



## Determination of optimal order for the full-logged I-D-F polynomial equation and significance test of regression coefficients

Park, Jin Hee<sup>a,b</sup> · Lee, Jae Joon<sup>c\*</sup>

<sup>a</sup>ABD, Department of Civil Engineering, Kumoh National Institute of Technology, Gumi, Korea

<sup>b</sup>Subsection Chief, Sewage Section, Gumi City Hall, Gumi, Korea

<sup>c</sup>Professor, Department of Civil Engineering, Kumoh National Institute of Technology, Gumi, Korea

Paper number: 22-078

Received: 15 September 2022; Revised: 26 September 2022; Accepted: 27 September 2022

### Abstract

In this study, to determine the optimal order of the full-logged I-D-F polynomial equation, which is mainly used to calculate the probable rainfall over a temporal rainfall duration, the probable rainfall was calculated and the regression coefficients of the full-logged I-D-F polynomial equation was estimated. The optimal variable of the polynomial equation for each station was selected using a stepwise selection method, and statistical significance tests were performed through ANOVA. Using these results, the statistically appropriately calculated rainfall intensity equation for each station was presented. As a result of analyzing the variable selection outputs of the full-logged I-D-F polynomial equation at 9 stations in Gyeongbuk, the 1<sup>st</sup> to 3<sup>rd</sup> order equations at 6 stations and the incomplete 3<sup>rd</sup> order at 1 station were determined as the optimal equations. Since the 1<sup>st</sup> order equation is similar to the Sherman type equation and the 2<sup>nd</sup> order one is similar to the general type equation, it was presented as a unified form of rainfall intensity equation for convenience of use by increasing the number of independent variables. Therefore, it is judged that there is no statistical problem in considering only the 3<sup>rd</sup> order polynomial regression equation for the full-logged I-D-F.

**Keywords:** Full-logged I-D-F polynomial equation, Optimal order, Stepwise selection, Statistical significance test

## 전대수 다항식형 확률강우강도식의 최적차수 결정 및 회귀계수에 대한 유의성 검정

박진희<sup>a,b</sup> · 이재준<sup>c\*</sup>

<sup>a</sup>금오공과대학교 토목공학과 박사수료, <sup>b</sup>구미시청 하수과 하수정비계장, <sup>c</sup>금오공과대학교 토목공학과 교수

### 요 지

본 연구에서는 임의지속기간의 확률강우량 산정을 위해 실무에서 주로 사용되고 있는 전대수 다항식형 확률강우강도식의 최적차수 결정을 위하여 경상북도 내 9개 지점을 대상으로 확률강우량을 산정하고 전대수 다항식형 강우강도식의 회귀계수를 추정하였다. 추정된 지점별 다항식을 대상으로 단계선택법을 이용하여 각 지점별 다항식의 최적변수를 선정하고 선정된 변수들의 통계적 유의성을 검토하기 위하여 분산분석을 통한 유의성 검정을 실시하였으며, 이들 결과를 이용하여 각 지점별 통계적으로 적절하게 산정된 강우강도식을 제시하였다. 경북 9개 지점의 전대수 다항식형 강우강도식의 변수선정 결과는 6개 지점에서 1~3차식이 최적식으로 나타났고 1개 지점이 불완전 3차식이 최적식으로 나타났다. 그 중 1차는 Sherman 식, 2차는 General 식의 형태와 유사하므로 독립변수의 수를 증가시켜 적합도를 높이고 사용 편의를 위해 통일된 형태의 강우강도식으로 제시한다면 전대수 다항식형 강우강도식은 3차 회귀식까지만 고려하여도 통계학적으로 문제가 없는 것으로 판단된다.

**핵심용어:** 전대수 다항식형 확률강우강도식, 최적차수, 단계 선택법, 유의성 검정

\*Corresponding Author. Tel: +82-54-478-7616  
E-mail: jhb365@kumoh.ac.kr (Lee, Jae Joon)

## 1. 서론

최근 국내에서 기후변화로 인하여 국지성 호우가 빈번하게 나타나고 있으며 우수배제시설의 설계빈도를 초과하는 강우사상의 발생으로 많은 지역에서 내수침수로 인한 재산 및 인명피해가 발생하고 있다. 이에 따라 내수침수 피해를 저감하고 예방하기 위한 방안들이 계획되고 있고 수자원 계획 및 수공구조물 설계를 위한 최근 강우사상을 반영한 계획홍수량 산정이 중요한 요소로 작용하고 있다. 국내에서는 홍수량 관측자료가 부족하므로 확률강우량을 산정하고 확률강우강도식을 유도하여 설계홍수량을 산정하는 방법을 사용하고 있다. 중·대규모 유역의 경우 임계지속시간을 1시간 단위로 적용하므로 확률강우량을 바로 사용할 수 있으나 소규모 유역의 경우 임계지속시간을 10분 단위로 짧게 적용하므로 강우강도식을 유도하여 확률강우량을 산정하므로 정확한 강우강도식의 유도가 요구된다.

