

## Box-Cox 변환에 의한 I-D-F 곡선의 평활화

### The Smoothing of Rainfall Intensity - Duration - Frequency Relationships curve by the Box-Cox Transformation

이 희 창\* / 성 기 원\*\*

Lee, Hee Chang / Seong, Kee Won

---

#### Abstract

When available rainfall data is not sufficient, a rough tendency of I-D-F relationship appeared frequently. In fact, rainfall intensity on the curve shows abnormally higher value the longer rainfall duration is applied that gives rise to great confusion to apply a rainfall I-D-F relationships curve to a practical work, however, the research work will present a way to solve above mentioned problem by the use of the Box-Cox transformation formula for a given rainfall data. The study came to a conclusion that the Box-Cox transformation formula is satisfied to utilize in a practical work on the ground of analysis for rainfall data of Sancheong and Yeongcheon.

**Keywords** : Box-Cox transformation, rainfall I-D-F relationships curve, smoothing

---

#### 요 지

I-D-F곡선을 유도할 때 강우자료의 보유연한이 충분하지 않을 경우 지속시간별 강우강도의 변화가 매끄럽게 연결되지 못하는 경우가 발생하기도 한다. 특히 곡선에서, 상대적인 장시간에 강우강도가 크게 되는 문제는 실무적으로 I-D-F 곡선을 이용하는데 큰 혼란을 야기 시킨다. 본 연구에서는 강우자료를 Box-Cox 변환을 이용하여 지속시간과 강우강도의 상관관계를 통해 이러한 문제를 해결하는 방법을 제시한다. 산청과 영천의 강우자료에 대한 분석결과 Box-Cox 변환의 실효성을 확인할 수 있었다.

핵심용어 : Box-Cox 변환, I-D-F 곡선, 평활화

---

\* 건국대학교 대학원 토목공학과 박사 수료  
Doctoral Candidate, Dept. of Civil Engrg., Konkuk Univ., Seoul 143-701, Korea  
(E-mail : leehc@cvnet.co.kr)

\*\* 건국대학교 공과대학 토목공학과 교수  
Professor, Dept. of Civil Engrg., Konkuk Univ., Seoul 143-701, Korea  
(E-mail : kwseong@konkuk.ac.kr)

## 1. 서론

수자원계획의 기준으로 널리 이용되는 확률강우량의 산정을 위해서 강우지속기간별로 적정확률분포형을 결정 후 확률강우강도식이나 강우강도 - 지속기간 - 생기빈도(I-D-F) 관계곡선을 유도하는 것이 일반적이다. 그러나 강우자료의 자료보유 년수 부족문제, 강우자료의 추출시 정확성 문제, 강우지속기간별 자료수의 불일치, 자기우량 기록지에서의 최대치 자료추출 기준에 대한 일반성 결여, 예기치 않은 결측으로 인하여 발생하는 시계열성 결여로 인한 강우자료의 정확성 결여 등과 같은 문제로 인하여 어떤 확률분포형을 사용하여야 하는 가는 여전히 불확실한 과제로 남아있다. 또한 확률강우량에 의해 유도된 확률강우강도식도 기존의 연구들이 서로 다른 결과를 보이고 있으며 강우 관측기록 보유기간 증가로 인하여 적정 확률분포형도 달라지고 있다. 이 경우 심각한 문제는 동일한 재현기간에 짧은 지속기간의 강우강도가 상대적으로 긴 지속기간의 강우강도보다 작아지는 현상이며 이는 실무적으로 매우 곤란한 문제를 야기한다.

따라서 본 연구에서는 기존의 I-D-F 곡선에서 짧은 지속기간의 강우강도가 상대적으로 긴 지속기간의 강우강도보다 작아지는 경우의 I-D-F 곡선을 선정하여 이러한 현상이 발생하지 않는 확률강우강도 또는 관계곡선을 유도함에 그 목적이 있다.

우리나라 기존의 확률 강우강도식은 1960년대 중반부터 연구되었으며, 인수한과 신용배(1964), 변근주(1966), 최영박과 박종훈(1966), 이원환(1967, 1980), 이원환과 변근주(1969), 조훈제(1977), 손광익(1981), 이재준등(1981), 이원환과 박상덕(1992), 이원환등(1993), 유동훈(1995), 한정훈(1996), 이재준과 이정식(1999), 건설교통부(2000), 이재준과 이정식, 박종영(2001)등에 의해 제시된 바 있다.

