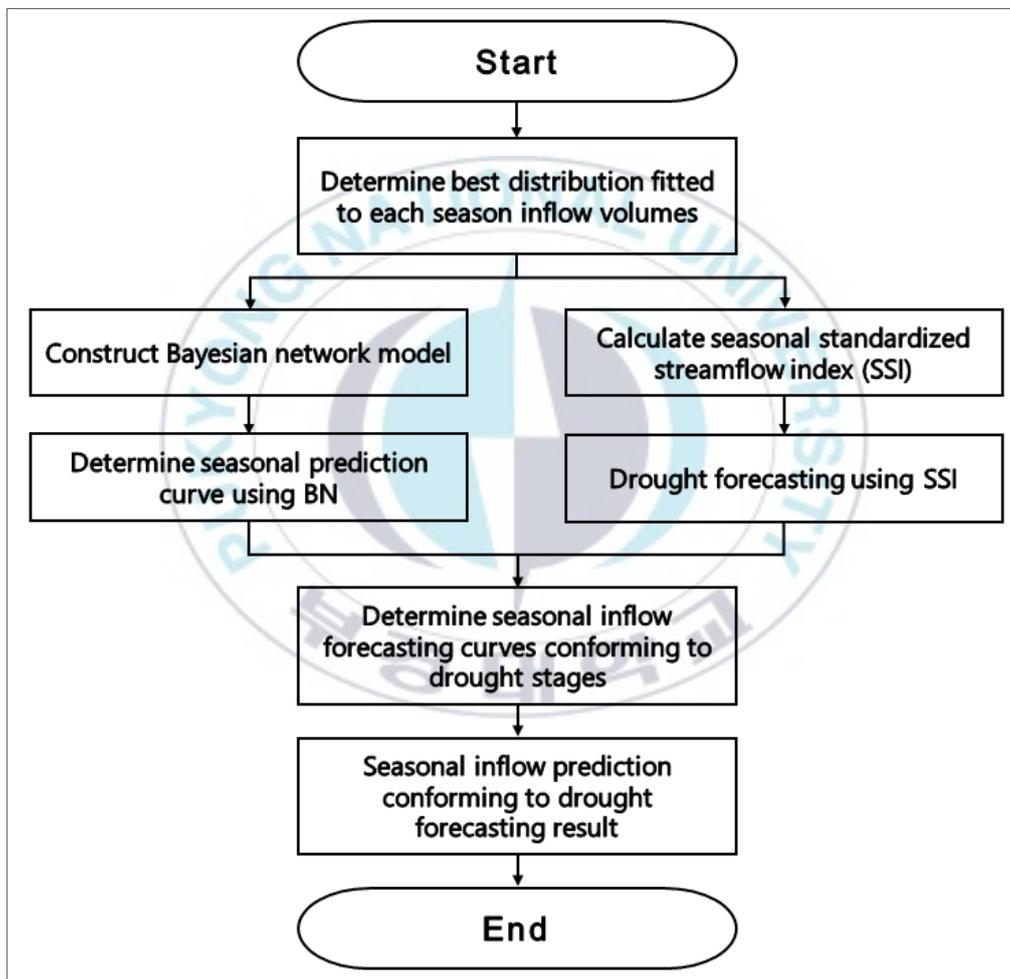


Fig. 1.1 Schematic diagram of seasonal inflow prediction applying a Bayesian network with drought forecasting

제2장 확률분포형 결정

2.1 확률분포형 이론

Copula 함수 기반의 베이지안 네트워크를 이용하기 위해서는 가장 먼저 각 변수들의 확률을 계산할 수 있어야 한다. 이를 위해서는 각 확률변수별 최적의 확률분포형을 결정해야 한다. 최적의 확률분포형을 결정하기 위해서는 후보가 되는 확률분포형들의 매개변수를 추정하고, 적합성 검정을 통해 최적의 하나를 선택하여야 한다. 수문학 분야에서 유량 및 강우량 등의 연속적인 수문자료 분석에 주로 사용되고 있는 확률분포형의 종류는 매우 다양하다. 그 중 본 연구에서 사용된 확률분포형은 빈도해석에 주로 사용되는 GEV 분포와 Log-Pearson type-III 분포를 제외한 lognormal 분포, gamma 분포, Gumbel 분포, Weibull 분포 및 Gaussian 분포이다.

2.1.1 Lognormal 분포

만일 무작위변수 x 가 몇 개의 다른 변수들의 곱으로 나타낼 수 있다면 x 에 대수(logarithm)를 취한 분포는 다른 인자들에 대수를 취한 것을 합산하게 되므로 정규분포에 접근하게 된다. 이렇듯이 변수에 대수를 취한 자료가 정규분포를 따르게 되면 원래의 변수는 lognormal 분포를 따른다고 할 수 있다. 수문자료의 대부분이 양의 값을 가지고 오른쪽으로 왜곡된 경향을 보이기 때문에 수문자료를 다루고자 할 때는 이 분포를 많이 사용한다. Lognormal 분포의 확률밀도함수는 Eq. (2.1)과 같다(이재수, 2006).

$$f(x) = \frac{1}{\sigma_y \sqrt{2\pi}} \frac{1}{x} \exp \left[-\frac{1}{2} \left(\frac{\ln x - \mu_y}{\sigma_y} \right)^2 \right] \quad (2.1)$$

여기서, $y = \ln x$, μ_y 는 y 의 평균, σ_y 는 y 의 표준편차이다.

2.1.2 Gamma 분포

Gamma 분포는 분포의 형태와 잘 알려진 수학적 특성 때문에 여러 수문자료 중 특히 홍수량의 분포를 결정하는데 많이 사용되며, 대수변환을 하지 않고도 왜곡도가 있는 수문변수를 설명하고자 할 때 이용된다. Gamma 분포의 확률밀도함수는 Eq. (2.2)와 같다.

$$f(x) = \frac{1}{\alpha \Gamma(\beta)} \left(\frac{x - x_0}{\alpha} \right)^{\beta-1} \exp \left(-\frac{x - x_0}{\alpha} \right) \quad (2.2)$$

여기서, α 와 β 는 각각 분포의 축척 및 모양을 결정하는 변수이며, $\Gamma(\beta)$ 는 gamma 함수이다. x_0 는 0이 아닌 값을 가지는 위치변수이다. 만약 축척변수 α 가 양수라면, $x_0 \leq x < \infty$ 이고, α 가 음수라면 $-\infty \leq x < x_0$ 가 된다(이정환, 2006).

2.1.3 Gumbel 분포

Gumbel 분포는 GEV(general extreme value)분포의 특별한 경우들 중 하나이다. GEV 분포는 주로 홍수나 가뭄 같은 사상의 빈도해석에 많이 사용되는 분포이다. 만약 GEV 분포에서 형상변수인 β 가 0이 된다면 그 분포는 Gumbel 분포가 된다. Gumbel 분포의 확률밀도함수는 Eq. (2.3)과 같다.

$$f(x) = \frac{1}{\alpha} \exp \left[-\frac{x-x_0}{\alpha} - \exp \left(-\frac{x-x_0}{\alpha} \right) \right] \quad (2.3)$$

여기서, α 는 축척변수이며 항상 양의 값을 가져야 한다. x_0 는 위치변수로서 최대 확률의 발생위치를 나타낸다.

2.1.4 Weibull 분포

Weibull 분포 또한 Gumbel 분포와 마찬가지로 GEV 분포의 특별한 경우 중 하나이다. Weibull 분포는 GEV 분포에서 형상변수인 β 가 양의 값을 가지는 분포이다. Weibull 분포의 확률밀도함수는 Eq. (2.4)와 같다(이정환, 2006).

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x-x_0}{\alpha} \right)^{\beta-1} \exp \left[-\left(\frac{x-x_0}{\alpha} \right)^\beta \right] \quad (2.4)$$

여기서, α 는 형상변수, β 는 축척변수이며 x_0 는 위치변수이다. 만약 x_0 가 0이되면 2변수 Weibull 분포가 된다.

2.1.5 Gaussian 분포

Gaussian 분포는 정규분포 또는 표준오차곡선(normal error curve)라고도 불리며 확률 및 통계분야에서 가장 중요한 분포이다. 그 이유는 일반적인 조건에서 독립확률변수들의 합이 커질수록 그 합인 분포는 정규분포에 가까워진다는 중심극한정리 때문이다. Gaussian 분포는 많은 물리 과정들을 개개 과정의 합으로 개념화 하므로 통계분야 뿐만 아니라 수문분야에서도 많이 적용하는 분포이다. Gaussain 분포는 일반적으로 평균치에 대하여 좌우가 대칭인 종 모양을 하고 있다. Gaussian 분포의 확률밀도함수는 Eq. (2.5)와 같다(이재수, 2006).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad (2.5)$$

여기서, μ 는 자료의 평균, σ 는 자료의 표준편차이다.

2.2 매개변수 추정

확률분포형의 매개변수를 추정하는 방법으로는 모멘트법, 최우도법, 확률가중모멘트법 및 L-모멘트 법이 있다. 이 중 본 연구에서는 가장 효율적인 모수 추정치를 산출시킨다는 점에서 이론적으로 가장 정확한 방법이라고 알려진 최우도법이 매개변수 추정기법으로 선정되었다(이정환, 2006).

최우도법의 기본 개념은 관측된 표본에 가장 적합한 모집단의 매개변수를 구하는 것이다. 즉, 다른 모집단보다 더 높은 가능성으로 관측된 표본을 산출할 수 있는 모집단의 매개변수 값을 구하는 것이다. 다시 말하면, 최우도법은 어느 한 모집단에서 한 매개변수의 여러 다른 값들은 서로 다른 표본들을 발생시키는 경향이 있다는 개념에 기초를 두고 있다. 따라서 어느 특정 표본은 어떤 특정 매개변수로부터 추출되어질 가능성이 다른 매개변수들로부터 추출될 가능성보다 크다. 그러므로 한 매개변수의 가능한 모든 값에서 관측된 어떤 특정한 값을 얻을 확률이 최대가 되는 매개변수를 추정할 필요가 있다(이재수, 2006).

주어진 표본자료에 대한 매개변수의 최우추정치(maximum likelihood estimator)는 표본의 확률 또는 확률밀도를 극대화하는 매개변수의 값이다. 만약에 표본자료가 모집단 $f(x; \theta)$ 에서 취한 확률표본이라면 표본의 확률은 Eq. (2.6)과 같은 우도함수(likelihood function)로 나타낼 수 있다(이재수, 2006).

$$L(\theta) = f(x_1; \theta)f(x_2; \theta) \cdots f(x_n; \theta) \quad (2.6)$$

따라서 최우추정치는 우도함수인 Eq. (2.6)을 극대화하는 매개변수 θ 값이다. 매개변수가 많아지면 매개변수를 여러 개로 확장하여 우도함수를 $L(\theta_i)$ 로 나타낼 수 있다. 계산의 편의를 위해 양 변에 대수를 취하여 대수 우도함수를 만들 수도 있다. 결론적으로 최우도법은 우도함수를 매개변수별로 미분한 뒤 0으로 놓고 추정하고자 하는 매개변수를 구하는 방법이다(이재수, 2006).

2.3 적합성 검정

본 연구에서는 계절별 유입량 자료를 대상으로 5가지 확률분포형의 적합성을 검정하고 가장 적합한 하나를 선택하기 위해 K-S 검정 (Kolmogorov-Smirnov test)이 적용되었다. K-S 검정은 관측된 자료의 누적분포함수 $F_n(x)$ 와 추정된 매개변수로 계산된 이론적인 누적분포함수 $F(x)$ 의 차이를 구하고 그 중 최대값을 한계치와 비교하여 해당 확률분포의 적합성을 판단하는 기법이다. K-S 검정의 첫 단계는 먼저 n 개의 관측 자료를 오름차순으로 정렬하고 Eq. (2.7)을 이용하여 누적확률을 계산하는 것이다(이재수, 2006).

$$F_n(x) = \frac{i}{n}, \quad x_i \leq x < x_{i+1} \quad (2.7)$$

여기서, n 은 자료의 개수이며 $F_n(x)$ 는 자료를 오름차순으로 정렬했을 때 i 번째 구간에 속하는 x 의 누적확률을 의미한다. 다음 단계는 관측 자료의 누적확률과 이론적인 누적확률을 계산해 각 구간에서 Eq. (2.8)을 이용하여 두 값의 최대 차 D_n 을 구하는 것이다.

$$D_n = \text{Max} |F_n(x) - F(x)| \quad (2.8)$$

D_n 은 자료의 총 수와 유의수준 α 에 따라 한계치를 가지며 그 값은 Table 2.1과 같다. Eq. (2.8)을 이용하여 계산된 D_n 이 한계값 D_n^α 보다 작으면 이 자료는 검정하는 확률분포형을 유의수준 α 를 가지고 따

른다고 할 수 있다. 작은 D_n 값은 관측된 자료의 누가주상도와 해당 확률분포형을 이용해 계산한 누가확률밀도함수의 차이가 작음을 의미한다. 즉, 후보가 되는 확률분포형 중 D_n 이 가장 작은 확률분포가 해당 자료를 가장 잘 표현한다고 할 수 있다.(이재수, 2006).



Table 2.1 The threshold values of the K-S test according to significance level and sample size

Sample size	Significance level α				
	0.2	0.15	0.1	0.05	0.01
1	0.900	0.925	0.950	0.975	0.995
2	0.684	0.726	0.776	0.842	0.929
3	0.565	0.597	0.642	0.708	0.829
4	0.494	0.525	0.564	0.624	0.734
5	0.446	0.474	0.510	0.563	0.669
6	0.410	0.436	0.470	0.521	0.618
7	0.381	0.405	0.438	0.486	0.577
8	0.358	0.381	0.411	0.457	0.543
9	0.339	0.360	0.388	0.432	0.514
10	0.322	0.342	0.368	0.409	0.486
11	0.307	0.326	0.352	0.391	0.468
12	0.295	0.311	0.338	0.375	0.450
13	0.284	0.302	0.325	0.361	0.433
14	0.274	0.292	0.314	0.349	0.418
15	0.266	0.283	0.304	0.338	0.404
16	0.258	0.274	0.295	0.328	0.391
17	0.250	0.266	0.286	0.318	0.380
18	0.244	0.259	0.278	0.309	0.370
19	0.237	0.252	0.272	0.301	0.361
20	0.231	0.246	0.264	0.294	0.352
25	0.210	0.220	0.240	0.264	0.320
30	0.190	0.200	0.220	0.242	0.290
35	0.180	0.190	0.210	0.230	0.270
40				0.210	0.250
50				0.190	0.230
60				0.170	0.210
70				0.160	0.190
80				0.150	0.180
90				0.140	
100				0.140	
$n \rightarrow \infty$	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

제3장 베이지안 네트워크

베이지안 네트워크는 그래픽 기반 모델로서 변수들의 관계를 나타내는 구조그래프와 각 노드(node)에 대한 조건부확률로 구성된 방향성 비순환 그래프(directed acyclic graph; DAG)로 정의할 수 있다(남기훈, 2014). 방향성 비순환 그래프란 방향성을 가지되, 한 노드에서 출발을 하면 다시 그 노드로 돌아가는 경로가 없는 그래프를 의미한다. 각 노드는 남과 여, 흡연 여부, 결혼의 유무 등으로 나타내지는 이산형 변수일 수도 있고, 유량, 기온 등으로 나타내지는 연속형 변수일 수도 있다. 만약 노드 A와 B가 있고, B가 A의 영향을 받는다면 A는 B의 부모노드가 되고, B는 A의 자식노드라고 부른다.

베이지안 네트워크의 유형으로는 가장 일반적이고 간단한 형태인 naive Bayesian network (NBN), NBN을 트리 형태로 확장한 tree augmented naive Bayesian network (TAN), NBN을 확장한 BN augmented naive Bayesian network (BAN), 그리고 일반 베이지안 네트워크인 general Bayesian network (GBN) 등이 있다(김덕정, 2015). NBN은 베이지안 네트워크 모형 중 가장 간단한 형태를 하고 있으며, 그 이유는 네트워크를 형성하는 모든 변수들이 조건부 독립이라는 가정을 하고 있기 때문이다. 이러한 특성상 여러 변수들의 결합 확률을 변수들 간 확률의 곱으로 나타내기 때문에 수식이 간단하다. 하지만 현실세계의 현상을 반영하는 데는 적합하지 않다는 문제점이 제기되어 왔으며, 이와 같은 약점을 보완하기 위해 개선된 모형들이 제시되어 왔다. TAN은 NBN이 자식노드들 사이에 지나치게 독립적이라는 가정에서 오는 문제점을 보완하기 위하여 자식노드들 사이에도

트리 형태의 관계가 있음을 가정한 베이지안 네트워크이다. BAN은 NBN과 달리 설명변수인 속성노드들 간에도 상호의존성이 존재한다고 가정한 모형이다. GBN은 베이지안 네트워크의 가장 일반화된 형태로 종속변수와 독립변수들의 차이를 두지 않고 모든 노드들 간의 상호의존성을 하나의 베이지안 네트워크로 표현하는 것이 특징이다. 따라서 GBN은 주어진 의사결정 문제에 속하는 여러 변수 간에 존재하는 확률적 인과관계 또는 상호의존성을 가장 자연스럽게 표시할 수 있다는 장점이 있다(선은정, 2010).

본 연구에서는 가장 일반적인 GBN이 댐의 계절별 유입량 예측에 사용되었으며, GBN의 간략한 모식도는 Fig. 3.1과 같다. Fig. 3.1에서 X_{21} 의 부모변수는 $X_{11}, X_{12}, \dots, X_{1n}$ 이며 X_{11} 의 자식노드들은 $X_{21}, X_{22}, \dots, X_{2n}$ 이 된다. 만약 Fig. 3.1이 댐의 봄철 유입량에 대한 여름철 유입량 예측을 위한 베이지안 네트워크라면, 왼쪽에 있는 봄철 유입량으로 이루어진 노드들은 오른쪽에 있는 여름철 유입량으로 이루어진 노드에 모두 영향을 끼친다고 할 수 있다.