국내에서는 Lee (1967), Lee and Byeon (1969)에 의해 확률강우강도식의 연구가 수행되기 시작하여 Lee (1980), Lee *et al.* (1981) 등을 통하여 연구가 발전되어 왔다. Lee *et al.* (1993)과 Heo *et al.* (1999)는 국내 실정에 맞는 강우강도식을 제안하였고 Kim *et al.* (2007)은 부산지점의 강우강도식을 유도하고 유전자 알고리즘을 이용한 강우강도식의 매개변수 산정방법을 제시하였다. NEMA (NEMA, 2010)에서는 전국 68개 강우관측소의 평균 확률강우량 및 확률강우강도식을 유도하여 실무에 적용가능하도록 제시하였다. MLTMA (MLTMA, 2011)에서는 확률강우강도식으로 전기간에서 정확도가 높은 전대수 다항식형 강우강도식을 제시하였고 MLTMA (MLTMA, 2012)에서 3변수인 General 형과 6차 다항식인 전대수 다항식형 등 두 가지 형태를 채택하여 사용하도록 하고 있으나 실무에서는 대부분 전대수 다항식형의 강우강도식을 일률적으로 사용하고 있다. 이후 Jeong *et al.* (2017)이 진주를 대상으로 General 형과 전대수 다항식형 강우강도식을 이용한 지역의 강우특성 분석 결과 5차 전대수 다항식의 적합도가 가장 높은 결과를 도출해 내었으나 그 외에 전대수 다항식형 강우강도식의 최적차수 결정에 관한 연구는 이루어지지 않고 있다.

현재 실무에서는 회귀식의 최적차수 결정에 대한 객관적 기준이 부재한 상황이므로 전대수 다항식 형의 강우강도식에 대한 유의성 검정이 필요한 실정이다. 따라서 본 연구에서는 실무에서 일반적으로 임의시간 확률강우량을 산정하기 위해 사용하고 있는 전대수 다항식형 강우강도식에 대해 확률강우량의 시간분포 회귀식의 유의성 검정을 통한 최적차수 결정의 연구를 진행한 Park and Lee (2022)의 연구와 동일하게 회귀

식을 유도하고 전대수 다항식형 확률강우강도식의 최적차수 결정 및 회귀계수에 대한 유의성 검정방법을 적용하여 통계적으로 적절한 강우강도식을 산정하였다.

## 2. 강우강도식의 유형 및 문제점

### 2.1 강우강도식의 유형

강우강도식은 산정된 확률강우량에 포함되지 않은 임의지속기간의 확률강우량을 산정하기 위한 목적으로 유도하고 있다. 현재까지 다양한 연구들로 인해 국내 실정에 맞는 여러 강우강도식이 유도되었으나 본 연구에서는 기본 형태 및 현재 가장 많이 사용되고 있는 강우강도식을 대상으로 연구를 진행하였다.

강우강도식의 기본 형태는 Eq. (1)과 같이 지속기간에 따라 강우강도가 감소하는 멱함수 형태를 가지고 있고 대수 변형시 선형에 가까운 형태를 띠게 된다. 이러한 형태의 강우강도식은 최근까지도 널리 사용되어 왔으나 현재에는 곡선의 적합도를 높이기 위해 전대수 다항식 형태로 사용하고 있다.

$$I(t) = \frac{a}{t^n + b} \quad (1)$$

일반적으로 가장 널리 알려진 Talbot 형, Sherman 형, Japanese 형도 이 General 형을 근간으로 하고 있다. 일반적으로  $n=1$ 인 경우는 Talbot 형,  $b=0$ 인 경우는 Sherman 형,  $n=0.5$ 인 경우는 Japanese 형이 된다.

전대수 다항식 형태의 강우강도식은 지속기간에 따른 확률강우강도를 대수화하여 고차 다항식에 적합시키는 방법이다. 지속기간에 따라 확률강우강도 곡선의 곡률이 차이가 큰 경우 기존 Sherman 형태의 확률강우강도식에 비해 적합의 정확도가 높고 비선형 적합도가 높기 때문에 확률 강우강도식을 지속기간별로 단기간 및 장기간으로 구분하여 작성할 필요가 없을 정도로 전 지속기간에 걸쳐 하나의 곡선식으로 표현이 가능하다는 장점이 있으며, 국토교통부에서는 Eq. (2)와 같은 형태의 식을 제시하였다(MLTMA, 2011).

$$\ln(I) = a + b \ln(t_h) + c(\ln(t_h))^2 + d(\ln(t_h))^3 + e(\ln(t_h))^4 + f(\ln(t_h))^5 + g(\ln(t_h))^6 \quad (2)$$

그러나 지금까지 사용해온 강우강도식의 지속기간은 분 단위를 사용하고 있으므로 1시간 이내의 짧은 지속기간에 대

해 강우강도를 용이하게 취할 수 있도록 본 연구에서는 Eq. (3)과 같은 형태의 6차 전대수 다항식형을 사용하기로 한다.

$$\ln(I) = a + b \ln(t_m) + c(\ln(t_m))^2 + d(\ln(t_m))^3 + e(\ln(t_m))^4 + f(\ln(t_m))^5 + g(\ln(t_m))^6 \quad (3)$$

여기서, I(t)는 강우지속기간에 따른 강우강도(mm/hr),  $t_h$ 는 강우지속기간(hr),  $t_m$ 은 강우지속기간(min), a, b, c, d, e, f, g 등은 회귀상수이다.

### 2.2 각종 강우강도식의 문제점

확률강우량도 개선 및 연구에서 전대수 다항식형 강우강도식이 제시된 이후, 강우강도식 유도 시 General 형과 전대수 다항식형 두가지 형태를 사용할 것을 제안하고 있다(MLTMA, 2012).

일반적으로 사용되고 있는 Talbot, Sherman, Japanese 형과 같은 강우강도식은 General 형 강우강도식으로 나타낼 수 있다. 이 General 형 강우강도식은 a, b, n 세 개의 매개변수를 사용하여 강우강도식을 산정하기 때문에 다른 강우강도식들에 비해 적은 수의 매개변수를 사용하므로 매개변수 추정이 용이하다는 장점이 있다.