## 2. 빈도해석과 I-D-F 곡선의 유도

본 연구에 적용하고자 하는 Box-Cox 변환은, G.E.P. Box and D.R. Cox에 의하여 제시된 근사 정규분포로 자료를 변환시키는 방법으로써 그 실용성은 많은 학문분야에서 검증 받은바 있다. 이 기법은 변환계수를 이용하여 관측 자료를 근사화된 정규분포를 따르는 변환자료로 전환시키고, 따라서 통계분석 및 극치 분석이 가능해지는 유용성을 갖고 있다. 이러한 변환기법을 강우강도와 지속기간에 각각 적용할 경우 식(1)과

식(2)를 통하여 변환된 강우강도 및 지속기간을 산출할 수 있게 된다. 이 기법은 Loganathan과 Parkin (1992)에 의하여 제시된바 있으나 실제적 적용성의 평가는 아직 미흡하다.

$$A_i = \begin{cases} \frac{H_i^{\lambda_1} - 1}{\lambda_1} & : \lambda_1 \neq 0 \text{ 일 경우} \\ \ln H_i & : \lambda_1 = 0 \text{ 일 경우} \end{cases} \quad (1)$$

$$B_i = \begin{cases} \frac{Q_i^{\lambda_2} - 1}{\lambda_2} & : \lambda_2 \neq 0 \text{ 일 경우} \\ \ln Q_i & : \lambda_2 = 0 \text{ 일 경우} \end{cases} \quad (2)$$

여기서,

$A_i$  : 변환된 강우 강도

$B_i$  : 변환된 강우 지속 기간

$\lambda_1, \lambda_2$  : 변환계수 (transform coefficients)

$H_i$  : 실측 강우 강도

$Q_i$  : 해당 강우 지속 기간 이다.

식(1),(2)에서 자료의 정규성을 조절하는 변수는  $\lambda_1$  및  $\lambda_2$  로서 이들의 적절한 산정이 필요하다. Loganathan과 Parkin(1992)은 왜곡도변수를 0으로 하는 기법을 이용하였으나 본 연구에서는 최우도기법(Maximum Likelihood Estimation, MLE)을 적용하였는데 이때 이용된 최우도함수는 식(3)과 같다.

$$f(x) = -\frac{n}{2} \ln \left[ \sum_{i=1}^n \frac{(x_i(\lambda) - \bar{x}(\lambda))^2}{n} \right] + (\lambda - 1) \cdot \sum_{i=1}^n \ln(x_i) \quad (3)$$

여기서  $\bar{x}(\lambda) = \frac{1}{n} \sum_{i=1}^n x_i(\lambda)$ 는 변환된 자료의 평균이며  $n$ 은 자료의 수이다.

자료가 정규분포를 따르는 자료로 변환되면, 변환된 강우강도와 지속기간의 상관관계를 이용하여 지속기간에 대한 조건부 강우강도 및 분산 등을 추정할 수 있다. 이러한 통계치를 추정하기 위해서 식(4), (5), (6)등이 이용되었으며 이는 변환된 자료들이 bivariate normal 분포를 이루기 때문이다.

$$E(A | B=b) = E(A) + \frac{Cov(A, B)}{Var(B)} [b - E(B)] \quad (4)$$

$$Var(A | B=b) = (1 - \rho^2) Var(A) \quad (5)$$

$$\text{여기서 } \rho = \frac{Cov(A, B)}{[Var(A)Var(B)]^{0.5}} \quad (6)$$

조건부 평균 및 분산이 주어지면 식(7)과 같은 빈도 계수법을 이용하여 극치값( $a_a$ )을 추정할 수 있다.

$$a_a = E(A | B=b) + z_a [Var(A | B=b)]^{0.5} \quad (7)$$

식(7)에서 계산된 극치값( $a_a$ )은 변환된 자료에 대한 극치값 이므로 이를 역변환을 통하여 변환관계를 해제해야 한다. 이를 식(8)에 나타내었다.