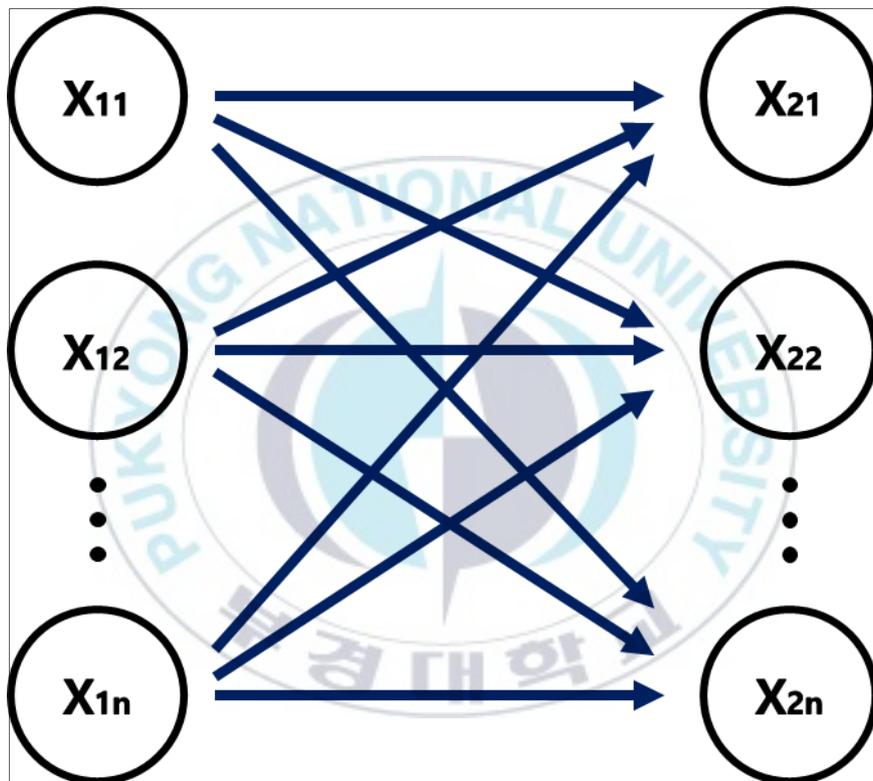


Fig. 3.1 A conceptual diagram of Bayesian network model

Madadgar and Moradkhani (2013)에 따르면 베이지안 네트워크에서 변수들의 집합인 X 의 결합확률분포는 Eq. (3.1)을 이용해 Eq. (3.2)와 같이 나타낼 수 있다.

$$f(A_1, \dots, A_n) = f(A_1)f(A_2|A_1) \dots f(A_n|A_1, A_2, \dots, A_{n-1}) \quad (3.1)$$

$$f(X) = f(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = \prod_{\forall i \in T} f(x_{t_i} | X_{pa(t_i)}) \quad (3.2)$$

Eq. (3.1)은 확률에서 사용하는 chain rule이다. Chain rule은 여러 확률변수들의 결합확률은 조건부 확률과 결합확률을 이용하여 연쇄적인 조건부 확률의 곱으로 표현가능하다는 정의이다. Eq. (3.2)에서 $X_{pa(t_i)}$ 는 시간 t_i 에서 x_{t_i} 의 부모변수를 의미하고 T 는 변수들이 포함된 전체 시간을 의미한다. 변수들이 시간의 흐름에 따라 연속적이고, x_{t_i} 의 부모변수가 모든 이전 변수들의 집합 $(x_{t_{i-1}}, x_{t_{i-2}}, \dots, x_{t_1})$ 이라면, Eq. (3.2)는 Eq. (3.3)과 같이 표현할 수 있다.

$$f(X) = f(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = \prod_{\forall i \in T} f(x_{t_i} | x_{t_{i-1}}, x_{t_{i-2}}, \dots, x_{t_1}) \quad (3.3)$$

Eq. (3.3)의 조건부 확률을 정리해서 나타내면 Eq. (3.4)와 같다.

$$f(x_{t_n} | x_{t_{n-1}}, \dots, x_{t_1}) = \frac{f(x_{t_n}, x_{t_{n-1}}, \dots, x_{t_1})}{f(x_{t_{n-1}}, \dots, x_{t_1})} \quad (3.4)$$

Eq. (3.4)의 우변에 있는 결합 확률밀도함수를 직접 계산하는 모형을 만드는 것은 상당히 어려우며 많은 분석적 노력이 필요로 한다. 하지만 Madadgar and Moradkhani (2013)는 copula 함수를 이용해 Eq. (3.4)의 우변을 조금 더 간단한 형태로 변환하여 쉽게 계산할 수 있도록 하였다. 다음 장에서는 여러 확률변수들의 결합확률을 쉽게 계산할 수 있도록 도움을 주는 copula 함수에 대해 설명하고자 한다.



제4장 Copula

4.1 Copula의 정의

Copula의 사전적 의미는 연결, 결합으로서 1959년 Sklar에 의해 처음으로 제시되었고, 그 이후로 Joe (1997)와 Nelsen (2006)에 의해 다양한 기본 이론 및 방법론이 정립되었다(김진영 등, 2016). Copula는 하나의 확률적 도구로서, 여러 개의 확률변수들이 서로 어떻게 연결되어 있는지를 나타내 준다. 확률변수 X_1, X_2, \dots, X_n 의 누적확률분포함수(cdf)를 각각 F_1, F_2, \dots, F_n 이라 하고 이들의 결합 확률분포함수를 F 라 하면, $F(x_1, x_2, \dots, x_n) = \text{prob}\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}$ 와 같이 표현할 수 있다. 이 확률변수들을 동시에 다루려면 n 차원 공간에 정의된 확률변수벡터 $X : (X_1, X_2, \dots, X_n)$ 가 고려되어야한다. 이러한 경우에 F_i 는 X 의 i 번째 주변(marginal)확률분포함수이다. 확률변수들간의 종속성을 알아보기 위하여 F 를 밀도함수부분과 종속성만을 나타내는 부분으로 분리하는 것이 copula를 정의하게 되는 동기이다. 다시 말해 유일한 결합 분포를 찾아낼 수 있으려면 밀도함수와 copula를 알아야 한다는 것이다(이재형, 2011).

Copula 함수는 복잡한 종속성 구조(dependence structure)를 고려하면서 다변량 누적분포함수를 추정하는데 유용한 도구를 제공한다. 즉, copula 함수는 다변량 분포함수와 주변확률 밀도함수를 연결시키는 함수를 의미한다. Copula 함수는 데이터 사이의 종속성 구조와 개별 자료의 분포를 분리하여 모형화 함으로서 추정과 시뮬레이션을 용이하게 하는 장점이 있으며, 다음과 같은 정의를 따른다(이재형,

2011).

n -차원의 copula C 는 다음과 같은 속성을 가진다.

1. $C: [0,1]^n \rightarrow [0,1]$
2. C 는 grounded 함수이며 n -증가함수이다.
3. C 는 모든 $u \in [0,1]$ 에 대하여 $C_n(u) = C(1, \dots, 1, u, 1, \dots, 1) = u$ 를 만족하는 주변확률 분포함수 C_n 을 갖는다

이러한 정의는 C 가 다변량 균일분포함수(multivariate uniform distribution)임을 보여준다. 즉, 만일 F_1, \dots, F_n 이 단변량 분포함수(univariate distribution function)라고 한다면, $C(F_1(x_1), \dots, F_n(x_n))$ 는 F_1, \dots, F_n 을 주변확률분포로 가진 다변량 누적분포함수이다. 왜냐하면, $u_n = F_n(x_n)$ 은 균일분포 확률변수(uniform random variable)이기 때문이다. 따라서, copula 함수는 다변량 분포함수의 추정이나 시뮬레이션을 쉽게 하는 유용한 도구라고 할 수 있다(이재형, 2001).

4.2 Sklar의 정리

Copula는 다변량 분포들 사이의 관계들을 이해하기 위한 방법으로 사용되며 그 형태는 1959년 Sklar에 의해 정립되었다. Sklar의 정리는 copula에서 가장 중요한 정리로서, 모든 copula 적용에 필수적으로 사용된다. Sklar 정리에 의하면 어떤 확률변수 X, Y 에 대해 각각 주변누가분포함수 $F_X(X), F_Y(Y)$ 가 존재하고 모두 연속인 경우 이를 결합할 수 있는 유일한 copula 함수 C 가 존재하게 된다(Schweizer and Sklar, 1983). 여기서, 주변누가분포함수 $F_X(X), F_Y(Y)$ 는 확률변수 X, Y 의 확률분포함수로부터 결정되며, 주변누가분포함수 $F_X(X), F_Y(Y)$ 의 결합 누가분포함수 $F_{X,Y}(X, Y)$ 는 copula 함수를 통해 Eq. (4.1)과 같이 정의된다(이재형, 2001).

$$F_{X,Y}(X, Y) = C(F_X(X), F_Y(Y)) \quad (4.1)$$

또한 두 변수의 주변 확률밀도함수 $f_x(x), f_y(y)$ 가 연속이라면, 결합 확률밀도함수 $f_{x,y}(x, y)$ 는 Eq. (4.2)와 같이 표현된다.

$$f_{x,y}(x, y) = c(F_X(x), F_Y(x)) f_x(x) f_y(y) \quad (4.2)$$

여기서, c 는 두 확률변수 X, Y 의 결합 확률밀도함수 $f_{x,y}(x, y)$ 와 주변 확률밀도함수 $f_x(x), f_y(y)$ 를 연결해주는 copula 함수이며, Eq. (4.3)과 같이 나타낼 수 있다(소재민, 2015).

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}, \quad u = F_x(x), \quad v = F_y(y) \quad (4.3)$$

4.3 Copula 함수의 종류

Copula는 그 특성에 따라 여러 가지의 종류로 나뉘게 된다. 수문학 분야 뿐만 아니라 경제, 통계학 분야에서도 가장 널리 적용되는 copula는 elliptical copula 와 Archimedean copula이다. Elliptical copula는 간단히 결합 확률분포로부터 도출할 수 있으며 함수의 종류로는 Gaussian copula 함수와 Student' s - t copula 함수가 있다. 이 두 copula 함수는 확률밀도함수의 형태가 타원형 분포이며 좌우 꼬리 부분이 대칭이라는 공통점이 있다(Fig 4.1). Archimedean copula는 elliptical copula와 다르게 매개변수를 추정할 필요가 있으며 함수의 종류로는 Clayton copula 함수, Frank copula 함수, Gumbel copula 함수가 있다. 이 세 종류의 copula 함수는 좌우 꼬리 부분이 대칭이 아니다(Fig. 4.2). Clayton copula 함수와 Gumbel copula 함수는 각각 좌측 꼬리 부분과 우측 꼬리 부분의 의존성이 더 높게 나타난다. 반면, Frank copula 는 대각선 축을 중심으로 좌우 대칭이며 꼬리 부분의 의존성이 없다(Bae and Choi, 2015). 위 5가지 copula 함수의 식은 Table 4.1에 나타나 있다.

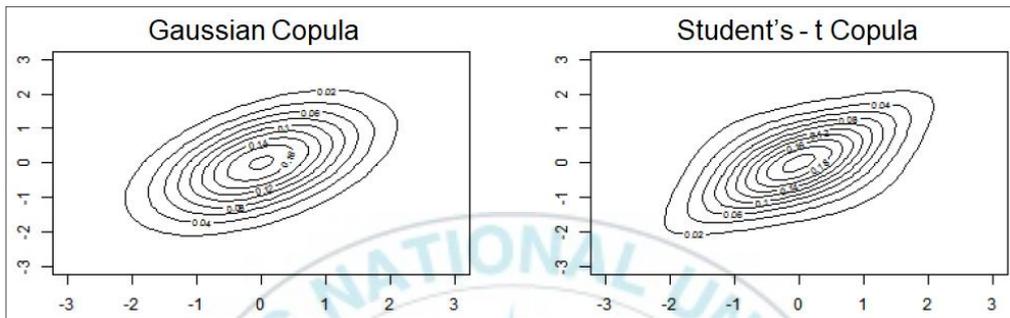


Fig. 4.1 Gaussian and Student's - t Copula
(Bae and Choi, 2015)

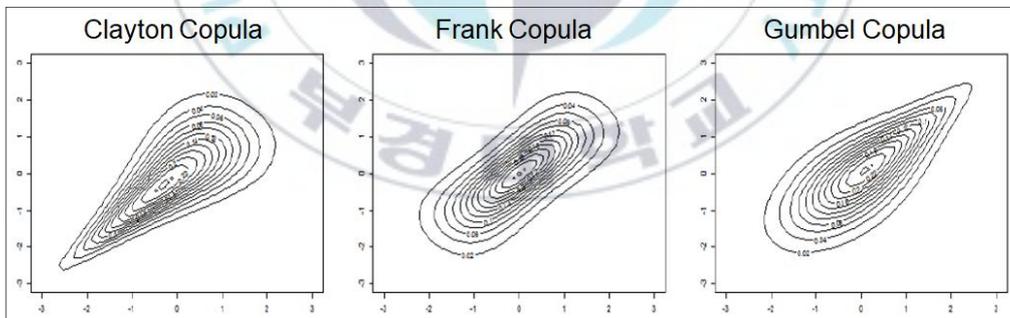


Fig. 4.2 Clayton, Frank, and Gumbel Copula
(Bae and Choi, 2015)

Table 4.1 The copula functions

Name of copula	Joint distribution function
Gaussian	$c(u_1, u_2, \dots, u_n; R) = \frac{1}{ R ^{1/2}} \exp\left[-\frac{1}{2} u' (R^{-1} - I) u\right]$
Student's t	$c(u_1, u_2, \dots, u_n; R, \nu) = R ^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left[\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \right]^n \frac{(1 + \zeta' R^{-1} \zeta)^{-\frac{\nu+n}{2}}}{\prod_{i=1}^n \left(1 + \frac{\zeta_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}}$
Frank	$-\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right]$
Gumbel	$\exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right)$
Clayton	$\max\left[(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, 0\right]$

Nelsen (2006)

수문학 분야에서 copula 함수의 사용에 대해 살펴보면, 유지영 등 (2013)은 전국 57개 기상관측소를 대상으로 Table 4.1의 copula 함수를 이용해 가뭄의 이변량 빈도해석을 실시하였다. 이 때 대상으로 했던 57개 기상관측소 중 가장 많은 32개 관측소의 자료에 Gaussian copula 함수가 가장 적합하다는 결과를 제시하였다. 또한, Madadgar and Moradkhani (2013)는 미국의 Gunnison 강을 대상으로 계절 간 유량의 결합밀도함수를 계산하는데 Gaussian copula 함수가 가장 적절하다는 결과를 도출하였다. Gaussian copula 함수는 매개변수를 추정할 필요가 없어 Archimedean copula 에 비해 계산이 쉽다. 또한 같은 elliptical copula 인 Student' s - t copula 함수에 비해 식이 비교적 간단하고, 상관계수 행렬만 구할 수 있다면 결합확률의 계산이 가능하다. 이러한 이유로 본 연구에서는 결합 확률밀도함수를 계산하기 위해 5가지의 copula 함수 중 Gaussian copula 함수가 사용되었다. Gaussian copula 함수에 사용되는 변수들 중 u 와 R 은 Eqs. (4.4) ~ (4.6)과 같다.

$$u = (u_1, u_2, \dots, u_n)' \quad (4.4)$$

$$u_i = \Phi^{-1}(F_i(X_i)) \quad (4.5)$$

$$R = \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{pmatrix} \quad (4.6)$$

여기서, $\Phi^{-1}(F_i(X_i))$ 는 정규누가분포함수의 역함수를 의미하고, R 은 변량 간 상관계수인 ρ 로 이루어진 상관계수 행렬을 의미한다(Zezula,

2009). 만약 두 변수 (X, Y) 를 0이 아닌 유한한 확률변수 벡터라고 하면 (X, Y) 의 상관계수 $\rho(X, Y)$ 는 Eq. (4.7)을 이용해 계산할 수 있다.

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)} \sqrt{Var(Y)}} \quad (4.7)$$

여기서, $Cov(X, Y)$ 는 (X, Y) 의 공분산, $Var(X)$, $Var(Y)$ 는 각각 확률변수 X , Y 의 분산이다(이재형, 2001).



4.4 Copula 함수 기반의 베이지안 네트워크

Copula 함수 기반의 베이지안 네트워크는 Madadgar and Moradkhani (2013)에서 소개되었다. Madadgar and Moradkhani (2013)는 앞선 절에서 소개한 copula 함수를 이용해 Eq. (3.4)를 Eq. (4.8)과 같이 변형하였다.

$$f(x_{t_n} | x_{t_{n-1}}, \dots, x_{t_1}) = \frac{c(u_{t_n}, \dots, u_{t_1}) \prod_{i=t_1}^{t_n} f_{x_i}(x_i)}{c(u_{t_{n-1}}, \dots, u_{t_1}) \prod_{i=t_1}^{t_{n-1}} f_{x_i}(x_i)} \quad (4.8)$$

여기서, \prod 는 곱셈 연산자이며 $f(\cdot)$ 는 주변 확률분포함수이다. Eq. (4.8)을 사용하면 여러 변수들 간 조건부확률이 간단히 계산된다. 본 연구에서는 연속된 두 계절 유입량의 조건부확률을 계산해 유입량을 예측하기 때문에, Eq. (4.8)은 두 변수를 가진 Eq. (4.9)와 같이 나타낼 수 있다.

$$f(x_{t_2} | x_{t_1}) = \frac{c(u_{t_2}, u_{t_1}) f_{x_{t_2}}(x_{t_2}) f_{x_{t_1}}(x_{t_1})}{f_{x_{t_1}}(x_{t_1})} = c(u_{t_2}, u_{t_1}) f_{x_{t_2}}(x_{t_2}) \quad (4.9)$$

여기서, $c(\cdot)$ 은 copula 함수, x_{t_2} 는 예측하고자 하는 계절의 유입량, x_{t_1} 은 직전계절의 유입량이다.

제5장 가뭄지수를 이용한 가뭄예보

5.1 가뭄의 정의

가뭄은 강수 부족에 따른 물수지 불균형 현상으로 정상적이면서 주기적으로 재발하는 기후의 한 특성으로 수개월에서 수년에 걸쳐 물 공급이 부족한 시기를 의미하며, 평균이하의 강수량이 지속적으로 이어지는 지역에서 나타난다. 가뭄은 일시적으로 발생하는 홍수와 달리 시작을 정확히 파악하기가 어렵다. 또한 가뭄은 상당기간 동안 서서히 누적되어 나타나고 가뭄이 해결된 후에도 경제적, 사회적으로 오랫동안 파급효과가 나타나기도 한다. 이 때문에 가뭄을 정확하고 정량적으로 정의하는 것이 어렵다(박명우, 2016).

일반적으로 가뭄은 기상가뭄, 농업가뭄, 수문가뭄 및 사회·경제적 가뭄으로 구분된다. 기상학적 가뭄은 강수량의 부족으로 인한 가뭄에 기반을 두고 있다. 농업가뭄은 농작물, 산림, 초지 등에 필요한 토양수분량의 부족을 말한다. 수문가뭄은 하천유량, 저수지와 호수의 수위 및 지하수에 의한 수자원 공급원에 연관된다. 또한, 사회·경제적 가뭄은 자연현상, 환경악화 및 사람들에 대한 영향의 복잡한 상호작용으로 정의된다(김한수 등, 2016).

5.2 가뭄지수

5.2.1 가뭄지수의 선정과 정의

가뭄지수는 가뭄을 정량적으로 평가하기 위한 지표로서, 강수량, 기온, 증발산량 등 여러 수문 및 기상인자를 이용해 하나의 수치로써 나타낸 것이다. 가뭄지수를 이용하면 가뭄의 상대적 비교가 가능하여 가뭄에 대한 정량적 평가가 가능하다. 가뭄지수는 관심 영역에 따라 크게 기상학적, 농업적 및 수문학적 가뭄지수로 분류할 수 있다.

먼저 기상학적 가뭄지수로는 강수 부족량을 산정해 가뭄을 구분하는 표준강수지수(standardized precipitation index; SPI), 수분의 부족량과 부족기간의 함수를 이용해 가뭄의 심도를 나타내는 팔머가뭄지수(Palmer drought severity index; PDSI)가 있다. 댐의 유입량, 즉 어느 유역의 유출량은 강수량뿐 아니라 토양함수비, 불투수면적비 등 여러 가지 요인들이 종합적으로 적용해 발생한다. 그렇기 때문에 여러 수문 요인 중 하나인 강수량의 부족량을 산정해 가뭄을 구분하는 SPI는 본 연구를 위한 가뭄지수로 적절하지 않다고 판단된다. PDSI는 강수량 자료, 기온, 일조시간 등의 자료를 이용하여 기후적으로 필요한 강수량과 실제 강수량을 비교하여 가뭄을 구분하는 지수이다. PDSI도 SPI와 유사한 이유로 본 연구를 위한 가뭄지수로 적절하지 않다고 판단된다.