그러나 General 형 강우강도식은 재현기간에 대해서는 고려할 수 없으므로 동일한 지속기간을 적용하더라도 짧은 재현기간에 해당하는 강우강도가 긴 재현기간에 해당하는 강우보다 커지는 역전현상이 다른 강우강도식들에 비해 많이 나타나며, 정확도 또한 다른 강우강도식들에 비해 낮게 나타

나는 단점이 있는 것으로 연구되었다(Kim *et al.*, 2007).

한편 전대수 다항식형 강우강도식은 3변수인 General 형 강우강도식에 비해 결정계수가 높고 단·장기간을 구분할 필요가 없으며 회귀측면에서 전기간이 모두 우수하므로 현재 실무에서는 6차 전대수 다항식형 강우강도식이 주로 사용되고 있다. 그러나 전대수 다항식형은 강우강도식은 확률강우량 자료가 없는 10분~60분 구간에서 움푹 패이는 형태가 자주 발생하고 있으며(Fig. 1), 이를 보완하기 위하여 MOIS (MOIS, 2021)에서는 전대수 다항식형을 4~5차식으로 조정하는 것을 제안하고 있으나 통계적 분석에 따른 각 차수에 대한 유의성 검정을 실시하지 않고 차수를 조정하도록 제시하는 등 전대수 다항식형의 최적차수를 결정하지 못하는 문제점이 있다.

그리고 기왕에 사용하던 다른 강우강도식과는 달리 강우지속기간을 시간(hr) 단위의 값을 사용함으로써 1시간 미만의 강우지속기간에 대해서는 대수변환을 실시하여 사용하여야 하는 번거로움과 통일된 독립변수 사용이 이루어지지 못함으로써 불편을 초래하고 있기도 하다. 또한, 6차 전대수 다항식형의 회귀식에서 확률강우량에 자연로그(ln)를 취했을 때 지속기간을 시간(hr)으로 할 경우, 지속기간 1시간에서의 값이 0으로 산정되어 독립변수 역할이 제한되는 문제점이 있다 (Tables 1 and 2).

따라서 본 연구에서는 전대수 다항식형 강우강도식의 최적 차수를 결정하고 회귀계수에 대한 유의성 검정을 실시하기 위하여 전대수 다항식형 회귀식에서의 강우지속기간을 분(min)으로 사용한 확률강우강도식을 산정하여 연구를 진행하였다.

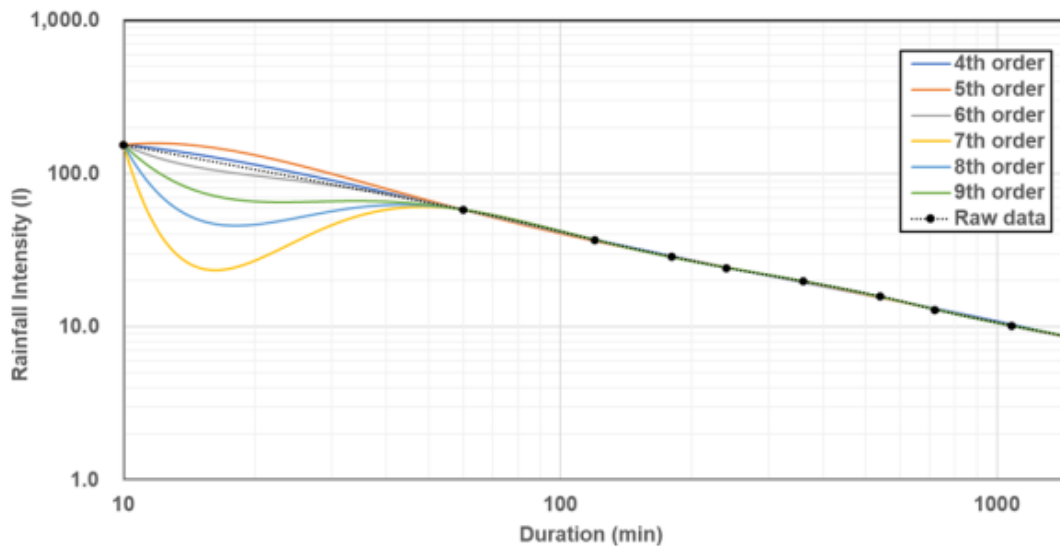


Fig. 1. The full-logged I-D-F polynomial equation curve (Gumi)

**Table 1.** Dependent and independent variables of probable rainfall intensity formula for Gumi (frequency 50 years, Duration (hr))

Probable rainfall intensity (mm/hr)	Duration (hr)	Dependent variable	Independent variables					
		ln(I)	ln(t)	(ln(t)) <sup>2</sup>	(ln(t)) <sup>3</sup>	(ln(t)) <sup>4</sup>	(ln(t)) <sup>5</sup>	(ln(t)) <sup>6</sup>
166.8	0.167	5.11679549	-1.791759469	3.210401996	-5.752268176	10.30668097	-18.46709323	33.08858916
62.60	1	4.136765278	0	0	0	0	0	0
39.45	2	3.675034047	0.693147181	0.480453014	0.333024652	0.230835099	0.160002698	0.110905419
30.53	3	3.418709807	1.098612289	1.206948961	1.32596896	1.456725794	1.600376859	1.758193683
⋮								

**Table 2.** Dependent and independent variables of probable rainfall intensity formula for Gumi (frequency 50 years, Duration (min))