$$h_a = [a_a \cdot \lambda + 1]^{1/\lambda} \quad (8)$$

### 3. 적용 및 결과분석

#### 3.1 기본자료

수문자료를 빈도해석하기 위해서는 자료들이 독립사상이어야 하며 전체적으로 동질성을 가져야 한다. 본 연구에서는 건설교통부의 I-D-F 곡선에서 동일 재현

기간의 짧은 지속기간의 강우강도가 상대적으로 긴 지속기간의 강우강도보다 작아지는 경우의 2개 지점을 선정한 결과 산청지점과 영천지점의 지속기간별 매년 최대치 계열의 동일한 강우자료를 이용하였다. 자료추출은 각 지속기간별로 자기우량기록지로부터 실측을 통해 판독하였으며 관측소와 강우 기록기간 및 보유 년수는 표 1과 같고 괄호 안은 강우자료의 기록 보유 년수이다.

표 1에서 볼 수 있듯이 산청과 영천의 강우자료 보유연한은 기존의 방법을 이용한 극치해석을 수행하기에는 충분하다고 볼 수는 없다. 이에 본 연구에서는 강우강도와 지속기간의 상관관계를 이용하여 기존의 지속기간별로 분리된 자료를 통합함으로써 계산자료의 수를 상승시키고 기존의 빈도곡선 대신 강우강도와 지속기간에 대한 빈도곡면(Frequency Surface)을 제시할 수 있는 절차를 이용하였다.

#### 3.2 Box-Cox 변환에 의한 통계 특성치

산청과 영천지점의 강우자료를 지속기간별로 강우강도로 환산하여 구한 Box-Cox 변환 결과의 통계 특성치는 다음 표2와 같으며, 변환된 자료의 왜곡도계수가 거의 0에 근접하고 있어 2개 지점 모두 거의 정규분포화 되었음을 보여주고 있다. 또한 강우강도와 지속기간의 상관관계가 대단히 높게 나타나게 됨으로서 변환된 자료들은 bivariate normal 분포로 가정함에 부족함이 없는 것으로 판단된다.

표 1. 관측지점과 기록보유 년 수

지점	기록기간 및 보유 년 수	
	10분, 20분, 40분, 50분	30분, 60분, 90분, 120분, 180분, 240분, 360분, 540분, 720분, 900분, 1080분, 1440분, 2880분
산 청	'88 ~ '99 (12개년)	'72 ~ '99 (28개년)
영 천	'88 ~ '99 (12개년)	'71 ~ '99 (29개년)

표 2. Box-Cox 변환에 의한 통계 특성치

구 분	산 청		영 천	
	Intensity ( $\lambda_1=0.1539$ )	Duration ( $\lambda_2=0.0708$ )	Intensity ( $\lambda_1=-0.0425$ )	Duration ( $\lambda_2=0.0720$ )
Mean	3.6430	6.8223	2.4474	6.8626
Variance	1.7209	4.7794	0.6881	4.8205
Standard Deviation	1.3118	2.1862	0.8295	2.1956
Skewness	-0.0327	-0.0350	0.0070	-0.0353
Correlation Coef.	- 0.8820		- 0.9229	

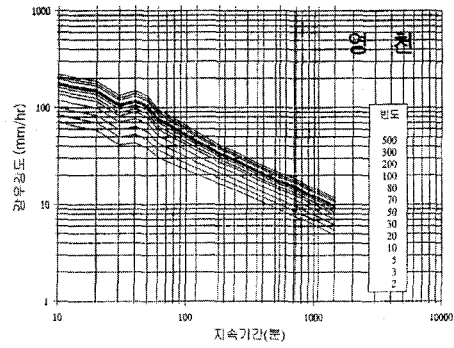
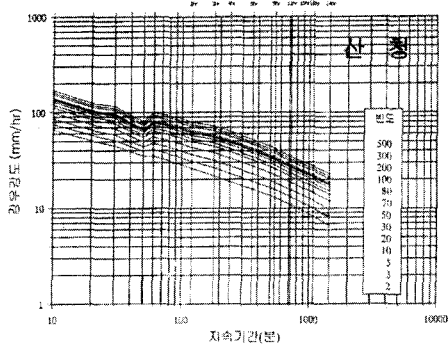