농업적 가뭄지수로는 토양수분지수(soil moisture index; SMI), 작물수분지수(crop moisture index; CMI) 등이 있다(박명우, 2016). 두 가뭄지수는 각각 토양의 수분, 작물의 수분량을 토대로 가뭄지수를 계산한다. 이 가뭄지수들 또한 유입량을 이용하는 본 연구의 특성상

적절하지 않은 가뭄지수로 판단된다.

마지막으로 수문학적 가뭄지수로는 SPI와 개념은 같으나 강수 부족량이 아닌 유출량 부족량을 산정해 가뭄을 구분하는 표준유출량지수 (standardized streamflow index; SSI)가 있다. SSI는 여러 가지 요인들이 적용된 관측 유출량 자료를 직접 이용해 가뭄을 구분한다는 장점이 있다. 댐의 유입량은 그 댐 구역의 유출량과 같은 의미를 지니고 있으므로 본 연구에서는 가뭄을 구분하고 가뭄예보를 실시하기 위한 가뭄지수로 SSI를 선정하였다. SSI는 관측된 유출량을 이용해 Mckee et al. (1993)이 개발한 SPI의 산정 방법을 기반으로 계산된다. SSI의 범위에 따라 가뭄 단계는 4가지로 분류되며, SSI의 범위에 따른 가뭄 단계의 구분은 Table 5.1과 같다.

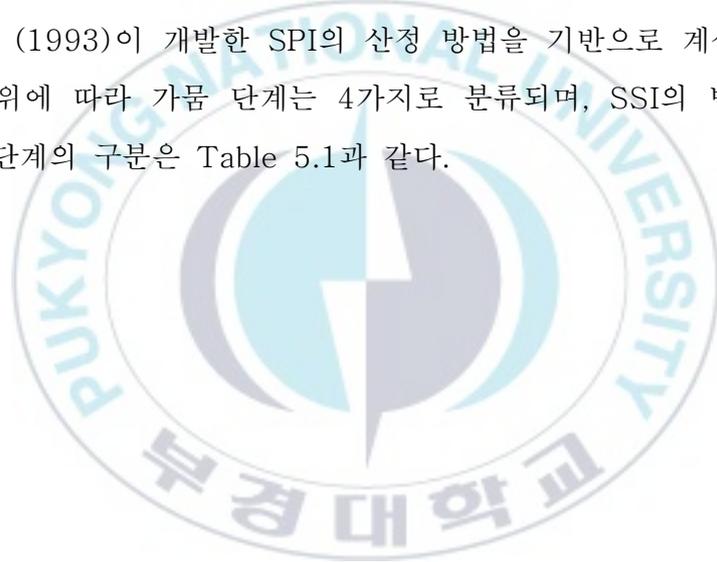


Table 5.1 Classification of drought stages for the range of SSI

SSI range	Drought category
$2.00 \leq Z$	Extreme wet
$1.50 \leq Z < 2.00$	Very wet
$1.00 \leq Z < 1.50$	Moderately wet
$0.00 \leq Z < 1.00$	Near normal
$-1.00 \leq Z < 0.00$	Mild drought (D1)
$-1.50 \leq Z < -1.00$	Moderate drought (D2)
$-2.00 \leq Z < -1.50$	Severe drought (D3)
$Z < -2.00$	Extreme drought (D4)

Son et al. (2011)

5.2.2 가뭄지수 산정

Mckee et al. (1993)은 강수량이 감소하게 되면 물 수요에 비해 물이 상대적으로 부족해져 가뭄이 시작하게 된다는 점에 착안하여 SPI를 개발하였다. 이는 강수 부족에 대한 확률적 특성을 적용한 것이라 할 수 있다. 과거기간의 강수량자료에 대한 해당 확률밀도함수의 누가확률분포함수를 통해서 실제 강수량의 누가확률을 계산한 후 실 강수량의 누가확률을 정규분포함수를 이용하여 지수로 환산해 가뭄을 구분하는 방식이다(류재희 등, 2002). SSI는 SPI의 산정 방법을 기반으로 강수량 대신 유량 자료를 이용해 계산이 가능하다.

SSI의 산정절차는 크게 댐의 월별 유입량 자료의 수집 및 계절별 유입량 자료 구축, 계절별 유입량 자료의 최적 확률분포형 결정 및 표준화를 통한 가뭄지수 산정으로 나눌 수 있다. 계절별 유입량 자료는 각 계절에 해당하는 3개월의 유입량을 합한 값이다. 계절별 유입량 자료가 구축되면 2장에서 설명한 방법을 이용해 계절별 유입량 자료의 최적 확률분포형이 결정되어야 한다. 각 계절별 유입량 자료의 확률분포형이 결정되면, 특정 유입량에 대한 누가확률이 계산된다. 마지막으로, 특정 유입량에 해당하는 누가확률과 표준정규분포의 누가확률이 같은 값을 찾아 그 Z값을 읽으면 그 값이 SSI가 된다. 본 연구에서는 유입량의 누가확률을 표준정규분포의 확률변수값으로 변환하기 위하여 Abromowitz and Stegun (1964)이 제시한 방법을 활용하였으며, 그 식들은 Eqs. (5.1) ~ (5.4)와 같다.

$$t = \sqrt{\ln\left(\frac{1}{(H(x))^2}\right)}, \quad 0.0 < H(x) \leq 0.5 \quad (5.1)$$

$$t = \sqrt{\ln\left(\frac{1}{1-(H(x))^2}\right)}, \quad 0.5 < H(x) \leq 1.0 \quad (5.2)$$

$$Z = -\left(t - \frac{c_0 + c_1t + c_2t^2}{1 + d_1t + d_2t^2 + d_3t^3}\right), \quad 0 < H(x) \leq 0.5 \quad (5.3)$$

$$Z = +\left(t - \frac{c_0 + c_1t + c_2t^2}{1 + d_1t + d_2t^2 + d_3t^3}\right), \quad 0.5 < H(x) \leq 1.0 \quad (5.4)$$

여기서, $H(x)$ 는 확률분포함수에서 계산된 누가확률, c_0 는 2.515517, c_1 은 0.802853, c_2 는 0.010328, d_1 은 1.432788, d_2 는 0.189267, d_3 는 0.001308이다(Son et al., 2011).

5.3 가뭄 예보

가뭄 예보는 Eq. (4.9)에 사용되는 두 변수를 각각 직전 계절의 SSI, 다음 계절의 SSI으로 바꾸어 다음 계절에 가뭄이 발생할 확률을 계산해 실시한다. 다음 계절에 가뭄이 발생할 확률은 다음 계절의 SSI가 0을 초과하지 않을 비 초과 확률이다. 예를 들면, Fig. 5.1은 봄의 SSI가 -1일 때 여름 SSI의 확률분포이다. Fig. 5.1에서 빗금으로 표시된 부분의 면적은 SSI가 0을 초과하지 않을 비 초과 확률이다. 만약 이 확률이 가뭄 예보 기준 이상이라면, 다음 계절에 가뭄이 발생할 것이라 예보한다. 가뭄 예보 기준이란 다음 계절에 가뭄이 발생할 확률이 얼마나 되어야 다음 계절에 가뭄이 발생할 것이라 판단하는지를 결정하는 기준이다. 만약 가뭄 예보 기준이 50 %라면, SSI가 0을 초과하지 않을 비 초과 확률이 51 %만 되어도 가뭄을 예보하는 것이다. 반대로, SSI가 0을 초과하지 않을 비 초과 확률이 가뭄 예보 기준 이하라면, 다음 계절엔 가뭄이 예보되지 않는다.

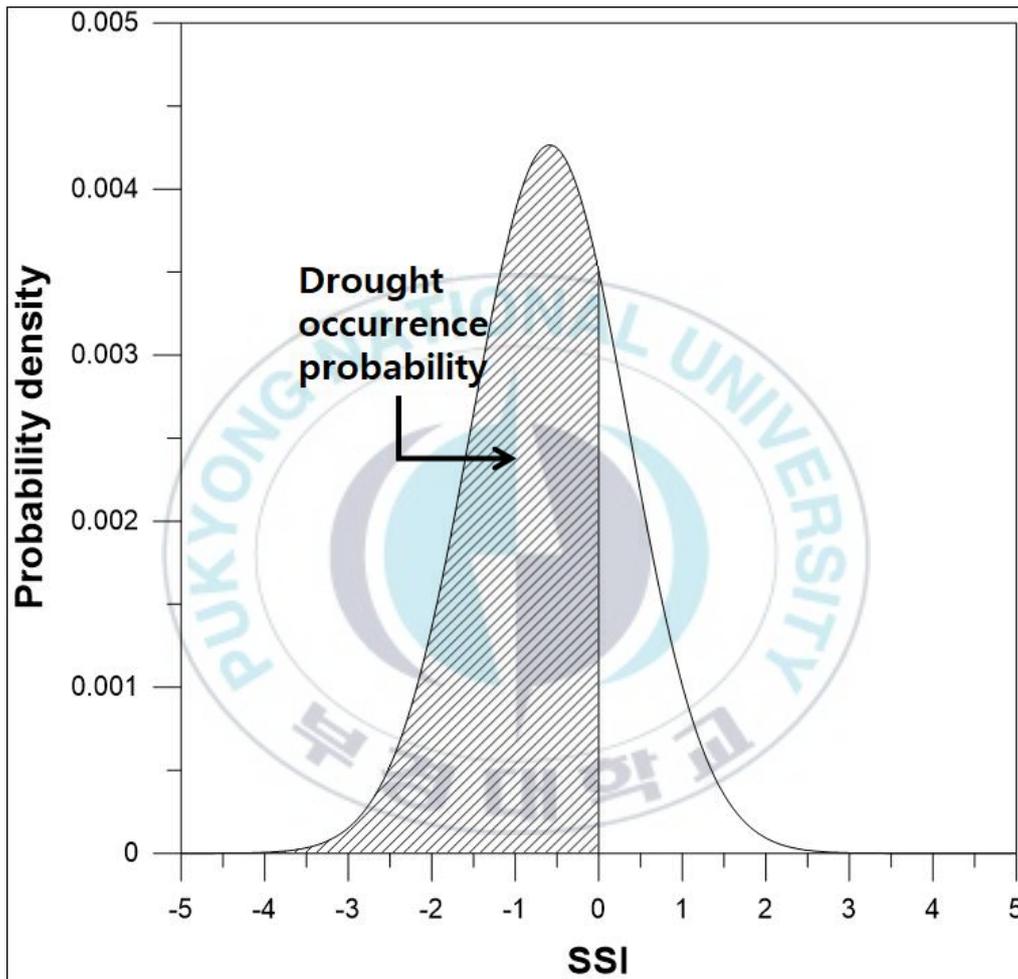


Fig. 5.1 Probability density function of summer SSI conditioned on a spring SSI of - 1

제6장 Copula 함수 기반의 베이지안 네트워크를 이용한 유입량 예측

6.1 연구 대상 댐

6.1.1 연구 대상 댐 선정

계절별 유입량을 확률론적으로 예측하고, 그 결과를 평가하기 위해서 기록 기간이 긴 자료가 있어야 한다. 또한, 본 연구의 목적은 가뭄 예보 결과를 고려한 계절별 유입량 예측을 하는 것이다. 그러므로 연구 대상 댐은 가뭄이 발생하면 큰 피해가 발생할 것으로 예상되어야 한다. 또한, 예측 결과의 적절성을 평가하기 위해 검증 기간에 최소 2회의 가뭄이 포함되어야 한다.

이에 본 연구에서는 낙동강권역 내 다목적댐들 중 댐 용량이 가장 크고 1977년 1월부터 현재까지 월별 유입량 자료를 제공하는 안동댐이 연구 대상 댐으로 선정되었다. 안동댐은 낙동강권역 내 다목적 댐들 중 용수 공급량이 가장 많으며, 2013년과 2015년에 가뭄이 발생해 피해를 입은 적이 있다.

6.1.2 안동댐의 제원 및 현황

안동댐은 낙동강권역 내에 있는 다목적 댐 중 하나이다. 낙동강 대권역의 유역면적은 23,384 km²이며, 그 중 안동댐의 유역면적은 1,584 km²로서 낙동강 전체 유역면적의 약 6.6 %에 해당된다(강신욱, 2011). 1976년 10월에 준공된 안동댐은 1977년 1월부터 댐 수문자료를 제공하고 있다. 안동댐의 연평균 유입량은 940 백만 m³, 저수면적은 51.5 km² 이다. 안동댐의 총 저수용량은 상시만수위를 기준으로 1,248 백만 m³으로 낙동강 내 다목적 댐 중 가장 크며, 유효저수용량은 1,000 백만 m³이다. 안동댐의 자세한 저수지 특성자료는 Table 6.1에 나타나 있다. 안동댐은 연간 약 930 백만 m³의 용수를 공급하도록 계획되어 있는데, 이 중 생·공용수는 450 백만 m³, 농업용수와 하천유지용수는 각각 300 백만 m³, 176 백만 m³를 공급하도록 계획되어 있다. 각 월별로 보면 안동댐은 다른 다목적댐들과 마찬가지로 농사철인 4월 ~ 10월엔 농업용수를 공급하고 있으며, 나머지 기간은 대부분 생·공용수를 공급하고 있다. 각 월별 안동댐의 용수공급계획량은 Fig. 6.1에 나타나 있다(수자원공사, 2014).

Table 6.1 Project data of Andong Dam

하천	낙동강	유역연평균유입량 (m ³ /s)	27.00
저수면적 (km ²)	51.50	계획홍수위 (El. m)	161.70
유역면적 (km ²)	1,584	상시만수위용량 (백만 m ³)	1,248
유역연평균강우량 (mm)	950.00	홍수조절용량 (백만 m ³)	110.00
상시만수위 (El. m)	160.00	저수위용량 (백만 m ³)	237.40
저수위 (El. m)	130.00	총저수용량 (백만 m ³)	1,248
용수공급가능수위 (El. m)	121.00	사수용량 (백만 m ³)	118.00
유효저수용량 (백만 m ³)	1,000.00	공급불가능용량 (백만 m ³)	118.00
비상용수공급량 (백만 m ³)	130.00	실제홍수량 (m ³ /s)	8,350
저수지길이 (km)	43.50		

한국수자원공사 (2014)

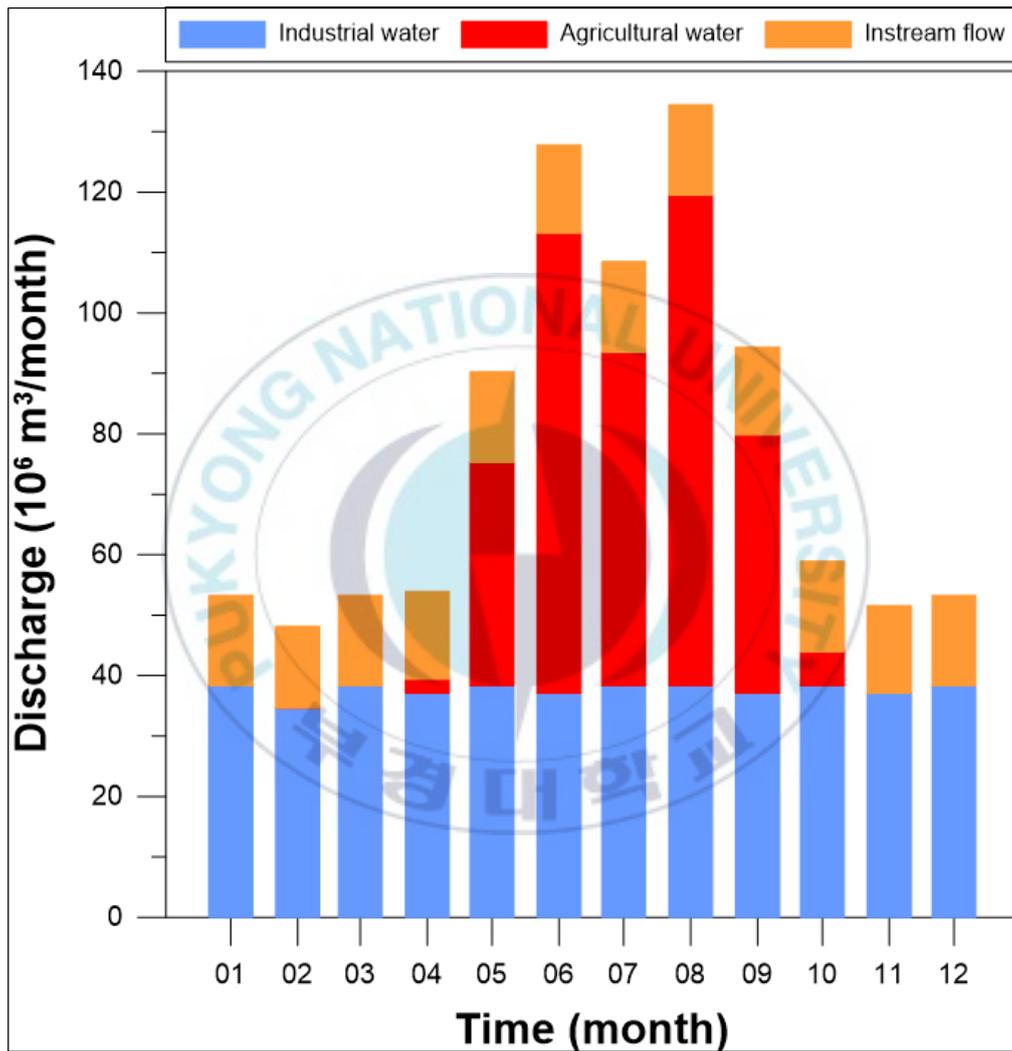


Fig. 6.1 Monthly water supply planning of Andong Dam

6.2 Copula 함수 기반의 베이지안 네트워크 구축 및 적용

6.2.1 Copula 함수 기반의 베이지안 네트워크 구축

본 절에서는 안동댐의 1977년부터 2016년까지 계절별 유입량 자료를 수집하고 계절별 유입량을 예측하기 위해 베이지안 네트워크를 구축하였다. 1977년부터 2016년까지 계절별 유입량은 Fig. 6.2와 같다. 1977년부터 2010년까지 계절별 유입량 자료는 베이지안 네트워크를 구축하는데 사용되었다. 2011년부터 2016년까지의 계절별 유입량 자료는 계절 유입량 예측 결과의 적절성을 평가하기 위해 사용되었다. 가뭄 예보 결과를 고려한 계절별 유입량 예측을 위해 가장 먼저 최우도법을 이용해 각 계절별 5가지 확률분포의 매개변수를 추정하였다. Fig. 6.3 ~ 6.6은 각 계절별 유입량 자료의 히스토그램과 5개의 확률분포형을 함께 나타낸 것이다.