Probable rainfall intensity (mm/hr)	Duration (min)	Dependent variable	Independent variables					
		ln(I)	ln(t)	(ln(t)) <sup>2</sup>	(ln(t)) <sup>3</sup>	(ln(t)) <sup>4</sup>	(ln(t)) <sup>5</sup>	(ln(t)) <sup>6</sup>
166.8	10	5.11679549	2.302585093	5.30189811	12.20807155	28.11012357	64.7259515	149.0370111
62.60	60	4.136765278	4.094344562	16.76365739	68.63618949	281.0202092	1150.593566	4710.926508
39.45	120	3.675034047	4.787491743	22.92007719	109.7296803	525.3299383	2515.012742	12040.60273
30.53	180	3.418818983	5.192956851	26.96680086	140.0374332	727.2083484	3776.361575	19610.48271
⋮								

### 3. 경북지역 전대수 다항식형 강우강도식 분석

#### 3.1 연구대상 강우관측소(경북)의 확률강우량 및 확률 강우강도식

본 연구에서는 강우강도식의 최적차수 결정 및 유의성 검정을 위해 2021년까지의 기상청 과거 시간별 기상관측자료를 이용하여 경북지역의 9개 기상청 관측소(울진, 안동, 포항, 대구, 봉화, 영주, 문경, 의성, 구미)의 확률강우량을 강우분석 프로그램인 FARD2020(Frequency Analysis of Rainfall Data Program 2020; ME, 2019)을 이용하여 산정하였다. 확률강우량 산정을 위한 확률분포함수의 모수 추정 방법은 확률가중모멘트법(PWM), 확률분포형은 Gumbel 분포를 채택하였다. 본 연구에서는 지방하천 및 홍수방어(조절)용 저수지의 최소 설계빈도인 50년의 확률강우량을 이용하여 연구를 진행하였다(Table 3).

앞서 산정된 경북지역의 울진 등 9개 지점의 빈도 50년 확률강우량을 이용하여, Eq. (3)의 형태를 가진 6차 전대수 다항식형 강우강도식을 유도하였으며, 각 지점별 강우강도식 회귀계수는 Table 4와 같다.

#### 3.2 변수선택법에 의한 최적차수 결정

앞 절에서 산정된 각 지점별 전대수 다항식의 회귀계수를

이용하여 변수선택법에 의한 최적차수를 결정하였다. 변수선택법에는 다양한 방법이 있으나 일반적으로 전진선택법(Forward Selection)과 후진제거법(Backward Elimination)을 모두 이용하는 단계선택법(Stepwise Selection)을 이용하여 선정된 모형이 적합도가 가장 높은 것으로 알려져 있다(Park and Lee, 2022). 본 연구에서는 변수가 모두 포함된 상태에서 시작하여 기여도가 낮은 변수를 삭제하는 방법으로 단계선택법을 이용하여 변수선정을 통한 최적차수를 결정하였고 또한 결정된 모형의 예측력을 판단하기 위한 근거로 PRESS 통계량(Predicted Residual Error Sum of Squares statistic)을 이용하였다.

$$PRESS = \sum_{i=1}^n e_{i(i)}^2 = \sum_{i=1}^n (Y_i - \widehat{Y}_{i(i)})^2 \tag{4}$$

여기서,  $\widehat{Y}_{i(i)}$ 는  $i$ 번째 자료를 제외하고 잔차의 제곱합(SSE: Error sum of squares)이 최소가 되는 모형으로부터  $i$ 번째 값을 추정하는 것이다.

PRESS 통계량은 여러 모형들의 예측력을 비교할 수 있는 유용한 통계량으로 PRESS 기준에 의해서 모형을 선택할 때에는 PRESS값이 작을수록 예측력이 높다고 판단할 수 있다. 또한 예측력의 비교를 위해 PRESS 잔차를 이용한 통계량  $R^2_{PRESS}$ 이 있는데,  $R^2_{PRESS}$ 가 높을수록 예측력이 높다고 할 수

**Table 3.** Probable rainfall of 9 stations in Gyeongbuk area (frequency 50 years)

Stations (No.)	Durations (min)								
	10	60	120	180	240	300	360	720	1,440
Uijin (130)	20.7	61.1	80.9	98.0	114.7	128.7	141.0	203.9	272.3
Andong (136)	25.4	60.1	77.6	86.6	94.7	100.5	106.2	147.2	191.5
Pohang (138)	25.8	74.1	97.6	122.4	139.2	155.1	169.1	242.2	321.0
Daegu (176)	23.8	69.9	90.5	105.7	117.0	127.0	137.0	186.7	244.7
Bonghwa (271)	27.3	75.0	98.8	127.5	143.8	166.4	185.9	250.5	319.9
Yeongju (272)	26.5	71.5	92.9	113.7	129.0	142.1	153.2	203.2	274.9
MunGyeong (273)	24.9	62.4	84.0	95.8	108.4	115.6	123.9	165.2	217.1
Uiseong (278)	28.9	62.3	79.6	93.8	104.4	113.7	118.8	155.3	198.5
Gumi (279)	27.8	62.6	78.9	91.6	103.1	115.5	127.2	168.4	225.0

**Table 4.** Estimation of regression coefficients of polynomial regression equation for the full-logged I-D-F (6th order equation)

Coefficients Stations (No.)	a	b	c	d	e	f	g	Determination coefficient (R <sup>2</sup> )
	Uijin (130)	12.035975	-11.918539	7.873583	-2.607961	0.449885	-0.03899	
Andong (136)	53.478554	-61.204294	30.70155	-7.882876	1.087921	-0.076796	0.00217	99.99%
Pohang (138)	-10.454046	19.035789	-8.981816	2.115819	-0.272363	0.018347	-0.00051	99.99%
Daegu (176)	33.249469	-38.187239	20.713252	-5.787349	0.873621	-0.068029	0.002146	99.99%
Bonghwa (271)	-36.345993	51.798821	-25.234327	6.204589	-0.826264	0.056862	-0.001589	99.97%
Yeongju (272)	-88.716248	124.510849	-65.769694	17.877007	-2.6635	0.207034	-0.006578	99.98%
MunGyeong (273)	62.402645	-75.494089	39.859094	-10.89812	1.627179	-0.126486	0.004015	99.99%
Uiseong (278)	-28.266408	44.929689	-23.773196	6.408424	-0.946127	0.072816	-0.002289	99.99%
Gumi (279)	35.58122	-43.482677	25.234191	-7.598459	1.240221	-0.104473	0.003562	99.98%