그림 1. 산청지점의 강우강도-지속기간-빈도곡선  
(건설교통부 2000.6)

그림 2. 영천지점의 강우강도-지속기간-빈도곡선  
(건설교통부 2000.6)

표 3. 산청과 영천지점의 재현기간별 지속기간별 강우강도

(단위:mm/hr)

지점	지속기간(분)	재현기간(년)					
		2	5	20	50	100	500
산청	10	67.6	88.3	112.7	127.2	137.8	161.3
	20	53.4	70.4	90.7	102.8	111.6	131.3
	30	46.1	61.1	79.2	90.0	97.8	115.5
	40	41.3	55.1	71.6	81.5	88.7	105.1
	50	37.8	50.6	66.0	75.3	82.1	97.4
	60	35.2	47.2	61.7	70.5	77.0	91.4
	120	26.1	35.5	47.0	54.0	59.1	70.8
	360	15.3	21.4	28.9	33.5	37.0	44.8
	720	10.4	14.8	20.4	23.9	26.4	32.4
1440	6.8	9.8	13.8	16.3	18.2	22.6	
영천	10	77.1	106.7	146.2	171.9	191.6	239.1
	20	54.3	74.9	102.1	119.8	133.3	165.8
	30	44.0	60.5	82.3	96.4	107.2	133.0
	40	37.9	51.9	70.4	82.4	91.6	113.5
	50	33.6	46.0	62.3	72.9	80.9	100.2
	60	30.5	41.6	56.4	65.8	73.1	90.4
	120	20.8	28.3	38.1	44.4	49.2	60.6
	360	11.1	14.9	19.9	23.1	25.5	31.3
	720	7.3	9.8	13.0	15.1	16.6	20.3
1440	4.8	6.3	8.4	9.7	10.6	13.0	

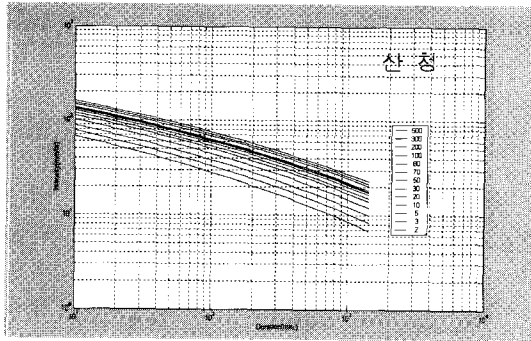


그림 3. 산청지점의 강우강도-지속기간-빈도곡선  
(본 연구)

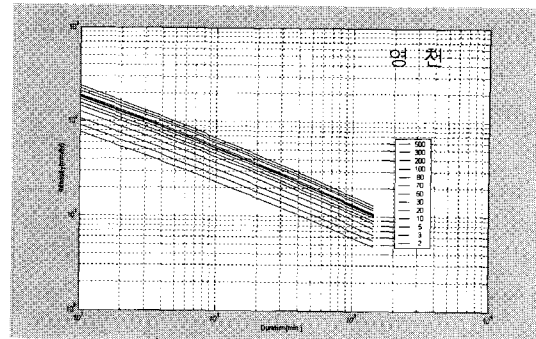


그림 4. 영천지점의 강우강도-지속기간-빈도곡선  
(본 연구)

표 4. 확률 강우량의 상대오차

(단위 : %)

지 점	지속기간(분)	재 현 기 간 (년)					
		2	5	20	50	100	500
산 청	10	-2.7	-2.8	4.3	0.9	1.7	4.6
	30	-2.7	-4.1	-3.1	-1.8	-0.8	1.0
	50	-7.9	-15.6	-20.0	-21.0	-21.0	-21.2
	60	-0.6	0.4	2.7	4.1	5.3	7.4
	1440	-0.3	0.1	0.4	0.7	0.8	1.0
영 천	10	-9.4	-7.9	-7.0	-6.7	-7.0	-8.4
	30	-5.8	-2.4	-16.1	0.4	0.6	0.0
	50	12.5	17.6	22.2	23.4	24.0	31.7
	60	1.6	4.6	6.2	6.4	6.4	5.8
	1440	1.4	-3.4	-7.0	-8.5	-9.5	-11.6

### 3.3 I-D-F 관계곡선의 유도

“1999년도 수자원 관리기법 개발 연구조사 보고서 제1권 한국 확률강우량도 작성, 2000.6 건설교통부”에 따르면 산청과 영천의 I-D-F 관계곡선은 그림 1 및 2로 나타나있는데 지속기간 20~60분 사이의 곡선은 극심한 이상현상을 보여주고 있다.