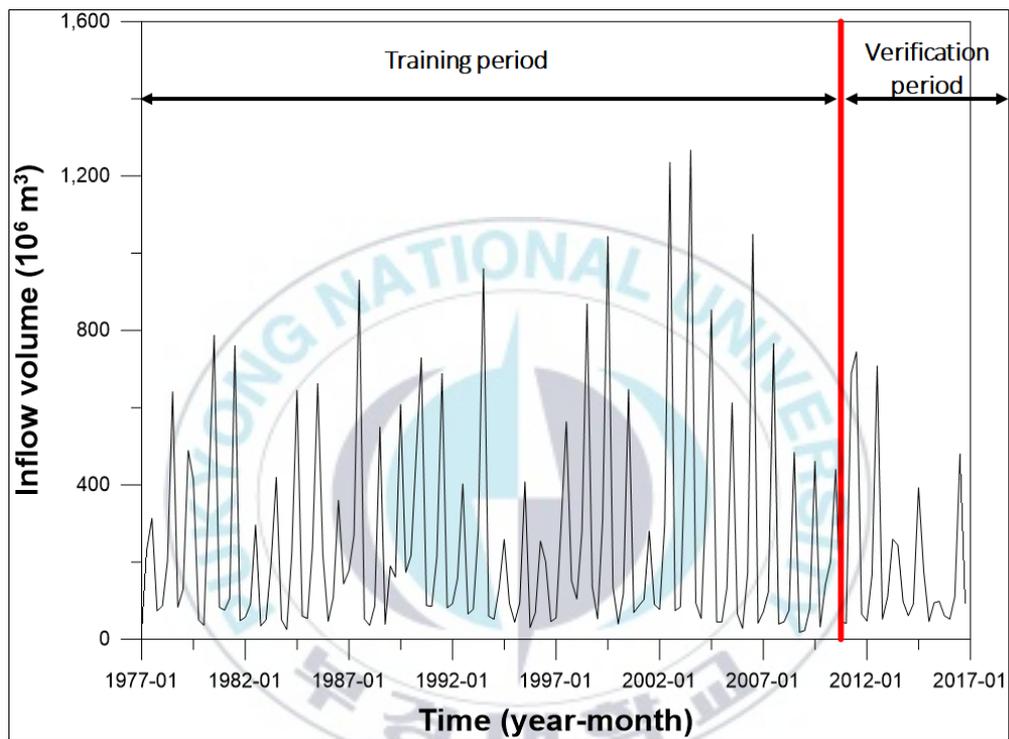


Fig. 6.2 Seasonal inflow of Andong Dam

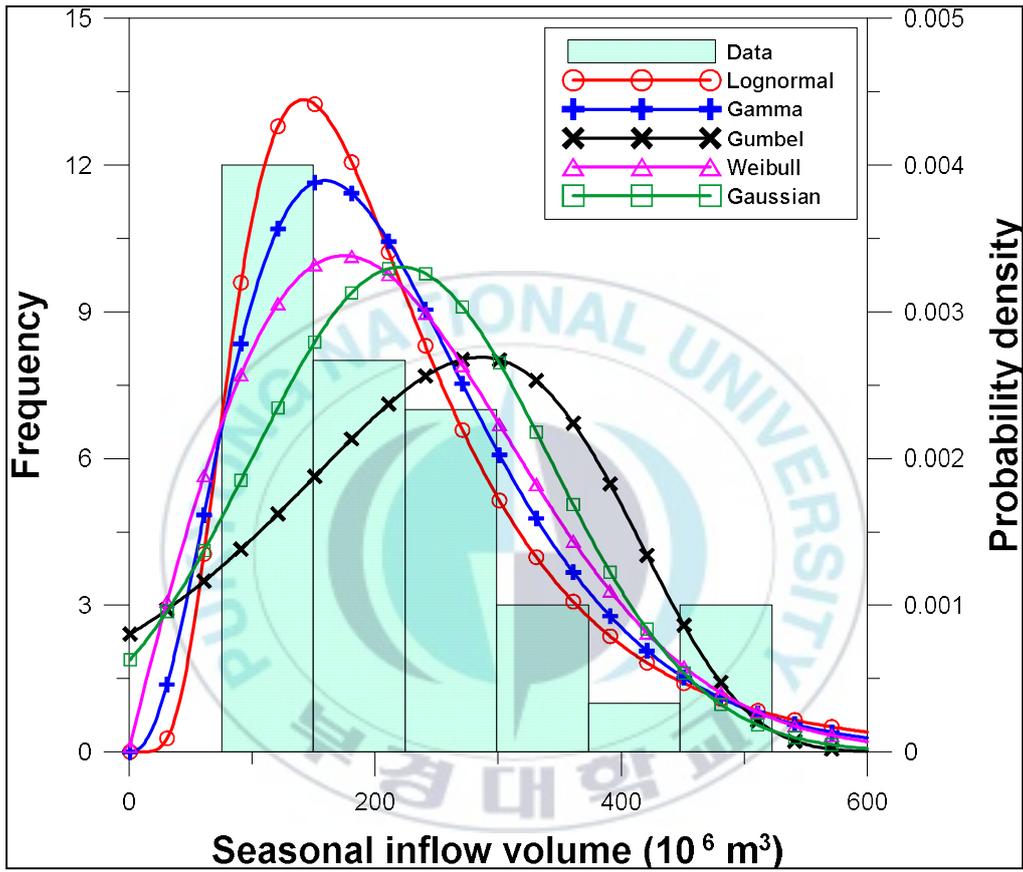


Fig. 6.3 Histogram against the five distributions of the spring inflow volumes during the training period of 1977–2010

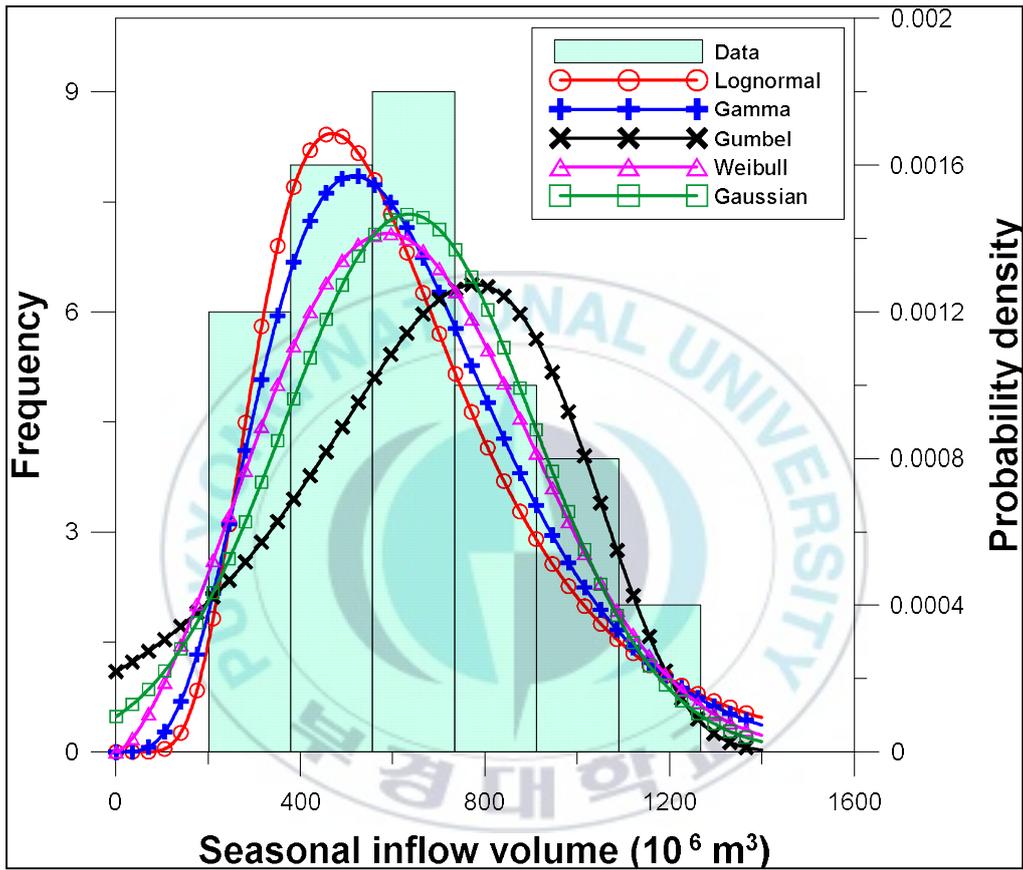


Fig. 6.4 Histogram against the five distributions of the summer inflow volumes during the training period of 1977–2010

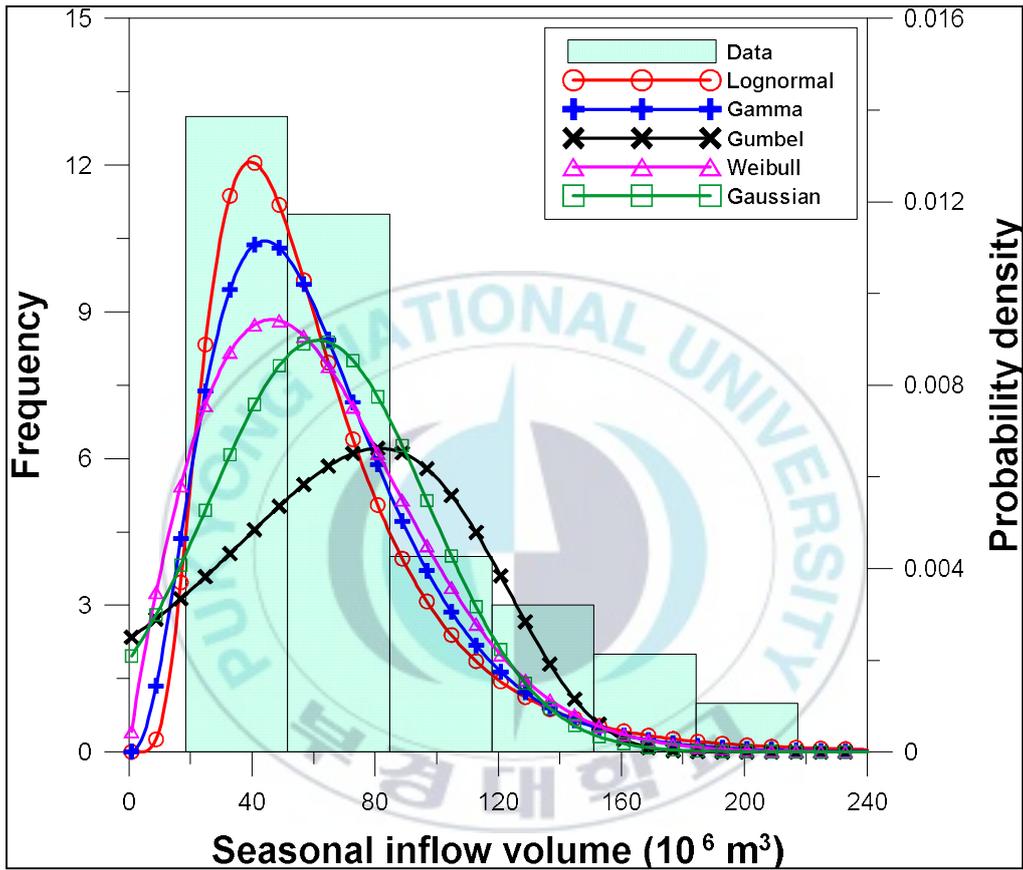


Fig. 6.5 Histogram against the five distributions of the fall inflow volumes during the training period of 1977–2010

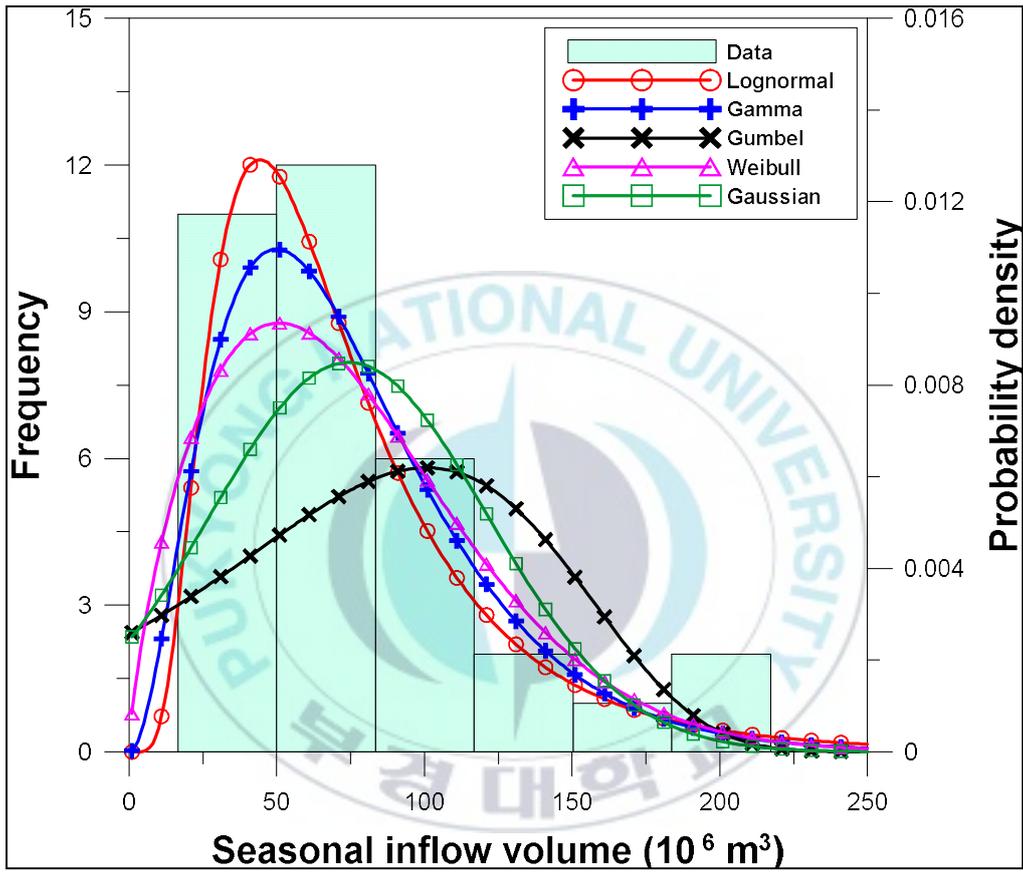


Fig. 6.6 Histogram against the five distributions of the winter inflow volumes during the training period of 1977–2010

계절별 유입량 자료에 가장 적합한 확률분포형은 유의수준 0.05의 K-S 검정을 통해 결정되었다. 유의수준이 0.05이고, 자료의 개수가 34개이므로 Table 3.1을 참고하면 K-S 검정의 한계치는 0.232이다. 각 계절별 K-S 검정의 결과는 Table 6.2와 같으며, K-S 검정 결과 D_n 이 가장 작은 수치는 굵게 표시되어 있다. Table 6.2를 보면, 굵게 표시된 수치는 모두 한계치인 0.232를 넘지 않는다. K-S 검정 결과, 여름 유입량 자료는 Weibull 분포, 여름을 제외한 봄, 가을, 겨울의 유입량 자료는 lognormal 분포형이 가장 적합한 것으로 나타났다.

다음으로, Gaussian copula 함수를 계산하기 위해 연속된 계절 간 유입량의 상관계수를 계산한 결과는 Table 6.3과 같다. 연속된 두 계절 간 유입량의 상관계수 계산 결과, 봄과 여름, 겨울과 봄의 유입량 간 상관계수가 다른 계절 유입량 간 상관계수에 비해 크게 계산되었다. 이는 봄과 여름, 겨울과 봄의 유입량 간 종속성이 크다는 것을 의미한다. 반면에 여름과 가을, 가을과 겨울의 유입량 간 상관계수는 비교적 작게 계산되었다. 그 이유는 가을과 겨울이 건조한 우리나라의 기후 특성상 여름철 홍수기 이후엔 큰 변동 없이 적은 유입량이 지속되기 때문으로 판단된다.

Table 6.2 The D_n values of the K-S test

Distribution	Season				Threshold value
	Spring	Summer	Fall	Winter	
Lognormal	0.1001	0.1057	0.0813	0.0915	0.232
Gamma	0.1123	0.0771	0.1072	0.1267	
Gumbel	0.1930	0.1288	0.2396	0.2457	
Weibull	0.1038	0.0736	0.1380	0.1394	
Gaussian	0.1205	0.0932	0.1713	0.1871	

Table 6.3 The correlation coefficient of each season's inflow to that of prior season

	Spring – Summer	Summer – Fall	Fall – Winter	Winter – Spring
Correlation coefficient	0.4828	0.1281	0.1831	0.3767

6.2.2 Copula 함수 기반의 베이지안 네트워크를 이용한 유입량 예측 결과 및 평가

본 절에서는 2011년 ~ 2016년을 대상으로 가뭄 예보 결과를 고려하지 않고 계절별 유입량을 예측하였다. 가뭄 예보 결과를 고려하지 않으면, 직전계절의 유입량에 대한 다음 계절 유입량의 확률밀도함수를 계산하고, 그 중 확률밀도함수 값이 가장 큰 값을 예측 유입량으로 결정하게 된다. 예를 들면, Fig. 6.7은 봄철 유입량이 100 백만 m^3 일 때 여름철 유입량의 확률분포를 나타낸 그림이다. 가뭄 예보 결과를 고려하지 않으면, 봄철 유입량이 100 백만 m^3 일 때 여름철 예측 유입량은 확률밀도함수 값이 가장 큰 약 480 백만 m^3 이 된다. 이러한 방법을 직전 계절의 모든 유입량에 대해 실시하면 직전 계절 유입량에 대한 다음 계절 예측 유입량 곡선이 결정된다.

Fig. 6.8 ~ Fig. 6.11은 안동댐의 계절별 유입량 예측 곡선이며 검은 점들은 2011년부터 2016년까지 각 계절별 관측 유입량이다. 안동댐의 봄철과 여름철 유입량 예측 곡선의 기울기는 계절 간 유입량의 상관계수의 영향으로 인해 다른 두 계절의 예측 유입량 곡선의 기울기에 비해 크다.