있다(Allen, 1971).

$$R^2_{PRESS} = 1 - \frac{PRESS}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (5)$$

여기서,  $\sum_{i=1}^n (Y_i - \bar{Y})^2$ 는 전체제곱합(Total sum of squares, SST)으로서, 회귀제곱합(SSR)과 잔차제곱합(SSE)으로 분해된다.

경북지역 9개 지점의 전대수 다항식의 최적차수 결정을 위하여 통계 분석 프로그램인 R 프로그래밍을 이용하였으며 앞서 설명한 바와 같이 단계선택법을 적용하였다. 변수 선정을 위한 유의수준은 일반적 유의수준인 0.15를 적용하였으며, PRESS 통계량을 이용하여 지점별 최적의 모형을 선정하였다.

Table 5는 경북 9개 지점의 3차부터 6차 회귀식을 대상으로 변수선정 결과를 나타낸 것이다.

Table 5에서 보면, 울진, 안동, 포항, 영주, 문경, 의성 및 구

미지점은 3차식을 적용하였을 때 1~3차식에서 PRESS 통계량이 가장 낮은 값을 가져 1~3차식의 적합도가 가장 높은 것으로 분석되었다. 반면에 대구와 봉화지점은 불완전 5~6차식일 때 PRESS 통계량이 가장 낮아 적합도가 가장 높은 것으로 분석되었으나, 봉화지점의 경우 3차식의 경우에도 4차, 5차식에 비해 적합도가 높은 것으로 나타나 전반적으로 모든 지역에서 3차 회귀식의 적합도가 높게 나타났다.

또한 선택된 회귀계수들을 살펴보면 안동, 구미 등 2개 지점은 1차식이 최적식으로 나타났고, 울진, 포항은 완전한 2차식, 문경과 의성 2개 지점은 완전한 3차식이 최적식으로 결정되었다. 영주는 3차식의 최대 차수가 포함된 불완전 3차식이 최적으로 선택되었다. 대구와 봉화 2개 지점은 각각 불완전 5차 및 6차식이 최적식으로 결정되었다.

경북 9개 지점의 전대수 다항식형 강우강도식의 변수선정 결과를 분석한 결과 9개 지점 중 6개 지점이 1~3차식이 최적식으로 나타났고 1개 지점이 불완전 3차식이 최적식으로 나타났다. 그 중 1차는 Sherman 식, 2차는 General 식의 형태와 유

Table 5. Results of statistic by stepwise selection method (frequency 50 years)

Stations	Max. Order	6th	5th	4th	3rd
Uljin (130)	Selected Parameter	a, b, c	a, b, c	a, b, c	a, b, c
	Adj R <sup>2</sup>	0.9986	0.9986	0.9986	0.9986
	Standard Deviation	0.0213406	0.0213406	0.0213406	0.0213406
	PRESS	0.0308745	0.0308745	0.0308745	<b>0.0308745</b>
	R <sup>2</sup> <sub>PRESS</sub>	0.9960	0.9960	0.9960	<b>0.9960</b>
Andong (136)	Selected Parameter	a, b	a, b	a, b	a, b
	Adj R <sup>2</sup>	0.9977	0.9977	0.9977	0.9977
	Standard Deviation	0.0332928	0.0332928	0.0332928	0.0332928
	PRESS	0.0428330	0.0428330	0.0428330	<b>0.0428330</b>
	R <sup>2</sup> <sub>PRESS</sub>	0.9963	0.9963	0.9963	<b>0.9963</b>
Pohang (138)	Selected Parameter	a,b,c	a,b,c	a,b,c	a,b,c
	Adj R <sup>2</sup>	0.9991	0.9991	0.9991	0.9991
	Standard Deviation	0.0175795	0.0175795	0.0175795	0.0175795
	PRESS	0.0234781	0.0234781	0.0234781	<b>0.0234781</b>
	R <sup>2</sup> <sub>PRESS</sub>	0.9971	0.9971	0.9971	<b>0.9971</b>
Daegu (176)	Selected Parameter	a, b, c, d, f	a, b, c, d, f	a, b, c, d, e	a, b, c, d
	Adj R <sup>2</sup>	0.9999	0.9999	0.9999	0.9999
	Standard Deviation	0.0066418	0.0066418	0.0066822	0.0162598
	PRESS	0.0218695	<b>0.0218695</b>	0.0764592	0.583226
	R <sup>2</sup> <sub>PRESS</sub>	0.9978	<b>0.9978</b>	0.9923	0.9413
Bonghwa (271)	Selected Parameter	a, c, f, g	a, b, c, d, e, f	a, b, c, d, e	a, b, c, d
	Adj R <sup>2</sup>	0.9992	0.9997	0.9993	0.9985
	Standard Deviation	0.0171781	0.0107769	0.0160600	0.0225655
	PRESS	<b>0.0169998</b>	5.35882	7.14076	1.02159
	R <sup>2</sup> <sub>PRESS</sub>	<b>0.9980</b>	0.3607	0.1482	0.8781
Yeongju (272)	Selected Parameter	a, c, d	a, c, d	a, c, d	a, c, d
	Adj R <sup>2</sup>	0.9995	0.9995	0.9995	0.9995
	Standard Deviation	0.0135848	0.0135848	0.0135848	0.0135848
	PRESS	0.0053238	0.0053238	0.0053238	<b>0.0053238</b>
	R <sup>2</sup> <sub>PRESS</sub>	0.9994	0.9994	0.9994	<b>0.9994</b>
MunGyeong (273)	Selected Parameter	a, b, c, d	a, b, c, d	a, b, c, d	a, b, c, d
	Adj R <sup>2</sup>	0.9998	0.9998	0.9998	0.9998
	Standard Deviation	0.0094979	0.0094979	0.0094979	0.0094979
	PRESS	0.0035926	0.0035926	0.0035926	<b>0.0035926</b>
	R <sup>2</sup> <sub>PRESS</sub>	0.9997	0.9997	0.9997	<b>0.9997</b>
Uiseong (278)	Selected Parameter	a, b, c, d	a, b, c, d	a, b, c, d	a, b, c, d
	Adj R <sup>2</sup>	0.9999	0.9999	0.9999	0.9999
	Standard Deviation	0.0058515	0.0058515	0.0058515	0.0058515
	PRESS	0.0072925	0.0072925	0.0072925	<b>0.0072925</b>
	R <sup>2</sup> <sub>PRESS</sub>	0.9994	0.9994	0.9994	<b>0.9994</b>
Gumi (279)	Selected Parameter	a, b	a, b	a, b	a, b
	Adj R <sup>2</sup>	0.9992	0.9992	0.9992	0.9992
	Standard Deviation	0.01936	0.01936	0.01936	0.01936
	PRESS	0.0111272	0.0111272	0.0111272	<b>0.0111272</b>
	R <sup>2</sup> <sub>PRESS</sub>	0.9990	0.9990	0.9990	<b>0.9990</b>