이러한 현상은 자료의 보유연한과 수집된 자료의 이상치가 지속기간별로 뚜렷한 경향을 유지하지 않고 분포되어 있기 때문에 발생한 것으로 보이며 특히 지속기간별로 일관성이 결여된 분포함수가 선택될 경우 이러한 이상 현상은 더욱 두드러질 수 있다.

이러한 현상을 완화시키기 위하여 본 연구에서는 지

속기간과 강우강도의 상관관계를 이용하였으며 특히 극치값 추정을 위해서 Box-Cox 변환을 이용하였다. 제시되었던 공식들을 이용하여 산청과 영천의 재현기간, 지속기간별 강우강도를 추정하면 표 3에 수록된 결과를 얻을 수 있으며 또한 이 결과를 도시하면 그림 3, 4와 같다.

그림과 표에서 나타내었듯이 본 연구를 통한 I-D-F 곡선은 이상현상의 재현 없이 평활화 되었음을 알 수 있다. 이러한 결과는 실무적으로 대단히 유용한 것으로 소규모 수공구조물 설계에 있어서 안정성 있는 기본자료로서의 가치를 갖는다.

그러나 자료가 평활화 되었다고 하더라도 전반적인

값들의 일치성은 평가되어야 마땅하다. 이를 위하여 “1999년도 수자원 관리기법 개발 연구조사 보고서 제1권 한국 확률강우량도 작성 보고서, 2000.6 건설교통부”와 본 연구에 대한 상대오차를 재현기간, 지속기간 별로 계산하고 이를 표 4에 나타내었다.

표 4에서 볼 수 있듯이 이상현상이 극심한 지속기간 50분에 해당되는 상대오차는 매우 크지만 나머지 부분 들은 거의 일치한다. 그러나 지속기간 10분 및 1440분 과 같은 한계자료에 대한 높은 재현기간의 수치들은 다 소 큰 상대오차를 보여주고 있다. 만일 “1999년도 수자 원 관리기법 개발 연구조사 보고서 제1권 한국 확률강 우량도 작성 보고서, 2000.6 건설교통부”의 결과를 신 뢰할 경우 이는 가정된 정규분포의 특수성에서 비롯된 것으로 추정된다. 일반적으로 정규분포를 통해 추정되 는 높은 재현기간에 대한 극한값의 신뢰성은 평가되기 어렵다. 따라서 높은 재현기간에 대한 적용성은 별도로 검토해 볼 필요가 있다.

#### 4. 결 론

본 연구는 I-D-F 곡선의 평활하지 못한 구간의 개 선을 위하여, 건설교통부의 I-D-F 곡선식과 동일한 강 우자료를 이용하여 Box-Cox 변환에 의하여 I-D-F 곡선을 유도한 것으로 다음과 같은 결과를 얻었다.

- (1) 기존의 연구에서는 I-D-F 곡선을 유도하기 위 하여, 강우자료 입수와 적합도 검정을 거쳐 대표 확률분포형을 선정하기까지는 많은 시간과 노력 이 필요하지만 본 연구에 적용한 Box-Cox 변 환은 자료를 정규 분포화 시킴으로서 일관된 분 포형을 만들 수 있었다.
- (2) 비교 대상이 된 건설교통부의 I-D-F곡선에서는 동일재현기간의 짧은 지속시간 강우강도가 긴 지속기간의 강우강도보다 작아지는 구간이 발생 되지만 본 연구에 적용한 Box-Cox 변환에 의한 I-D-F곡선은 문제되는 구간이 없이 매끄럽게 연결되었다.
- (3) 확률강우량의 상대오차를 비교해 본 결과 건설 교통부 I-D-F곡선의 평활하지 못한 구간에서 그 상대오차가 비교적 크게 나타났다.
- (4) 따라서 I-D-F곡선을 유도할 때, 강우자료의 보 유연한이 충분하지 않거나, 일관성이 없을 경우 발생될 수 있는 평활하지 못한 구간은 Box-Cox 변환을 통하여 개선할 수 있었다.