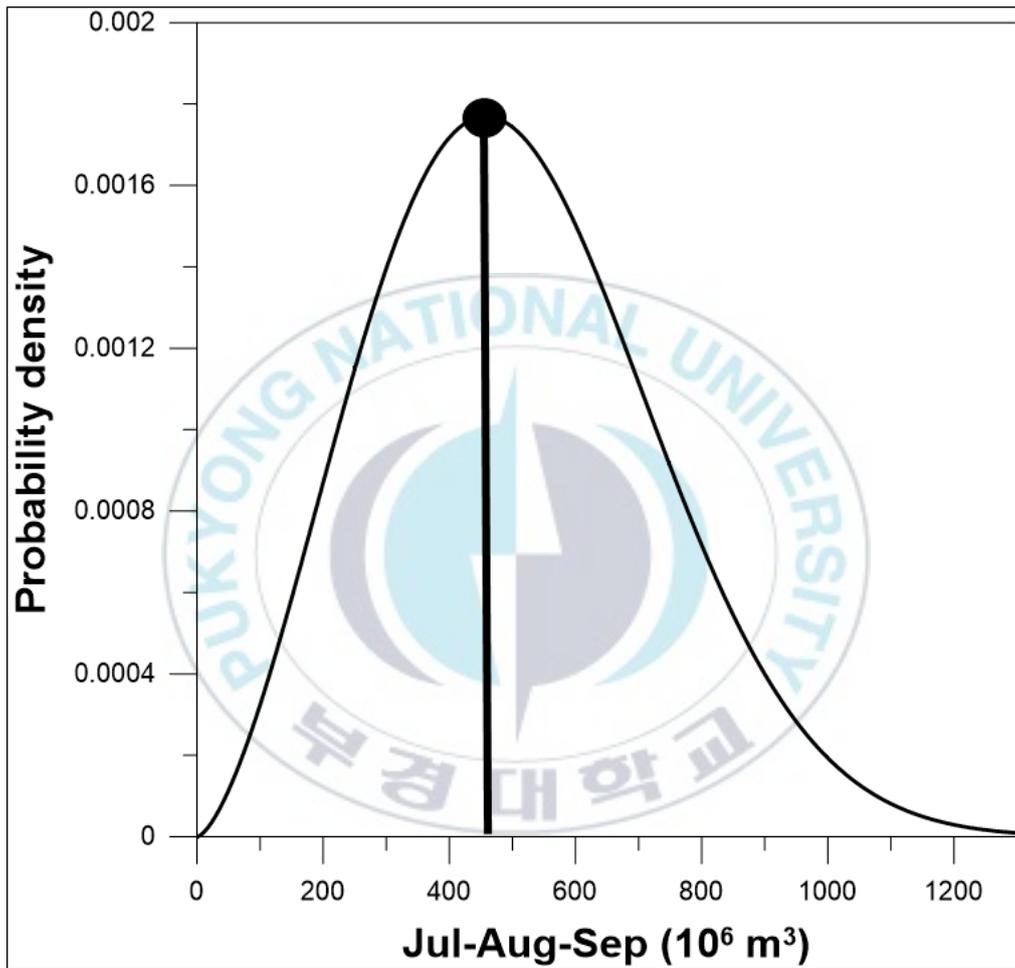


Fig. 6.7 Probability density function of summer inflow volume conditioned on a spring inflow volume of $100 \times 10^6 \text{ m}^3$

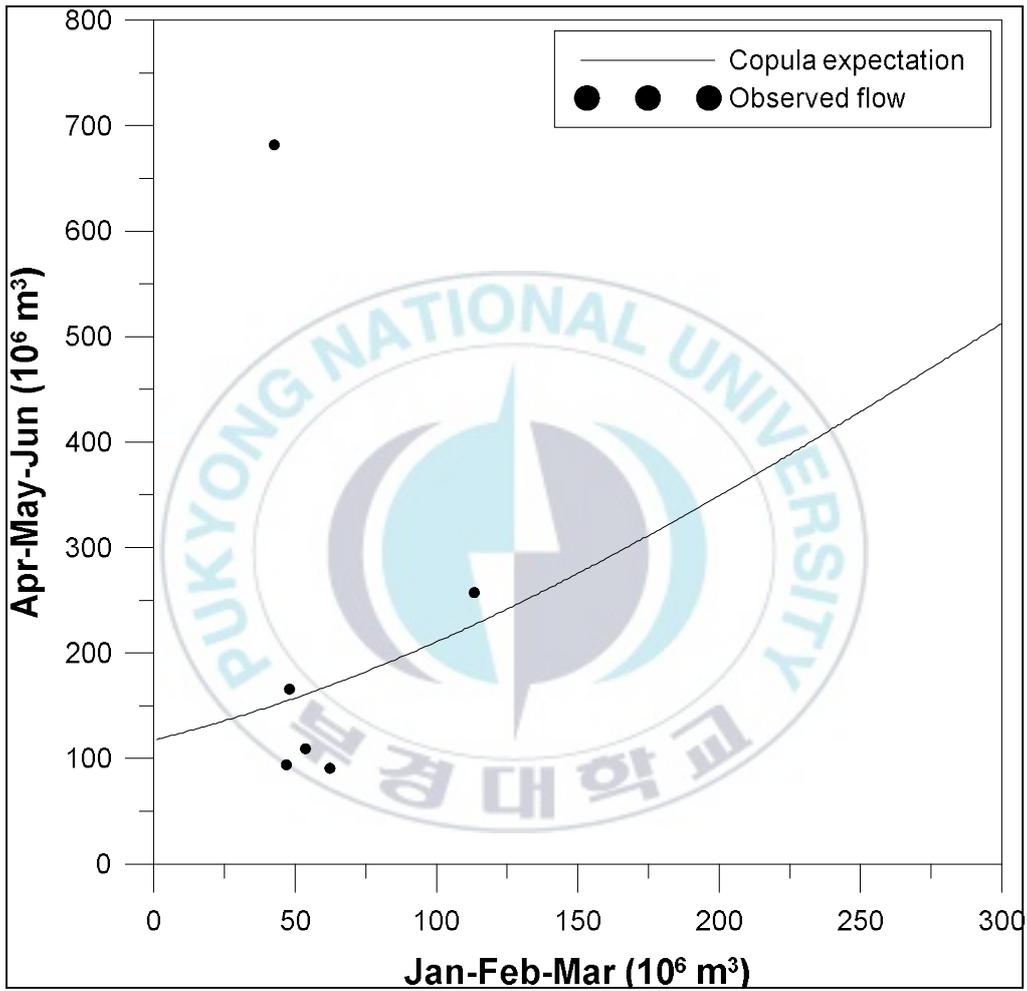


Fig. 6.8 Spring inflow forecasting curve

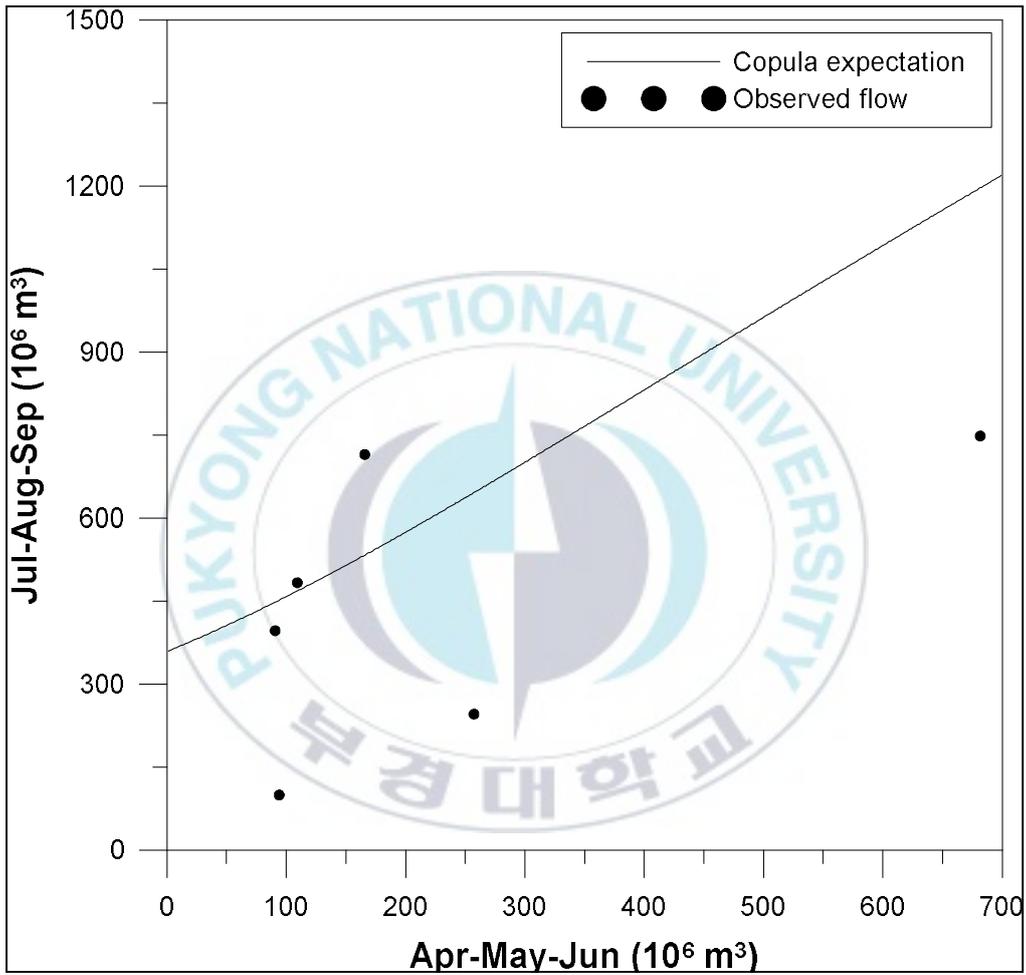


Fig. 6.9 Summer inflow forecasting curve

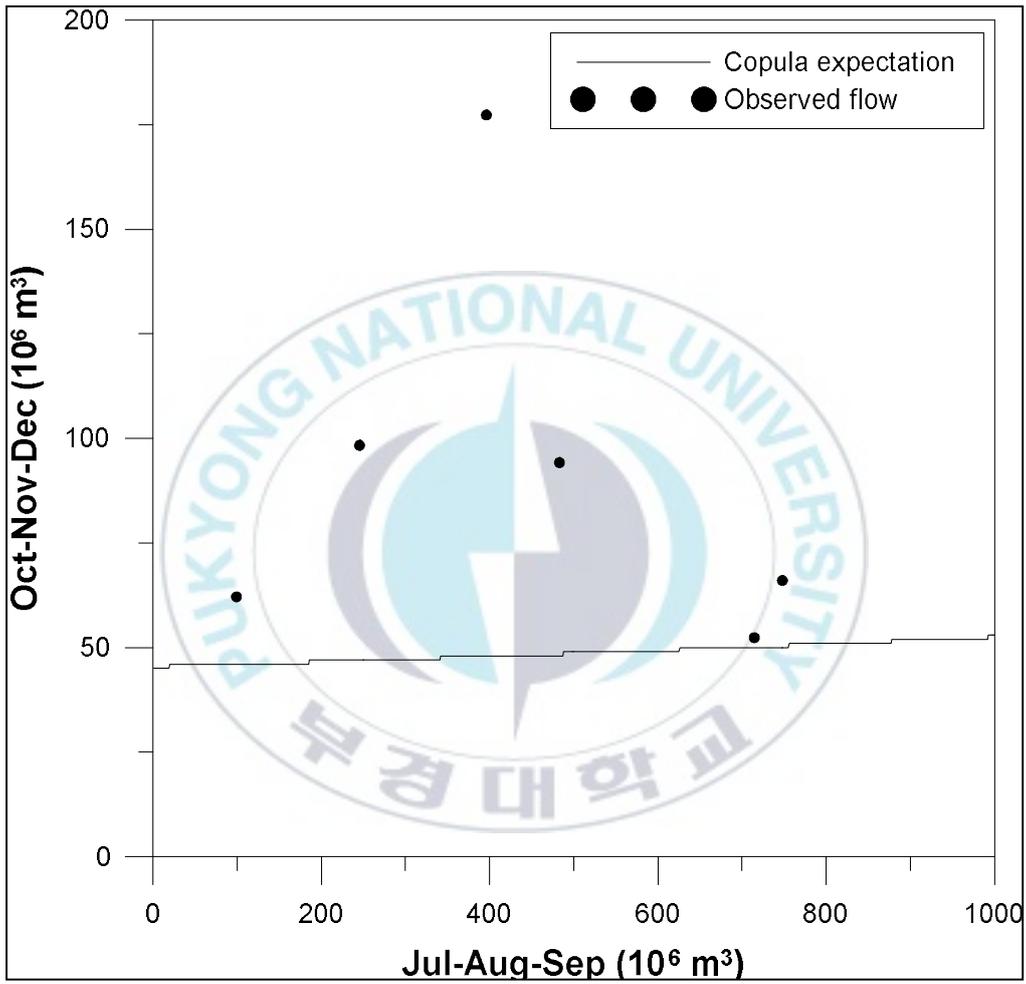


Fig. 6.10 Fall inflow forecasting curve

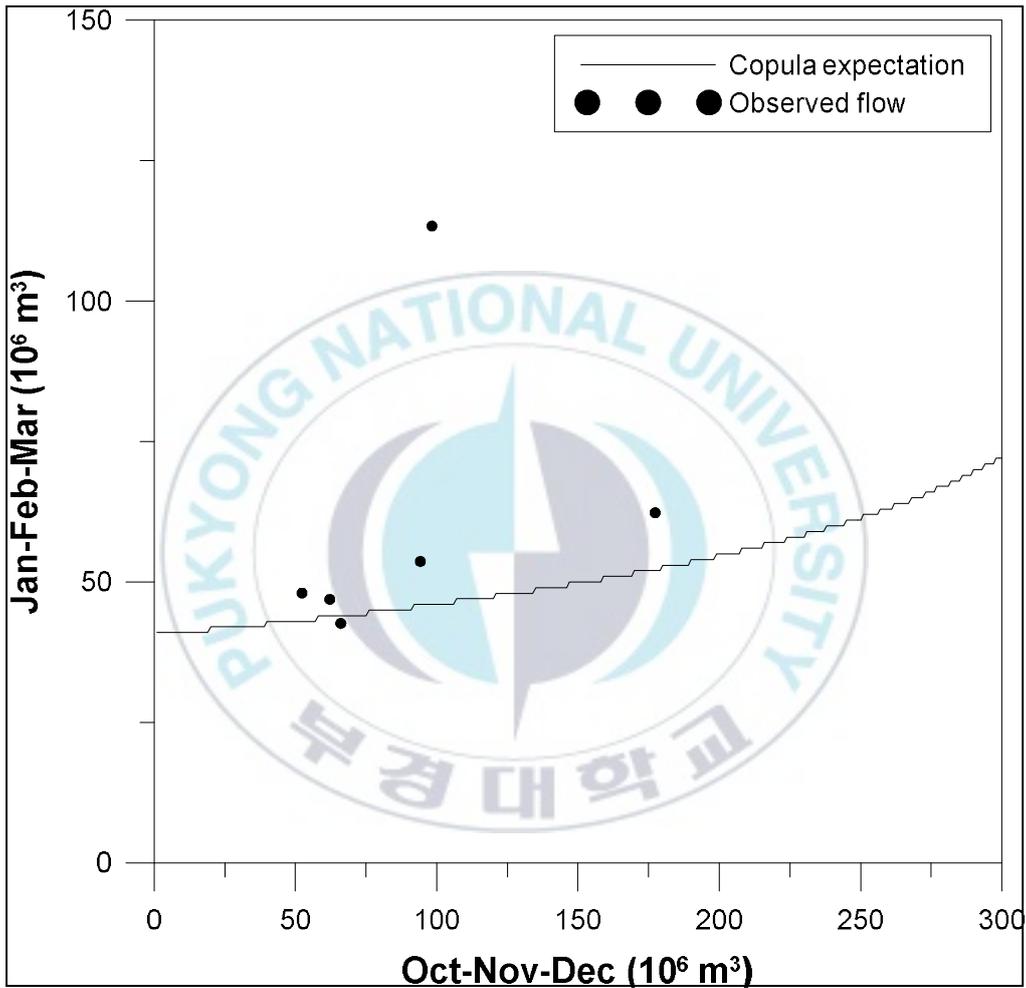


Fig. 6.11 Winter inflow forecasting curve

가뭄 예보 결과를 고려하지 않고 2011년부터 2016년까지 안동댐의 계절별 유입량을 예측한 결과는 Fig. 6.12와 같으며, Table 6.4는 여름철 및 전체 기간 유입량 예측의 절대오차를 계산한 결과이다. Fig. 6.12를 보면, 안동댐에 가뭄이 발생하지 않았던 2012년, 2014년 및 2016년에 예측된 유입량은 관측값과 큰 차이가 없었다. 2011년 여름의 경우, 봄철 유입량이 다른 년도에 비하여 상당히 컸기 때문에, 봄철 유입량과 상관계수가 높은 여름철의 유입량도 상당히 크게 예측되어 오차가 발생하였다. 또한, 안동댐에 D2에 해당하는 가뭄이 발생했던 2013년 여름과 D4에 해당하는 매우 큰 가뭄이 발생했던 2015년 여름철 예측 유입량은 관측값에 비해 매우 컸다. Table 6.4를 보면, 2013년과 2015년 여름의 절대오차는 각각 169.7 %, 355.6 %로 나타났다. 이는 봄과 여름의 유입량 간 상관계수가 높고, 두 해 모두 봄철 유입량이 아주 적지는 않았지만 여름에 가뭄이 발생해 발생한 오차이다.

다음 장에서는 과대 예측된 2013년 여름과 2015년 여름의 유입량을 조금 더 관측값과 유사하게 예측하기 위해 가뭄 예보 결과를 고려한 계절별 유입량을 예측하고, 그 결과를 본 절의 결과와 비교하고자 한다. 만약 2013년과 2015년 여름에 가뭄이 예보된다면, 기존 예측 유입량보다 적은 유입량을 예측해 좋은 결과가 도출될 것으로 판단된다.

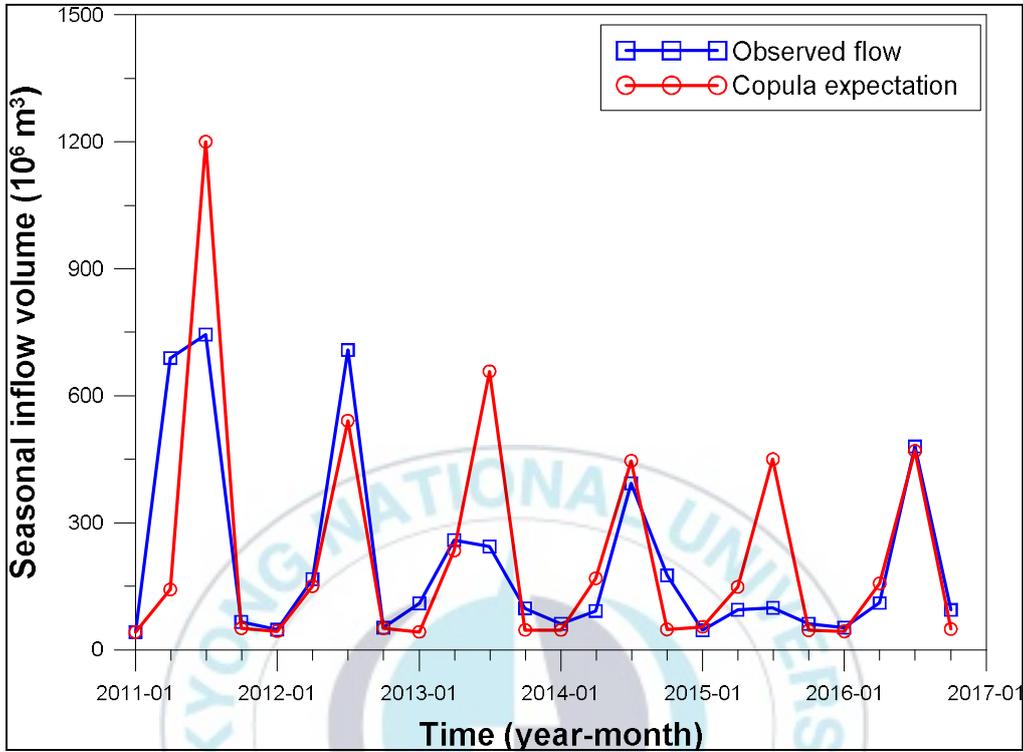


Fig. 6.12 The result of seasonal inflow forecasting without drought forecast

Table 6.4 The errors of seasonal inflow forecasting without drought forecast

Case	Absolute error for summer 2013 (%)	Absolute error for summer 2015 (%)	Range of absolute error for summer (%)	Range of absolute error for all seasons (%)
Seasonal inflow forecasting without drought forecast	169.7	355.6	2.2 ~ 355.6	1.2 ~ 355.6

제7장 SSI를 이용한 계절별 가뭄 예보

7.1 안동댐의 계절별 SSI 산정

계절별 가뭄 예보를 실시하기 위하여 각 계절의 SSI가 계산되어야 한다. 앞서 결정한 각 계절의 확률분포형을 이용해 계절별 SSI를 계산한 결과는 Fig. 7.1과 같다. Fig. 7.1에서 검은색으로 채색된 부분은 SSI가 0 이하인 부분으로, Table 5.1의 기준에 따라 가뭄이라고 정의할 수 있는 구간을 나타낸다. Fig. 7.1을 보면, 1977년부터 2010년까지 가뭄 상태 D3에 해당하는 가뭄은 약 5회 정도 발생하였으며, 가뭄 상태 D4에 해당하는 가뭄은 2009년에 1회 발생하였다. 또한 유입량 예측 결과 검증기간인 2011년 ~ 2016년 중에는 2013년 여름에 가뭄 상태 D2에 해당하는 가뭄이 발생하였으며, 2015년 여름에는 가뭄 상태 D4에 해당하는 큰 가뭄이 발생하였다.