사하므로 독립변수의 수를 증가시켜 적합도를 높이고 사용 편의를 위해 통일된 형태의 강우강도식으로 제시한다면 전대수 다항식형 강우강도식은 3차 회귀식까지만 고려하여도 통계학적으로 문제가 없는 것으로 판단된다.

#### 4. 전대수 다항식형 강우강도식의 결정

##### 4.1 전대수 다항식형 강우강도식의 유의성 검정

3장에서 단계선택법으로 산정된 전대수 다항식형 강우강도식에 대한 유의성 검정을 수행하였다. 경북 9개 지점을 대상으로 3차부터 6차까지의 전대수 다항식을 대상으로 분산분석(ANalysis Of VAriance, ANOVA)을 통한 회귀계수들의 통계적 유의성에 대한 검정을 실시하였다. 유의성 검정 과정은

3장에서 수행한 단계선택법의 과정과 동일하게 예측변수가 없는 상황에서 유의수준  $\alpha=0.15$ 를 적용하여 변수들을 추가하고 제거하는 과정을 반복하여 변수를 선정하여 회귀계수에 대한 가설검정을 수행하였다. Tables 6~14는 경북 9개 지점의 각 차수별 선정된 회귀계수들의 분산분석 결과로서 상수항을 제외한 선정된 변수들의 기여도 및 F-통계량, p-값 등을 나타내고 있다. 분산분석 결과 봉화와 영주 지점을 제외한 7개 지점에서 1차항의 회귀계수인 변수 b의 기여도가 99% 이상인 것으로 분석되었고 나머지는 1.5% 이하로 저차항의 회귀계수들의 기여도가 높게 나타났으며, 선정된 변수들이 포함된 다항회귀식은 통계적으로 적합도가 높은 것으로 나타났다.

Table 6. ANOVA of 6th-order regression equation (Uljin)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
b	1	7.65888	99.67%	0.09565	0.09565	210.03	0.000000
c	1	0.01496	0.19%	0.01496	0.01496	32.84	9.8E-13
Residual	22	0.01002	0.13%	0.01002	0.00046		9.16E-6
Total	24	7.68386	100.00%				

Table 7. ANOVA of 6th-order regression equation (Andong)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
b	1	11.4966	99.78%	11.4966	11.4966	10372.18	0.000000
Residual	23	0.0255	0.22%	0.0255	0.0011		
Total	24	11.5221	100.00%				

Table 8. ANOVA of 6th-order regression equation (Pohang)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
b	1	7.99488	99.72%	0.09916	0.09916	320.86	1.32E-14
c	1	0.01589	0.20%	0.01589	0.01589	51.42	3.45E-07
Residual	22	0.0068	0.08%	0.0068	0.00031		
Total	24	8.01757	100.00%				

Table 9. ANOVA of 3th-order regression equation (Daegu)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
b	1	9.87588	99.43%	0.00594	0.00594	134.76	2.44E-10
c	1	0.04024	0.41%	0.00714	0.00714	161.75	4.84E-11
d	1	0.01064	0.11%	0.00601	0.00601	136.29	2.21E-10
f	1	0.00467	0.05%	0.00467	0.00467	105.86	1.96E-09
Residual	20	0.00088	0.01%	0.00088	0.00004		
Total	24	9.93231	100.00%				

**Table 10.** ANOVA of 3th-order regression equation (Bonghwa)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
c	1	8.29757	98.98%	0.28154	0.28154	954.08	0.000000
f	1	0.04378	0.52%	0.04019	0.04019	136.20	1.21E-10
g	1	0.03516	0.42%	0.03516	0.03516	119.16	4.09E-10
Residual	21	0.0062	0.07%	0.0062	0.0003		
Total	24	8.38272	100.00%				