- (5) 100년 이상의 극치 강우강도를 추정하기 위해서 는 추후 연구를 통한 비교 및 보완이 필요하다.

#### 감사의 글

본 연구를 수행함에 있어서 해당 강우자료를 흔쾌히 제공해 주신 한국건설기술연구원의 김규호 박사님께 감 사드립니다.

#### 참고문헌

- 건설교통부 (2000.6). **1999년도 수자원 관리기법 개발 연구조사 보고서 제1권**, 한국건설기술연구원, p.289, p.291.
- 기상청 (1971-1999). **기상월보**
- 기상청 (1971-1999). **기상년보**
- 변근주 (1966). 국내 지역별(서울 외 5개 지방) 확률 강우강도에 관한 연구, 석사학위논문, 연세대학교
- 손광익 (1981). 우리나라 확률강우량 산정에 관한 수문 학적 연구, 석사학위논문, 연세대학교
- 신창동 (1993). 우리나라 강우의 적정분포형 결정에 관 한 연구, 석사학위논문, 금오공과대학교
- 안수환, 신용배 (1964). “서울지방의 강우특성에 관한 연구” **대한토목학회지, 대한토목학회**, 제12권, 제4 호, pp.17~35.
- 유동훈 (1995). “확률 강우강도식의 일반화” **한국수자 원학회 학술발표회 논문집**, 한국수자원학회, pp. 325~330.
- 이원환 (1967). “국내 지역별 강우특성과 확률강우량 산정에 관한 연구” **대한토목학회지, 대한토목학회**, 제15권, 제3호, pp. 28~38, 1967.
- 이원환 (1980). “도시하천 및 하수도 개수계획상의 계 획강우량 설정에 관한 추계학적 해석” **대한토목학 회논문집, 대한토목학회**, 제28권 제4호, pp.81~ 94.
- 이원환, 박상덕 (1992). “서울지방의 통합형 확률강우 강도식” **대한토목학회 논문집, 대한토목학회**, 제 12권, 제4호, pp.135~143.
- 이원환, 박상덕, 최성열 (1993). “한국 대표확률강우 강 도식 유도에 관한 연구” **대한토목학회 학술발표 개요집(III)**, 대한토목학회, pp.135~138
- 이원환, 변근주 (1969) “중소하천 및 도시하수도 계획 설계에 필요한 확률강우강도식의 유도” **대한토목학 회지, 대한토목학회**, 제16권, 제4호, pp.1~11.

- 이재준, 손광익, 이원환, 이길춘 (1981). “강우빈도 및 강우강도의 시공적 분포분석에 관한 연구” **한국수문학회지**, 한국수문학회, 제14권, 제4호, pp.53~72.
- 이재준, 이정식, 박종영 (2001). “계획강우량 설정을 위한 권역별 확률강우 강도식의 유도” **대한토목학회 논문집**, 대한토목학회, 제21권 제1-B호, pp.1~10.
- 조홍제 (1977). 한국 지점강우의 장시간 확률강우강도식의 유도, 석사학위논문, 연세대학교
- 최영박, 박종훈 (1966). “한국 주요도시 강우강도형의 지역별 특성연구” **대한토목학회지**, 대한토목학회, 제14권, 제1호, pp.15~29.
- 한정훈, 김경덕, 허준행, 조원철 (1996). “선형화 기법에 의한 확률강우강도식의 유도” **한국수자원학회 학술발표회 논문집**, 한국수자원학회, pp.232~237.
- G.E.P. Box and D.R. Cox(1964). “An Analysis of Transformations.” *J. Royal Stat. Soc.*, Series B. vol. 26, pp.211~252.
- G.V. Loganathan M. and M.A. Parkin (1992). “A Frequency Surface for Rainfall Intensity and Duration.” *Water Resources Planning And Management*, pp.387~389.
- Nathabandu T.Kottegoda and Renzo Rosso (1997). *Statistics, Probability, and Reliability for Civil and Environmental Engineers*. The McGraw-Hill Companies, Inc., pp.86~162, p.381.
- (논문번호:02-90/접수:2002.11.27/심사완료:2003.02.11)