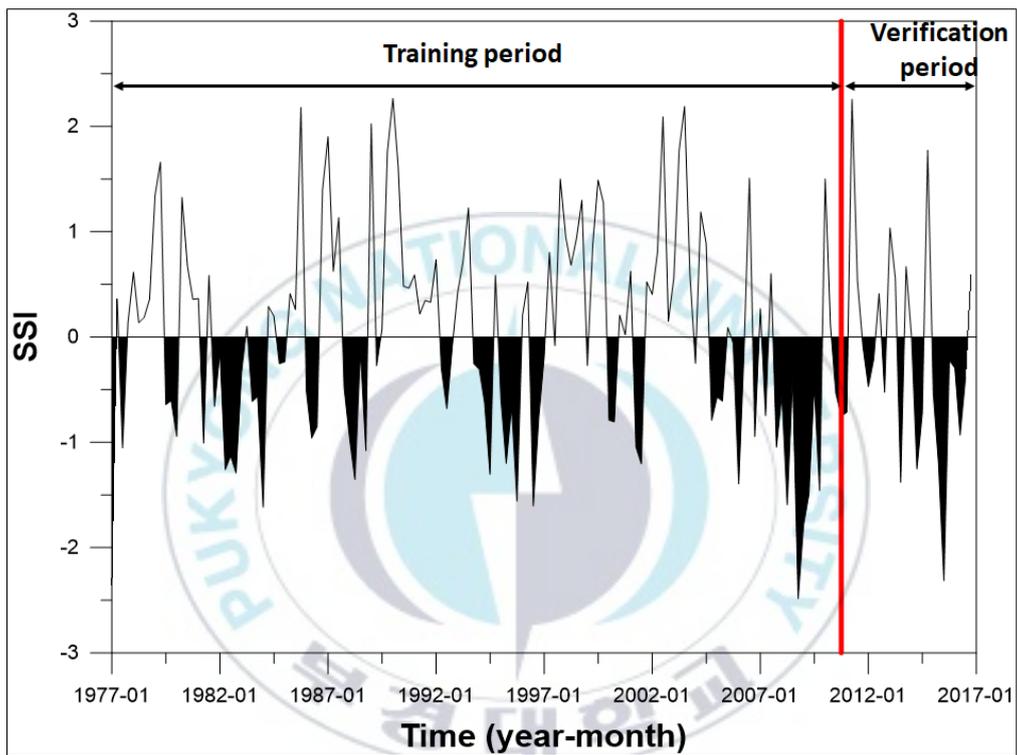


Fig. 7.1 Seasonal SSI for Andong Dam

7.2 SSI를 이용한 안동댐의 계절별 가뭄 예보 결과

본 절에서는 안동댐의 SSI를 이용하여 계절별 가뭄 발생 확률을 계산하였다. Gaussian copula 함수를 계산하기 위한 연속된 두 계절 간 SSI의 상관계수는 Table 7.1과 같다. Table 7.1을 보면, 연속된 두 계절 간 SSI의 상관계수는 연속된 두 계절 간 유입량의 상관계수와 유사하게 봄과 여름, 겨울과 봄철 간 SSI의 상관계수가 크게 계산되었다.

Table 7.1 The correlation coefficient of each season's SSI to that of prior season

	Spring – Summer	Summer – Fall	Fall – Winter	Winter – Spring
Correlation coefficient	0.4831	0.1830	0.2419	0.3539

가뭄 예보를 실시하기 이전에, 자료의 총 보유기간을 대상으로 가뭄 예보의 정확성을 평가하였다. 가뭄 예보 기준을 50 %로 설정한 가뭄 예보 결과, 가뭄 예보의 성공확률은 65 %로 3번의 가뭄 중 2번의 가뭄을 예측하는데 성공한 것으로 계산되었다.

2011년 봄부터 2016년 가을까지 각 계절의 가뭄 발생 확률을 계산한 결과는 Fig. 7.2와 같다. Fig 7.2을 보면, D2 상태에 해당하는 가뭄이 발생한 2013년 여름에 가뭄이 발생할 확률은 50 % 이하로 계산되었다. 반면에 D4 상태에 해당하는 큰 가뭄이 발생한 2015년 여름에 가뭄이 발생할 확률은 50 %보다 훨씬 크게 계산되었다. D2 상태의 가뭄을 예보하지 못한 이유는 2013년 봄철 유입량이 평균보다 많았기 때문이다. 다시 말하면, 2013년 봄에 관측된 유입량이 많았기 때문에 봄의 유입량과 상관계수가 높은 여름철 유입량이 많을 것이라 예상되어 가뭄 발생 확률이 작게 계산된 것이다. 반면 2015년 여름의 경우, 봄철 유입량이 적었기 때문에 가뭄 발생 확률이 크게 계산되었다. 즉, 본 연구에서 사용한 가뭄 예보는 직전 계절의 유입량이 많았지만 갑자기 발생한 ‘적당한 수준’의 가뭄을 예측하는데 한계가 있었다. 하지만 가뭄이 연속해서 발생하는 도중 발생한 매우 큰 가뭄은 적절히 예측하였다.

2015년 여름에 가뭄이 발생할 것으로 예보되었기 때문에, 이 때의 예측 유입량을 단순히 다음 계절 유입량의 확률분포에서 확률밀도함수 값이 가장 큰 유입량으로 결정하는 것은 부적절하다고 판단된다. 이에 다음 장에서는 가뭄 예보 결과를 고려한 계절 유입량 예측을 실시하고, 그 결과를 제시하고자 한다.

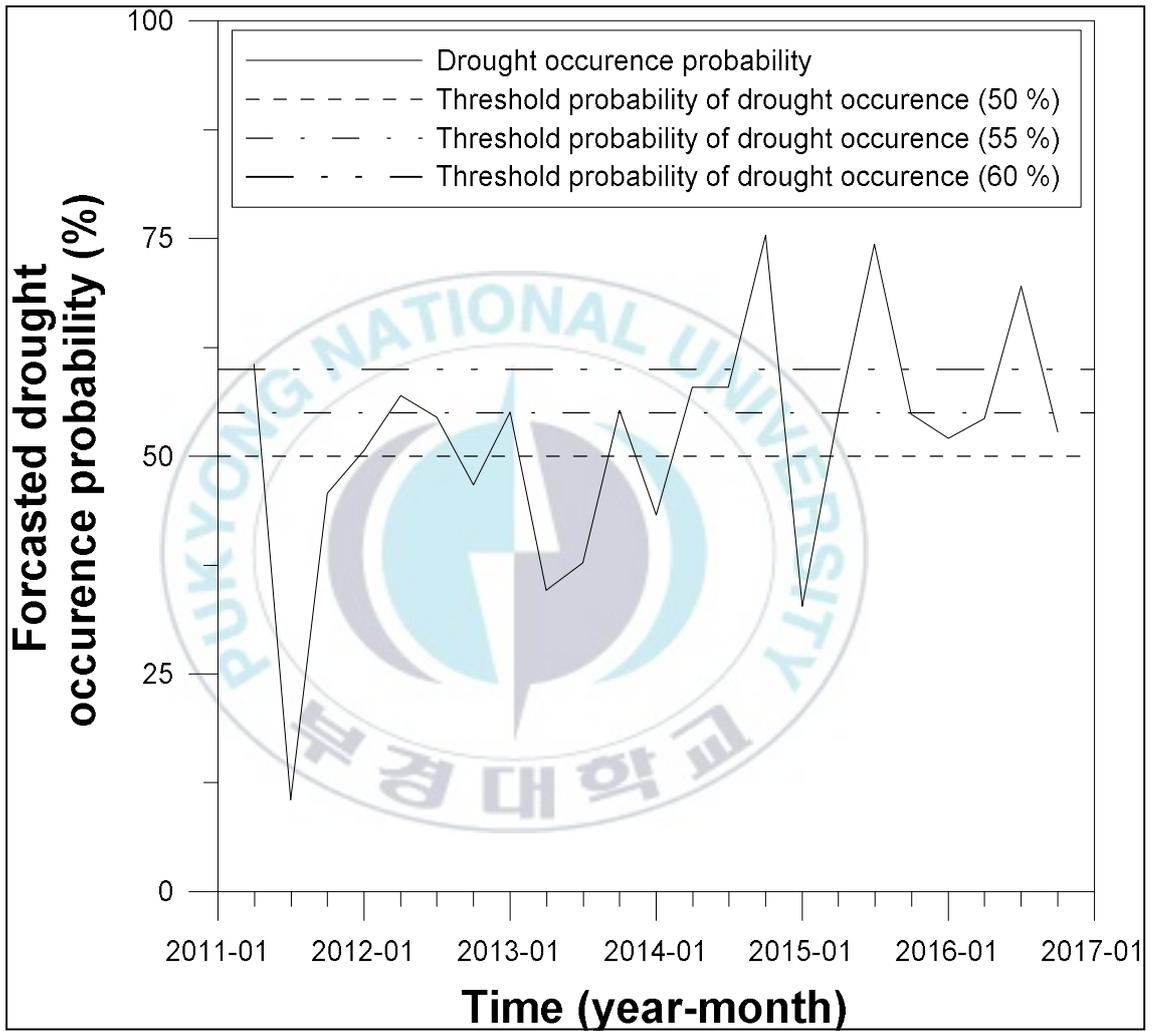


Fig. 7.2 The result of seasonal drought forecast

제8장 Copula 함수 기반의 베이지안 네트워크와 가뭄 예보를 결합한 계절별 유입량 예측

8.1 안동댐의 가뭄 단계별 계절 유입량 예측 곡선

가뭄 예보 결과를 고려한 유입량 예측을 위하여 각 가뭄 단계를 대표하는 유입량을 추정하여야 한다. 이를 위해 안동댐의 가뭄 단계별 계절 유입량 예측 곡선이 추정되었다. 안동댐의 가뭄 단계별 계절 유입량 예측 곡선은 아래와 같은 방법으로 추정되었다.

- (1) Fig. 8.1과 같은 다음 계절 유입량의 확률분포를 이용해 특정 유입량에 해당하는 누가확률 계산.
- (2) 계산된 누가확률을 이용해 SSI 계산.
- (3) 계산된 SSI가 각 가뭄 단계(D1 ~ D4)의 하한값과 일치하면 그에 상응하는 유입량을 각 가뭄 단계를 대표하는 유입량으로 결정. (단, D4의 경우 SSI가 -2.5인 유입량)
- (4) (1) ~ (3)을 모든 계절과 유입량을 대상으로 반복.

위와 같은 방법으로 결정된 가뭄 단계별 계절 유입량 예측 곡선은 Fig. 8.2 ~ Fig 8.5와 같다. Fig. 8.3을 보면, D2와 D3에 해당하는 곡선 근처에 위치하고 있는 두 점이 있다. 이 두 점들은 2013년과 2015년 여름에 관측된 유입량이다. 이 두 점에 해당하는 유입량은 2013년과 2015년 봄에 가뭄이 발생하였고 여름에 가뭄이 예보되었다면 예측할 수 있었을 것으로 판단된다.

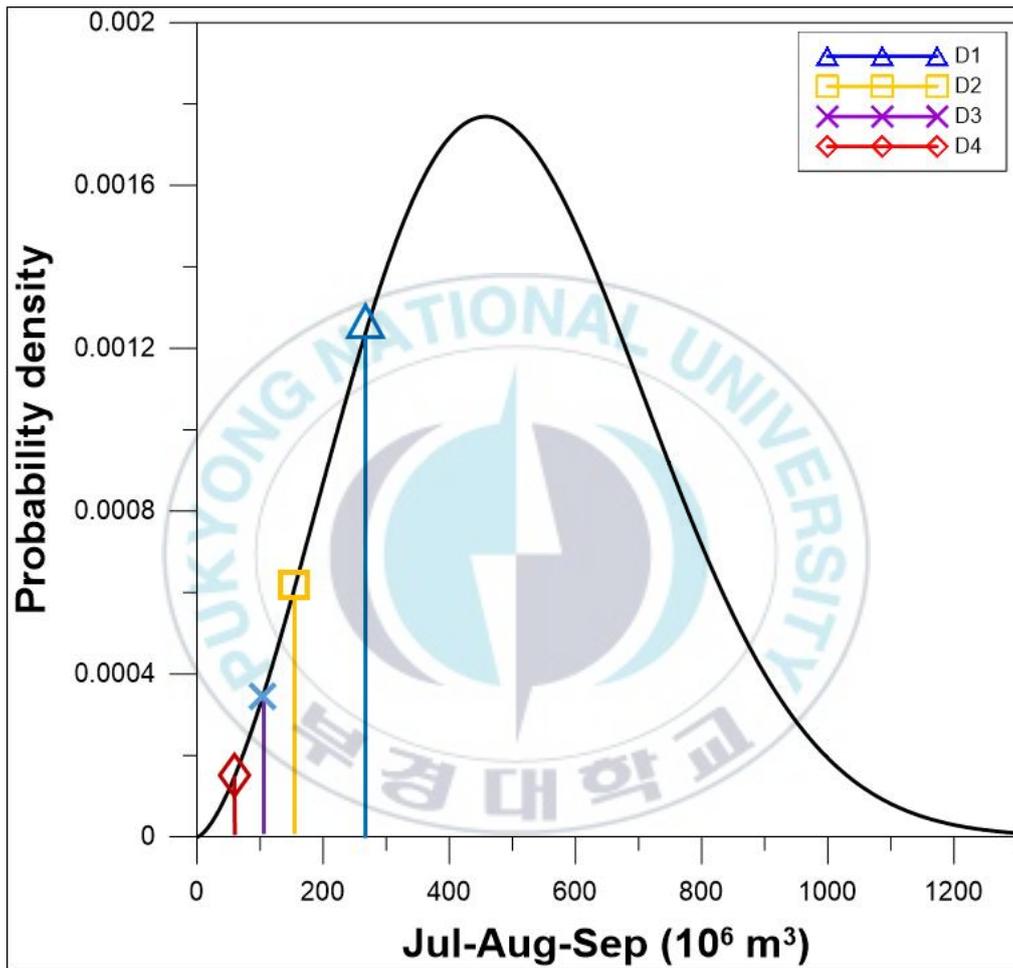


Fig. 8.1 Example of estimating representative inflow corresponding to lower bounds of SSI ranges of drought stages