**Table 11.** ANOVA of 6th-order regression equation (Yeongju)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
c	1	9.33039	98.62%	0.53163	0.53163	2880.74	0.000000
d	1	0.12657	1.34%	0.12657	0.12657	685.82	0.000000
Residual	22	0.00406	0.04%	0.00406	0.00018		
Total	24	9.46102	100.00%				

**Table 12.** ANOVA of 6th-order regression equation (Mungyeong)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
b	1	10.88330	99.740%	0.0004	0.00036	4.00	0.058633
c	1	0.01950	0.180%	0.009	0.00903	100.14	1.91E-09
d	1	0.00720	0.070%	0.0072	0.00715	79.28	1.42E-08
Residual	21	0.00190	0.020%	0.0019	0.00009		
Total	24	10.91180	100.000%				

**Table 13.** ANOVA of 6th-order regression equation (Uiseong)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
b	1	12.22890	99.940%	0.0030	0.0030	88.44	5.62E-09
c	1	0.00570	0.050%	0.0010	0.0010	29.08	2.38E-05
d	1	0.00070	0.010%	0.0007	0.0007	19.51	2.40E-04
Residual	21	0.00070	0.010%	0.0007	0.0000		
Total	24	12.23600	100.000%				

**Table 14.** ANOVA of 6th-order regression equation (Gumi)

Parameter	DF	Seq SS	Contribution	Adj SS	Adj MS	F-statistic	p-value
b	1	10.7622	99.92%	10.7622	10.7622	28695.36	0.000000
Residual	23	0.0086	0.08%	0.0086	0.0004	-	-
Total	24	10.7708	100.00%	-	-	-	-

#### 4.2 지점별 전대수 다항식형 강우강도식의 결정

3장에서 단계선택법을 이용하여 선정된 전대수 다항식의 회귀계수들과 앞 절에서 분산분석을 통해 선정된 회귀계수들을 비교분석한 결과 봉화 지점을 제외하곤 두 개의 방법이 동일한 결과를 나타내어 해당 결과를 이용하여 지점별 전대수 다항식형 강우강도식을 결정하였다(Table 15). 결정된 강우

강도식을 살펴보면 봉화 지점은 6차식의 최대 차수가 포함된 불완전 6차 강우강도식이 최적식으로 나타났으나 단계적 선택으로 인한 변수선정 결과를 보면 6차식과 3차식의 적합도가 비교적 근사한 값을 보여주고 있어 3차식으로 선정하여도 무방할 것으로 판단된다. 대구, 영주, 문경, 의성 등 4개 지점의 강우강도식에는 3차 회귀식의 최대 차수들이 포함된 식을 보



**Table 15.** Estimation of regression coefficients of best equation for the full-logged I-D-F (stepwise selection method)

Coefficients Stations (No.)	a	b	c	d	e	f	g	Determination coefficient (R <sup>2</sup> )
Uljin (130)	5.715978	-0.352905	-0.013361	-	-	-	-	99.47%
Andong (136)	6.468954	-0.601273	-	-	-	-	-	99.77%
Pohang (138)	5.952171	-0.359312	-0.013773	-	-	-	-	99.91%
Daegu (176)	2.397779	2.638634	-0.843643	0.080131	-	-0.000269	-	99.99%
Bonghwa (271)	5.566041	-	-0.090698	-	-	0.00036	-0.000037	99.92%
Yeongju (272)	5.533901	-	-0.100551	0.005806	-	-	-	99.95%
MunGyeong (273)	5.333638	0.126301	-0.135097	0.008043	-	-	-	99.98%
Uiseong (278)	6.20658	-0.365935	-0.044855	0.002458	-	-	-	99.99%
Gumi (279)	6.470344	-0.581752	-	-	-	-	-	99.92%

여주고 있으며, 울진, 안동, 포항, 구미 지점에서는 1차와 2차 회귀식의 최대 차수들이 포함된 식으로 나타났다. 울진, 안동, 포항, 구미 등 4개 지점에서는 1차와 2차 회귀식이 통계적으로 가장 적합도가 높은 것으로 나타났으나 이 식을 그대로 사용할 경우 저차항의 특성에 따라 오차가 비교적 크게 나타나므로 다른 지점들과 동일하게 3차 회귀식을 사용하는 것이 좋을 것으로 판단된다.

## 5. 결론

본 연구에서는 실무에서 일반적으로 임의시간 확률강우량을 산정하기 위해 사용하고 있는 전대수 다항식형 강우강도식에 대해 변수선택법을 이용하여 확률강우강도식의 최적차수를 결정하고 결정된 회귀식의 회귀계수에 대한 유의성 검정을 실시하여 간결하고 최적화된 강우강도식을 제시하였다. 최적차수 결정을 위하여 단계선택법을 이용하였으며, 분산분석을 통해 통계적 유의성을 검정하였다. 이상의 연구결과를 요약하면 다음과 같다.

- 1) 현재 실무에서는 임의시간 확률강우량을 산정하기 위하여 전대수 다항식형 강우강도식을 사용하고 있으며, 전반적으로 높은 결정계수를 나타내는 6차식을 획일적으로 사용하고 있다. 본 연구에서는 전대수 다항식형 강우강도식의 최적차수 결정을 위하여 경북 9개 지점을 대상으로 확률강우량을 산정하고 전대수 다항식형 강우강도식의 회귀계수를 추정하였다.
- 2) 추정된 지점별 다항식을 대상으로 단계선택법을 이용하여 각 지점별 다항식의 최적변수를 선정하고 선정된 변수들의 통계적 유의성을 검토하기 위하여 분산분석을 실시

하였으며, 이들 결과를 이용하여 각 지점별 간결하게 산정된 강우강도식을 제시하였다.