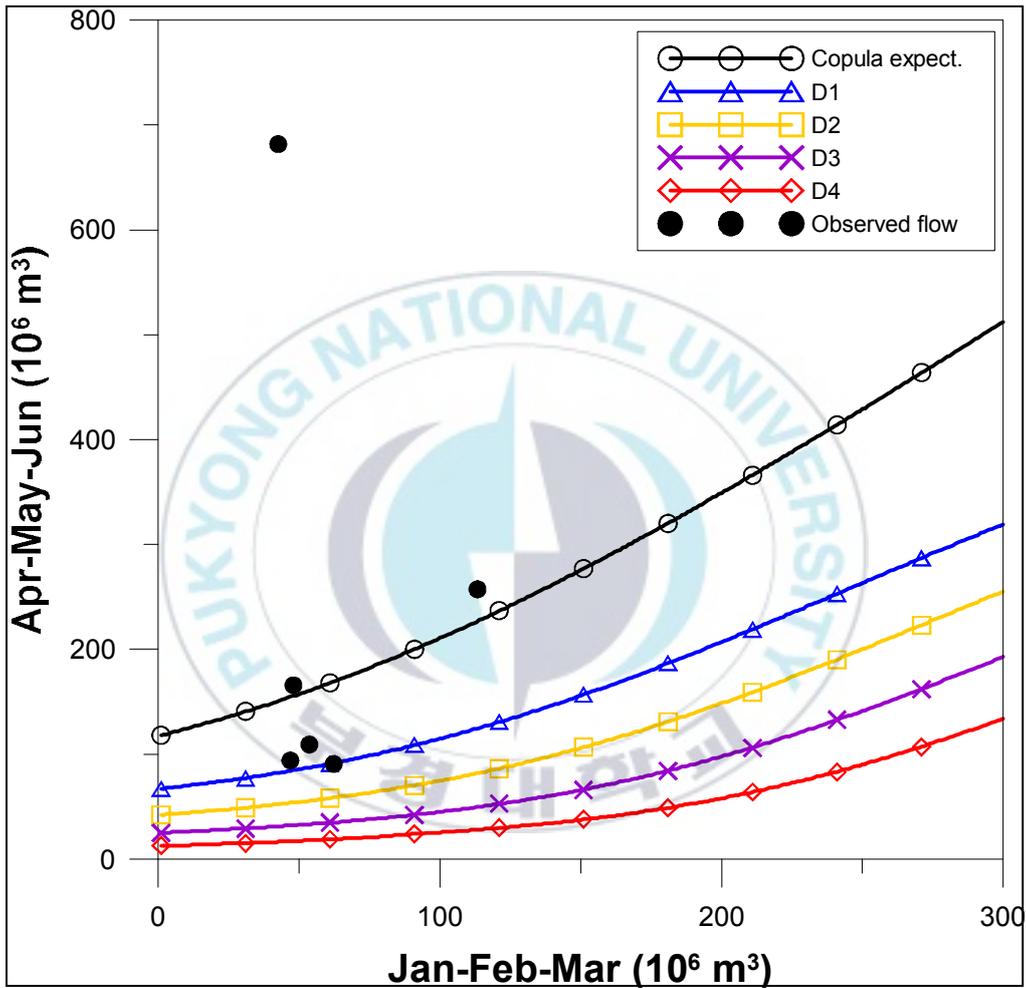


Fig. 8.2 Spring inflow forecasting curves conforming to drought stages

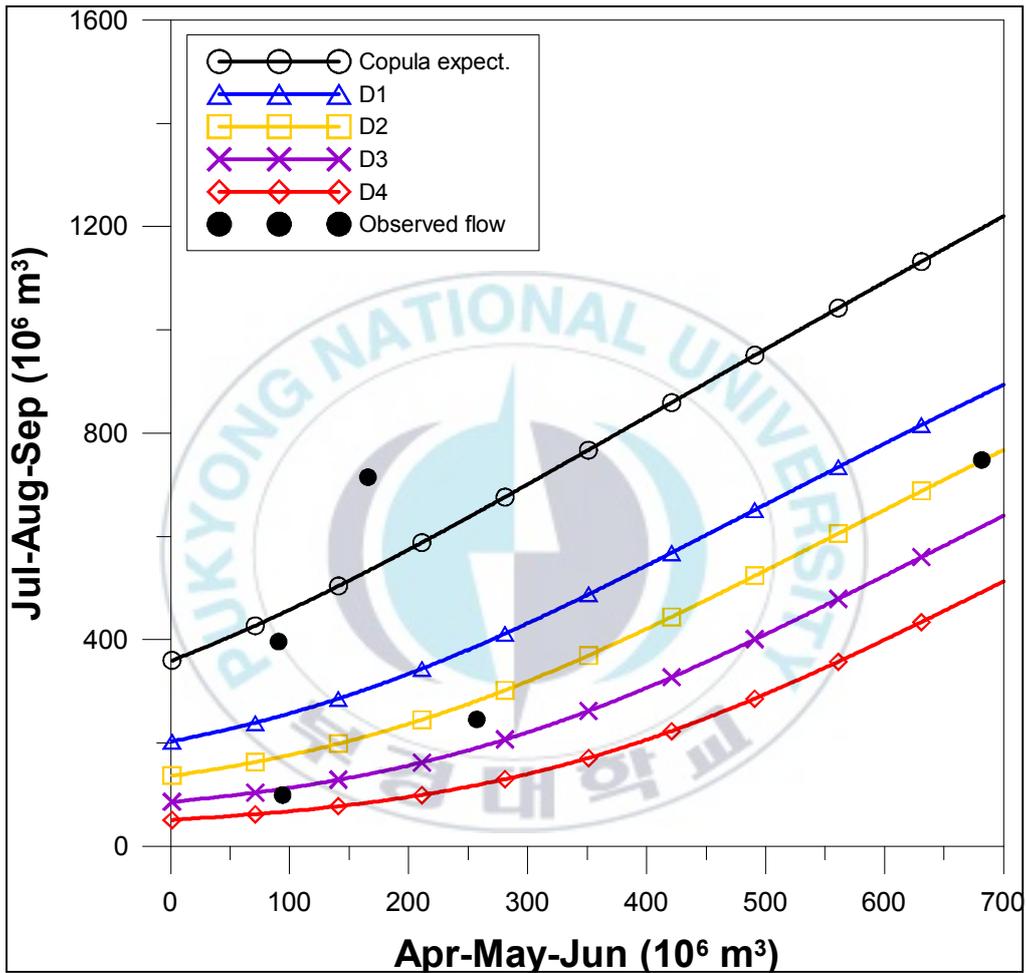


Fig. 8.3 Summer inflow forecasting curves conforming to drought stages

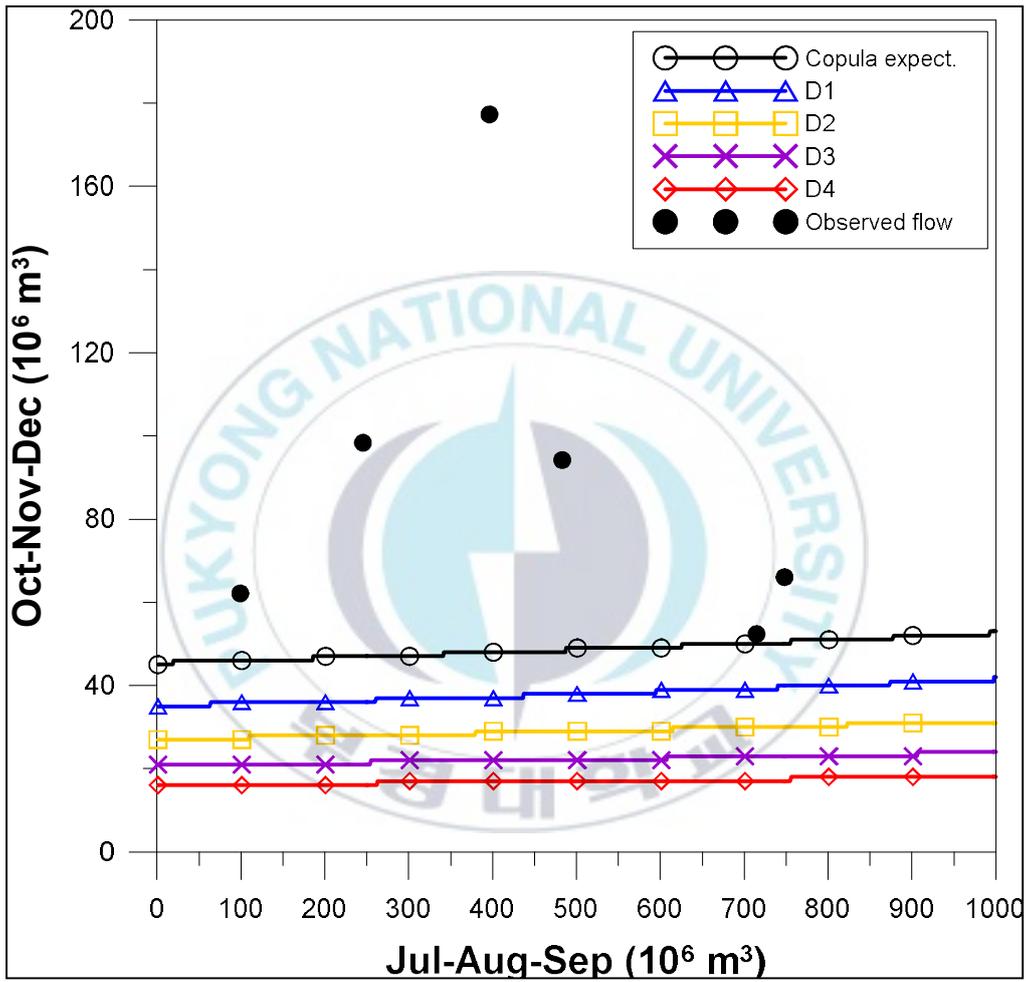


Fig. 8.4 Fall inflow forecasting curves conforming to drought stages

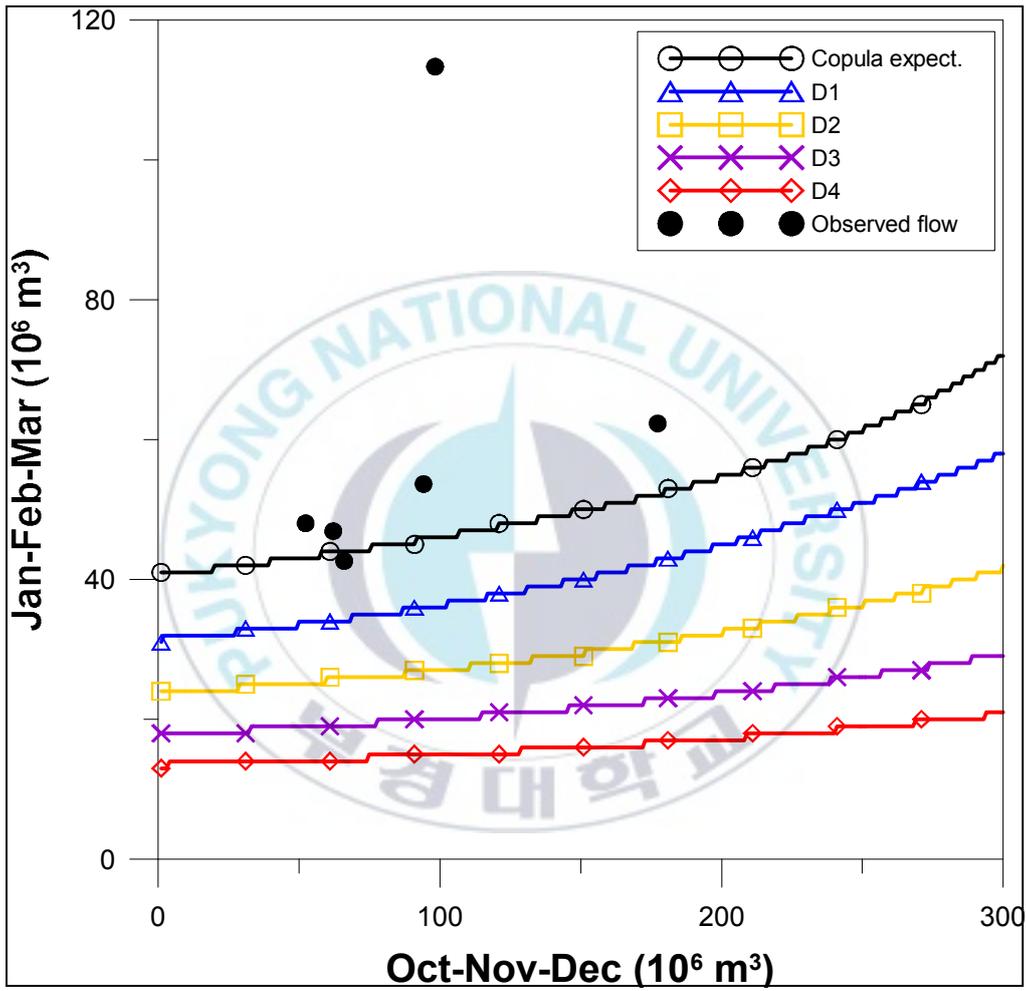


Fig. 8.5 Winter inflow forecasting curves conforming to drought stages

8.2 베이지안 네트워크만 사용한 계절별 유입량 예측과 결합예측의 결과 비교

가뭄 단계별 계절 유입량 예측 곡선을 만들고 나면, 가뭄이 예보되었을 때 어느 곡선을 선택하여 유입량을 예측할지가 결정되어야 한다. 댐을 운영하는 데, 특히 이수의 목적에서, 연속된 두 계절의 가뭄은 물의 공급에 아주 치명적일 수 있다. 더욱이 여름철에 강우가 집중되는 우리나라의 기후특성상 봄과 여름철에 연속되는 가뭄은 매우 치명적일 수 있다. 이러한 이유로, 본 연구에서는 직전 계절에 가뭄이 발생하였고 다음 계절에 가뭄이 예보되면, 다음 계절 예측 유입량은 직전 계절 가뭄 단계보다 한 단계 격상된 가뭄 단계에 해당하는 유입량 예측 곡선을 이용해 결정하였다. 예를 들면, 만약 봄철에 안동댐의 가뭄상태가 D1이었고, 여름철에 가뭄이 예보되었다면 여름철 유입량은 Fig. 8.3에서 사각형으로 표시된 D2 곡선을 이용해 예측하는 것이다. 반대로, 다음 계절에 가뭄이 예보되지 않았다면 다음 계절 유입량은 원으로 표시된 곡선을 이용해 예측하였다.

가뭄 예보 기준을 50 %로 설정한 가뭄 예보 결과와 가뭄 단계별 계절 유입량 예측 곡선을 이용해 안동댐의 2011년부터 2016년까지 계절별 유입량을 예측한 결과는 Fig. 8.6과 같다. Fig. 8.6을 보면, 가뭄 예보를 결합한 유입량 예측 결과는 가뭄 예보를 결합하지 않은 예측 결과에 비해 D4 상태에 해당하는 가뭄이 발생했던 2015년 여름철 유입량을 적절히 예측하였다. 하지만, D2 상태에 해당하는 가뭄이 발생한 2013년 여름의 유입량은 앞서 언급한 가뭄 예보의 한계로 인해 예측하지 못하였다. 또한, 대부분의 계절에 적은 유입량을 예측하는 경향이 나타났다. 이러한 경향을 보이는 원인은 가뭄 예보 기준을 50

%로 설정한 것과 봄과 여름철 유입량 간 높은 상관관계수인 것으로 판단된다. 안동댐은 봄과 여름철 유입량의 상관관계수가 높기 때문에 봄철 유입량에 여름철 예측 유입량이 선형적으로 변한다. 이러한 특성은 연속된 두 계절 간 SSI의 상관관계수에도 영향을 끼치고, 봄철 유입량이 아주 약한 가뭄 단계에 해당하는 유입량이라도 여름에 가뭄이 발생할 확률이 커지게 된다. 즉, 가뭄 예보 기준을 50 %로 설정하는 것은 빈번한 가뭄 예보를 야기할 수 있음을 의미한다. 실제로, 안동댐의 유입량 자료가 존재하는 40년 중 실제 가뭄지수가 0 이하인 계절은 160개 계절 중 67회였다. 반면 가뭄 예보 기준을 50 %로 설정했을 때 가뭄을 예보한 횟수는 실제 가뭄 발생횟수에 비해 15회나 많은 82회였다.

이러한 경향을 개선하기 위한 방법은 기존 50 %였던 가뭄 예보 기준을 상향시키는 것이다. 가뭄 예보 기준이 높아지면, 직전 계절의 관측 유입량이 평균 유입량에 비해 아주 조금 적더라도 다음 계절에 가뭄이 예보되지 않아 빈번한 가뭄 예보를 하는 경향이 개선될 것이다. 이에 기존 50 %였던 가뭄 예보 기준을 55 %와 60 %로 상향시켜 자료 보유 기간을 대상으로 가뭄 예보 횟수를 계산하였다. 우선, 가뭄 예보 기준을 55 %와 60 %로 상향하였을 때 가뭄 예보의 성공률은 각각 67.2 %, 64.5 %로 가뭄 예보 기준이 50 %였을 때와 비슷한 성공률을 보였다. 가뭄 예보 기준을 55 %로 상향하면, 총 160개 계절 중 58회의 계절에 가뭄을 예보하였다. 반면 가뭄 예보 기준을 60 %로 설정한 결과 총 160개 계절 중 31회의 계절에 가뭄을 예보하였다. 이러한 결과는 가뭄 예보 기준을 60 %까지 상향시키면 가뭄을 예보하는데 너무 엄격해질 수 있음을 의미한다.

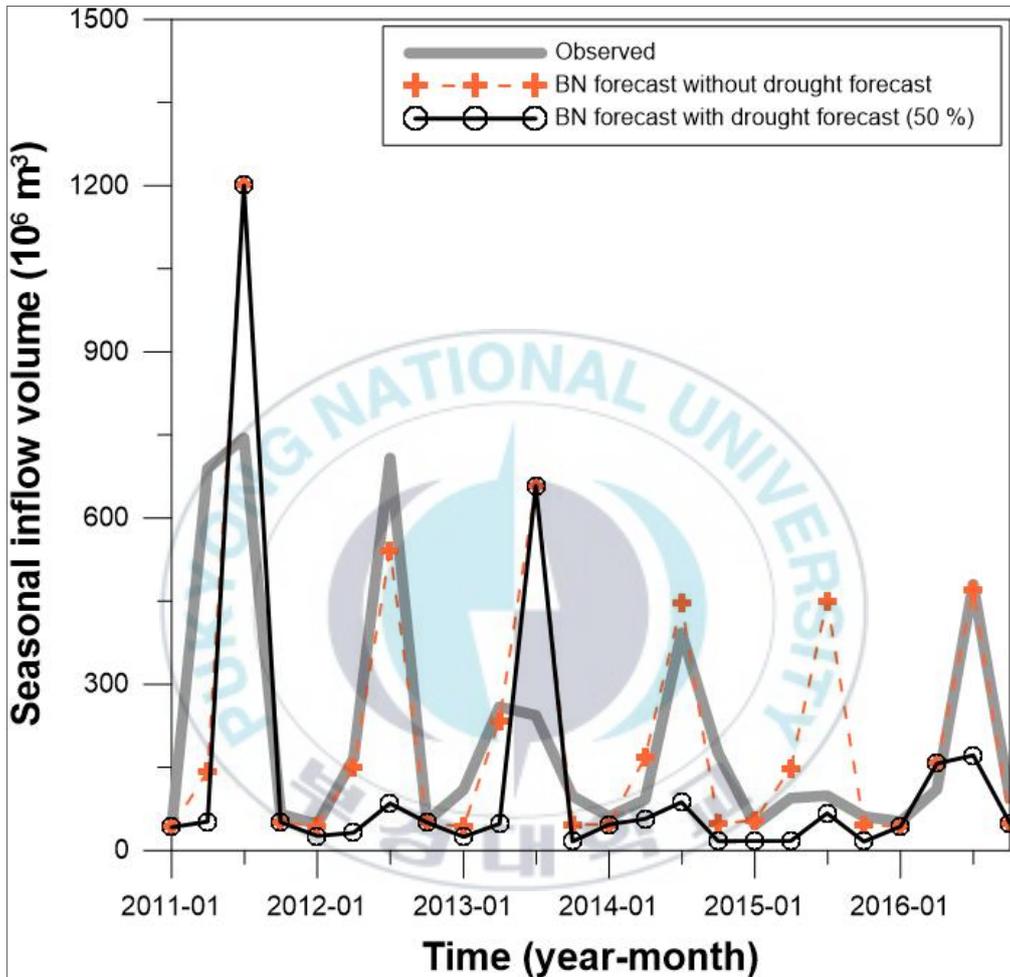


Fig. 8.6 The results of seasonal inflow forecasting without or with drought forecast under the threshold probability of drought occurrence to 50 %

가뭄 예보 기준을 55 %와 60 %로 상향시켜 안동댐의 계절별 유입량을 예측한 결과는 Fig. 8.7과 같으며, Table 8.1은 모든 경우의 예측 결과를 절대오차를 이용해 나타낸 표이다. Fig. 8.7과 Table 8.1을 보면, 가뭄 예보 기준이 55 %일 때와 60 %일 때 예측결과는 큰 차이를 보이지 않는다. 하지만 매 계절 적은 값을 예측하던 경향은 개선된 것을 알 수 있다. 특히 2012년 여름은 실제 유입량과 비슷한 유입량을 예측하였다. 또한, D4 상태에 해당하는 심한 가뭄이 발생했던 2015년 여름철 유입량 예측의 절대오차가 9.4 %로 매우 크게 감소하였다.