- 3) 경북 9개 지점의 전대수 다항식형 강우강도식의 변수선택 결과를 분석한 결과 9개 지점 중 6개 지점이 1~3차식이 최적식으로 나타났고 1개 지점이 불완전 3차식이 최적식으로 나타났다. 그 중 1차는 Sherman 식, 2차는 General 식의 형태와 유사하므로 독립변수의 수를 증가시켜 적합도를 높이고 사용 편의를 위해 통일된 형태의 강우강도식으로 제시한다면 전대수 다항식형 강우강도식은 3차 회귀식까지만 고려하여도 통계학적으로 문제가 없는 것으로 판단된다.
- 4) 통계적 유의성 검정 결과, 대구, 영주, 문경, 의성 등 4개 지점의 강우강도식에는 3차 회귀식의 최대 차수들이 포함된 식을 보여주고 있으며, 봉화 지점은 6차 회귀식의 최대 차수가 포함된 식이 가장 적합도가 큰 것으로 나타났으나 3차식도 비교적 근사한 적합도를 보여주므로 3차식을 사용하여도 무방한 것으로 판단된다. 또한 울진, 안동, 포항, 구미 지점에서는 1차와 2차 회귀식의 최대 차수들이 포함된 식으로 나타났다. 이들 지점에서는 1차와 2차 회귀식이 통계적으로 가장 적합도가 높은 것으로 나타났으나 이 식을 그대로 사용할 경우 저차항의 특성에 따라 오차가 비교적 크게 나타나므로 다른 지점들과 동일하게 3차 회귀식을 사용하는 것이 좋을 것으로 판단된다.
- 5) 본 연구의 결과는 강우빈도해석의 결과로 얻어진 확률강우량(확률강우강도) 자료로부터 얻어진 다중회귀분석 및 회귀계수의 유의성 검정을 통해 얻어진 것이므로 강우빈도해석 시 자료년수 증가와 적용 확률분포형이 달라지면 그 결과가 약간 달라질 수는 있겠으나, 전반적으로 통계적 분석을 통한 결과이므로 기상학적 요인과 지리적 요인에 따른 강우량을 기본 변량으로, 강우빈도해석 결과를 2차

변량으로 사용한 것이므로 우리나라 전 지역에 대해서도 유사한 결과가 얻어질 것으로 기대할 수 있다고 생각한다.

## 감사의 글

이 연구는 금오공과대학교 교수연구년제에 의하여 연구된 실적물입니다.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

- Allen, D.M. (1971). *The prediction sum of squares as a criterion for selecting predictor variables*. Technical Report no.23, Department of Statistics, University of Kentucky, KT, U.S.
- Heo, J.H., Kim, K.D., and Han, J.H. (1999). "Derivation of rainfall Intensity-Duration-Frequency equation based on the appropriate probability distribution." *Journal of Korea Water Resources Association*, Vol. 32, No. 3, pp. 247-254.
- Jeong, J.H., Ryu, G.H., and Kim, J.C. (2017). "A practical method for derivation of 'rainfall intensity - Duration' curve." *The Magazine of The Korean Society of Hazard Mitigation*, Vol. 17, No. 3, pp. 49-54.
- Kim, T.S., Shin, J.Y., Kim, S.Y., and Heo, J.H. (2007). "Parameter estimation of Intensity-Duration-Frequency curve using genetic algorithm (I): Comparison study of existing estimation method." *Journal of Korea Water Resources Association*, Vol. 40, No. 10, pp. 811-821.
- Lee, J.J., Son, K.I., Lee, W.H., and Lee, K.C. (1981). "A study on the analysis of time-regional distribution of precipitation frequency and rainfall intensity in Korea." *Journal of Korean Association of Hydrological Sciences, KWRA*, Vol. 14, No. 4, pp. 53-72.
- Lee, W.H. (1967). "A study on rainfall characteristics and determination of probability rainfall amount at various district in Korea." *Journal of the Korean Society of Civil Engineers, KSCE*, Vol. 15, No. 3, pp. 28-38.
- Lee, W.H. (1980). "A stochastic analysis on determination of design precipitation for planning of urbanized stream & sewerage." *Journal of the Korean Society of Civil Engineers, KSCE*, Vol. 28, No. 4, pp. 81-94.
- Lee, W.H., and Byeon, K.J. (1969). "Study of the rainfall intensity probability formula required for designing city sewerage and river plans." *Journal of the Korean Society of Civil Engineers, KSCE*, Vol. 16, No. 4, pp. 1-11.
- Lee, W.H., Park, S.D., and Choi, S.Y. (1993). "A derivation of the typical probable rainfall intensity formula in KOREA." *Journal of the Korean Society of Civil Engineers, KSCE*, Vol. 13, No. 1, pp.115-120.
- Ministry of Environment (ME) (2019). *Standard guidelines for flood estimation*.
- Ministry of Land, Transport and Maritime Affairs (MLTMA) (2011). *Study of improvement and supplement of probability rainfall*.
- Ministry of Land, Transport and Maritime Affairs (MLTMA) (2012). *Design flood calculation tips*.
- Ministry of the Interior and Safety (MOIS) (2021). *Consultation on assessment of accidental impact, etc. Practical Guidelines*.
- National Emergency Management Agency (NEMA) (2010). *A study for setting urban disaster prevention performance goals in consideration of climate change*.
- Park, J.H., and Lee, J.J. (2022). "Temporal distribution analysis of design rainfall by significance test of regression coefficients" *Journal of Korea Water Resources Association*, Vol. 55, No. 4, pp. 257-266.