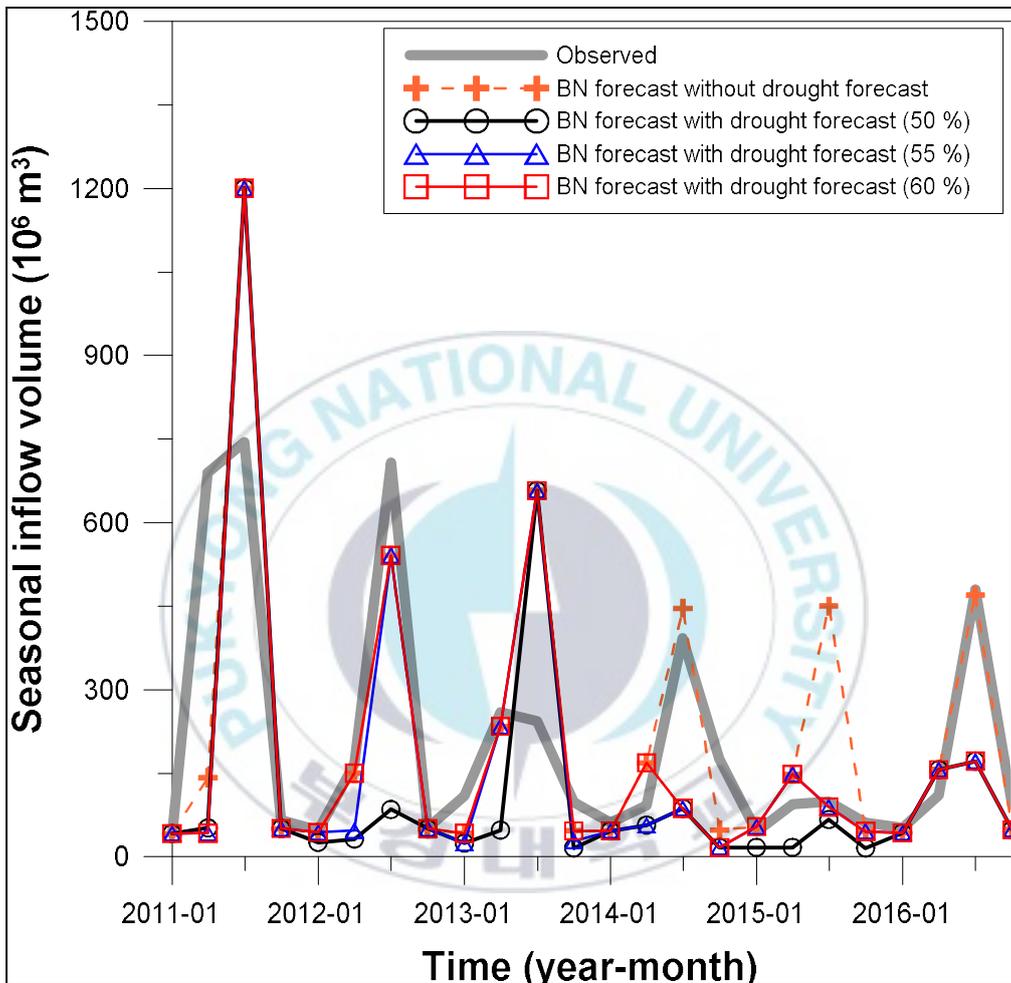


Fig. 8.7 The results of seasonal inflow forecasting without or with drought forecast under the threshold probability of drought occurrence range 50 % to 60 %

Table 8.1 Seasonal inflow forecasting errors

Cases	Absolute error for summer 2013 (%)	Absolute error for summer 2015 (%)	Range of absolute error for summer (%)	Range of absolute error for all seasons (%)
BN forecast without drought forecast	169.7	355.6	2.2 ~ 355.6	1.2 ~ 355.6
BN forecast with drought forecast by 50 % criteria	169.7	32.1	32.1 ~ 169.7	1.2 ~ 169.7
BN forecast with drought forecast by 55 % criteria	169.7	9.4	9.4 ~ 169.7	1.2 ~ 169.7
BN forecast with drought forecast by 60 % criteria	169.7	9.4	9.4 ~ 169.7	1.2 ~ 169.7

Fig. 8.7을 보면, 2012년 여름철에 가뭄 예보 기준이 50 %인 가뭄 예보 결과를 고려한 유입량 예측 결과와 가뭄 예보 기준이 55 %인 가뭄 예보를 고려한 유입량 예측 결과의 차이가 매우 크다. 그 이유는 연속적인 가뭄 예보에 있는 것으로 판단된다. 자세히 설명하면, Fig. 7.2에 나타난 계절별 가뭄 발생 확률을 보면 2012년 겨울에 가뭄이 발생할 확률은 50 %보다 크고 55 %보다 작다. 이로 인해 가뭄 예보 기준을 50 %로 설정하게 되면 2012년 겨울에 가뭄이 예보 되어 한 단계 격상된 가뭄단계에 해당하는 예측을 실시하게 된다. 더불어, 2012년 봄에 가뭄이 발생할 확률은 55 %이상이고, 2012년 여름에 가뭄이 발생할 확률도 50 %보다 크고 55 %보다 작다. 이러한 이유로 가뭄 예보 기준을 50 %로 설정하면 2012년에는 3계절 연속적인 가뭄이 예보가 되어 2012년 여름에 아주 적은 유입량이 예측된다. 반면 가뭄 예보 기준을 55 %로 설정하면 가뭄이 예보되지 않기 때문에 이러한 차이가 발생한 것이다.

제9장 결론 및 향후 연구

9.1 결론

본 연구에서는 가뭄 예보 결과를 고려해 다목적 댐의 계절별 유입량을 예측하기 위해 가뭄 예보 결과를 결합한 copula 함수 기반의 베이지안 네트워크를 사용하였다. 연구 대상 댐은 안동댐이며 2011년부터 2016년의 유입량 자료가 예측 결과의 적절성을 평가하기 위해 사용되었다. 1년을 4계절로 나누었으며 각 계절은 봄(4월 ~ 6월), 여름(7월 ~ 9월), 가을(10월 ~ 12월), 겨울(1월 ~ 3월)이다. 계절별 유입량 자료의 확률분포형 결정 결과, 여름철 유입량 자료는 Weibull 분포, 나머지 계절 유입량 자료는 lognormal 분포가 가장 적합한 것으로 결정되었다. 겨울과 봄, 봄과 여름의 유입량 간 상관계수는 높았으나 나머지 두 경우는 유입량 간 상관계수가 작게 계산되었다. Copula 함수 기반의 베이지안 네트워크를 이용해 안동댐의 가뭄 단계별 계절 유입량 예측 곡선을 결정한 후 가뭄지수 중 하나인 SSI를 이용해 계절별 가뭄을 예보하였다. 가뭄 예보는 직전 계절의 SSI에 대한 다음 계절 SSI의 확률밀도함수를 계산하고, SSI가 0을 초과하지 않을 비초과 확률과 가뭄 예보 기준을 비교하여 실시하였다.

연구 결과 얻은 결론들은 다음과 같다. 가뭄 예보를 결합하지 않은 계절별 유입량 예측 결과는 댐을 운영하기 위한 미래 유입량 자료로 매우 부적절하다는 것이다. 가뭄 예보를 결합하지 않는다면 단순히 직전 계절의 유입량에 대한 다음 계절의 유입량 중 확률밀도함수 값이 가장 큰 유입량을 예측 유입량으로 결정하게 된다. 이러한 특성상 가

물이 발생했던 계절의 유입량 예측 결과는 매우 불량하였다.

가뭄 예보를 결합할 때 가뭄 예보 기준을 잘 설정할 필요가 있다. 가뭄 예보 기준이 50 %인 가뭄 예보 결과를 고려한 유입량 예측 결과는 많은 계절에 적은 유입량을 예측하는 경향을 보였다. 하지만 가뭄 예보 기준을 55 %로 상향하면 가뭄을 예보하는 횟수가 줄어들어 적은 유입량을 예측하는 경향이 개선되었다. 가뭄 예보 기준을 60 %까지 상향시켜 보았으나 그 결과는 55 %와 거의 동일하게 도출되었다. 하지만 가뭄 예보 기준을 60 %까지 상향시키면 실제 가뭄 발생횟수에 비해 매우 작은 횟수의 가뭄을 예보하는 것을 확인하였다. 이는 가뭄 예보가 너무 엄격해질 수 있음을 의미하고, 가뭄 예보가 너무 엄격해지면 약한 가뭄이 연속해서 이어지는 상황에 대처할 수 없을 것으로 판단된다. 그러므로, 안동댐의 경우에는 가뭄 예보 기준이 55 %인 가뭄 예보 결과를 고려해 계절 유입량을 예측하는 것이 가장 적절한 것으로 사료된다.

결론적으로, 심한 가뭄 시 안동댐의 계절 유입량은 가뭄 예보 기준을 적절히 설정한 가뭄 예보와 copula 함수 기반의 베이지안 네트워크를 결합한 확률론적 예측을 이용하여 적절히 예측할 수 있었다.

9.2 향후 연구

향후에는 안동댐이 아닌 다른 댐들을 대상으로 본 연구에서 사용한 기법을 적용시켜 보며 댐의 수문학적 특성에 따라 최적의 가뭄 예보 기준을 찾아야 할 필요가 있다. 최적화 기법을 이용해 절대오차를 최소화하는 가뭄 예보 기준을 찾는 방법도 고려될 수 있다. 또한, 단순 유입량 예측 결과의 절대 오차가 아닌, 예측된 유입량을 이용한 댐 운영 결과를 비교해 하류의 물 부족량이 가장 적은 가뭄 예보 기준을 찾는 연구도 진행될 수 있다. 기상 예보의 신뢰성이 있다면, 이를 가뭄 예보에 활용할 수 있다. 현재 행정안전부에는 2015년에 발생한 큰 가뭄 이후 매 월 가뭄 예·경보를 실시하고 있다. 가뭄 예·경보는 전국의 현재 가뭄 상태와 1개월, 3개월 가뭄 전망을 제시한다. 현재는 가뭄 예·경보 자료가 많지 않지만, 이후에 자료가 쌓이고 나면 본 연구에서 활용한 가뭄 예보 대신 가뭄 예·경보 자료를 이용해 다음 계절의 가뭄을 예측하는 것이 가능할 지도 모른다. 본 연구에서 사용한 복잡한 과정을 거친 가뭄 예보 결과보다 쉽게 확인할 수 있는 가뭄 예·경보 자료를 이용해 가뭄 예보를 실시한다면 계절별 유입량을 예측하는 것이 조금 더 간단해질 것으로 사료된다.

참고문헌

- 강민석 (2014). SWAT 모형과 앙상블 기법을 이용한 안동댐 유입량 예측. 석사학위논문, 아주대학교.
- 강신욱 (2011). 차원변화 탐색기법을 이용한 장기 저수지 운영의 수위구간 결정 방법 개발 및 적용. 박사학위논문, 부경대학교.
- 권현한, 김장경, 이종석, 나봉길 (2012). “Bayesian 모형을 이용한 단일 사상 강우-유출 모형의 불확실성 분석.” **한국수자원학회 논문집**, 한국수자원학회, Vol. 45, No. 5, pp. 505-516.
- 김광훈, 진영규, 이상호 (2017a). “안동댐의 여름철 가뭄 단계별 예측 유입량 곡선 소개.” **물과 미래**, 한국수자원학회, Vol.50, No. 10, pp. 59-64.
- 김광훈, 진영규, 이상호, 장수형 (2017b). “미국 수자원 조사, 관리 기관의 앙상블 하천유량 예측기법 활용.” **물과미래**, 한국수자원학회. Vol. 50, No. 1, pp 59-66.
- 김덕정 (2015). 베이지안 네트워크의 능력모수 복원에 관한 모의실험 연구. 석사학위논문, 성신여자대학교.
- 김선호, 소재민, 강신욱, 배덕효 (2017). “베이지안 기법 기반의 댐 예측 유입량 산정기법 개발 및 평가.” **한국수자원학회 논문집**, 한국수자원학회, Vol. 50, No. 7, pp-489-502.
- 김영오, 정대일, 김형섭, 이길성 (2001). “앙상블 예측을 통한 물공급전망 개선방안.” **2001년 한국수자원학회 학술대회 논문집(1)**, 한국수자원학회, pp. 304-308.
- 김주철, 김정곤, 이상진 (2011). “확률기상예보를 이용한 중장기 ESP기

- 법 개선.” **한국수자원학회 논문집**, 한국수자원학회, Vol. 44, No. 10, pp. 843-851.
- 김진영, 김진국, 최병한, 권현한 (2015). “Bayesian network(BN)를 이
활용한 수문학적 댐 위험도 해석 기법 개발.” **한국수자원학회 논문
집**. 한국수자원학회, Vol. 48, No. 10, pp. 781-791.
- 김진영, 소병진, 김태웅, 권현한 (2016). “Copula 함수를 이용한 결합분
포 함수 추정 및 연관성 비교.” **한국수자원학회 논문집**, 한국수자원
학회, Vol. 49, No. 10, pp. 823-833
- 김한수, 김현식, 전근일, 강신욱 (2016). “2014년 ~ 2015년 가뭄의 평
가” **물과 미래**, 한국수자원학회, Vol. 49, No. 7, pp. 61-75.
- 남기훈 (2014). 베이지안 네트워크를 이용한 복합재난 위험성 평가에 관
한 연구. 박사학위논문, 인제대학교.
- 류재희, 이동률, 안재현, 윤용남 (2002). “가뭄평가를 위한 가뭄지수의
비교 연구.” **한국수자원학회 논문집**, 한국수자원학회, Vol. 35, No.
4, pp. 397-410.
- 박명우 (2016). 실시간 가뭄지수를 이용한 AR5 RCP 기후변화 시나리오
별 미래 가뭄전망. 석사학위논문, 부경대학교
- 선은정 (2010). 베이지안 네트워크를 이용한 신용평가에 영향을 주는 요
인에 관한 연구. 석사학위논문, 성균관대학교.
- 소재민 (2015). Copula 함수 기반의 이변량 결합가뭄지수 산정 및 가뭄
특성 분석. 석사학위논문, 세종대학교.
- 음형일 (2007). Sampling SDP와 ESP를 이용한 금강 수계 다목적댐군의
이수기 운영방안. 박사학위논문, 서울대학교.
- 이재수 (2006). 수문학. 구미서관
- 이재형 (2001). Copula 함수를 이용한 다변량분포의 VaR추정. 석사학위

- 논문, 성균관대학교.
- 이정한 (2006). 우리나라 강우의 권역별 대표확률분포형 결정에 관한 연구. 석사학위논문, 금오공과대학교.
- 유지영, 신지예, 김동균, 김태웅 (2013). “추계학적 강우발생모형과 Copula 함수를 이용한 가뭄위험분석.” **한국수자원학회 논문집**, 한국수자원학회, Vol. 46, No.8, pp. 425-437.
- 정준영 (2014). Copula 함수를 이용한 결합분포 함수 추정 및 연관성 비교. 석사학위논문, 고려대학교.
- 진영규 (2016). 앙상블 하천유량 예측과 용수 감량 공급 기법을 이용한 합천댐의 가뭄 시 운영 방안. 석사학위논문, 부경대학교.
- 진영규, 정태훈, 이상호, 강신욱 (2016). “앙상블 하천유량 예측 기법과 용수 감량공급 기법을 이용한 합천댐의 가뭄 시 운영 방안.” **한국방재학회 논문집**, 한국방재학회, 제16권, 1호, pp. 93-101.
- 진영규, 김광훈, 이상호 (2017). “댐 용수 공급 조정기준의 가뭄 대응능력 평가” **물과 미래**, 한국수자원학회, Vol. 50, No. 2, pp. 35-40.
- 한국수자원공사 (2014). 댐 운영 실무편람. 한국수자원공사 물관리센터.
- Abramowitz, M., and Stegun, I.A. (1964). Handbook of mathematical function, *U.S. Dept. of Commerce, National Bureau of Standards, Applied Mathematics Series*, No. 55.
- Bae, Y., and Choi, J. (2015) “A study on asymmetric relationship between financial and real sector by copula approach.” *Journal of Korean National Economy*, Vol. 33, No. 2, pp.29-59.
- Fisher, R. A. (1922). “On the mathematical foundation of theoretical statics.” *Philosophical transactions of the royal society of London, Series A*, Vol. 222, pp. 309-368.

- Humphrey G. B., Gibbs M. S., Dandy G.C., Maire H.R. (2016). “A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network.” *Journal of Hydrology*, Vol. 540, pp. 623–640.
- Joe H. (1997). *Multivariate models and dependence concept*, Chapman & Hall, London.
- Madadgar, S., and Moradkhani, H. (2013). “A bayesian framework for probabilistic seasonal drought forecasting.” *Journal of hydrometeorology*, Vol. 14, No. 6, pp. 1685–1705.
- Madadgar, S., and Moradkhani, H. (2014). “Spatio-temporal drought forecasting within Bayesian networks.” *Journal of Hydrology*, Vol. 512, pp. 134–146.
- Mckee, T. B., Doesken, N. J., and Kleist, J. (1993). *The relationship of drought frequency and duration of time scales*. 8th Conference on Applied Climatology, Jan. Anaheim, CA.
- Nelsen, R. B. (2006). *An introduction to copula*, Springer, New York.
- Schewizer, B., and Sklar, A. (1983). *Probabilistic metric spaces*, North-Holland, New York.
- Shin, J. Y., Ajmal, M., Yoo, J. Y., Kim, T. W. (2016). “A Bayesian network-based probabilistic framework for drought forecasting and outlook.” *Hindawi Publishing Corporation, Advances in Meteorology*, Vol. 2016, Article ID 9472605,
- Son, K.H., Bae, D.H., Chung, J.S. (2011). “Drought analysis and assessment by using land surface model on South

Korea.” *Journal of Korea Water Resources Association*, Vol. 44, No. 8, pp. 667–681.

Stedinger, J. R., Sule, B. F., and Loucks, D. P. (1984). “Stochastic dynamic programming models for reservoir operation optimization.” *Water Resources Research*, Vol. 20, No. 11, pp. 1499–1505.

Zezula, I. (2009). “On multivariate Gaussian copulas.” *Journal of statistical planning and inference*, Vol. 139, No. 11, pp. 3942–3946.





감사의 글

전역 후 여름방학 때 대학원에 들어가고자 이상호 교수님과 면담을 하고 연구실에 들어와 공부를 시작하고 어느덧 3년 반이 지나 학위논문을 완성하고 졸업을 하게 되었습니다.

연구실에 들어온 후 짧지도 길지도 않은 기간 동안 저에게 많은 지식을 쌓고 다양한 경험을 하게 해주신 저의 지도교수님인 이상호 교수님께 매우 감사드립니다. 논문 심사 주심을 맡아주시고, 대학원 생활 동안 많은 도움을 주신 김수한 교수님과 바쁜 와중에도 논문 심사를 흔쾌히 허락해 주신 이종섭 교수님께 감사드립니다. 대학원 생활을 하는데 많은 도움을 주신 김상단 교수님께도 감사드립니다.

학위논문을 작성하는데 많은 조언을 해주신 박종표 선배님, 오고가며 많은 도움을 주신 강태욱 선배님, 연구실 생활 중 많은 도움과 조언을 해주시고 좋은 추억 만들어주신 정태훈 형님, 진영규 형님께 감사드립니다. 연구실 막내를 벗어나게 해준 공부 잘하는 김성민, 몇 달 전 연구실에 들어와 공부를 시작하고 나와 같은 석사의 길을 걸어갈 정택문에게도 고맙다는 말을 하고 싶고 힘내라는 말을 하고 싶습니다. 항상 밝으신 멋쟁이 배재영 형님, 함께 대학원에 입학했지만 먼저 취직해 일하고 있는 김한아, 석사과정이자 조교로서 많은 도움을 주신 이동기 형님에게도 감사드립니다. 박정규 형님, 김재현 형님 농구할 때 재미있었습니다. 갓웨이를 마음에 들어했던 박지혜 누나 취직 축하드립니다. 동갑내기 친구지만 잠수병을 고쳤으면 하는 박영웅, 동갑내기 친구지만 박사과정을 밟고 있는 멋진 친구 이옥정에게도 고맙습니다.

어릴 때부터 함께 지내온 “낭만에 대하여” 계모임 친구들인 김도훈, 김상구, 김대진, 석상혁, 송신욱, 박주홍 좋은 친구로 남아줘서 고맙다. 김도훈은 열심히 돈 모아서 장사 잘 하고, 상구는 좋은 회사에 다시 취업해서 열심히 일하고, 대진은 더우나 추우나 나라 잘 지켜줘서 고맙다. 석상은 항상 운전 조심하고 피곤한데 불철주야 고생한다. 챙이는 얼른 좋은데 취직하고 많이 먹고 살 좀 썩라. 박주홍은 낭만에 대하여는 탈퇴했지만 가장 오래된 친구로서 항상 고맙다. 우리 친구들 모두 앞으로도 좋은 추억 만들면서 오래오래 지내면서 항상 건강하고 하는 일 다 잘 되었으면 좋겠다.

마지막으로 대학원에 진학하고 공부를 하는데 아낌없이 지원을 해주신 사랑하는 저의 부모님과 하나뿐인 동생과 모든 가족에게 진심으로 감사드립니다. 사랑합니다.

2018. 01.
김 광 훈