갈수기 발전용댐 운영을 위한 Lindley 확률분포모형의 적용성 검토

Applicability of Lindley Distribution for Hydroelectric Dam Operations in the Dry Season

신홍준* · 안현준** · 정창삼***

Shin, Hongjoon*, Ahn, Hyunjun**, and Jeong, Changsam***

Abstract

The long-term low-flow data are necessary for efficient planning of water resources and for estimating accurate quantiles via runoff data analysis at point. However, the short recording time period, low confidence, inconsistent distribution model, and parameter estimation method, make it difficult to estimate a proper low-flow quantile for each return period. In this study, the Lindley distribution model, which is a mix of the exponential and the gamma distribution models and has been verified as efficient by previous studies, was used to analyze the low-flow frequency using dam inflow data. We studied its applicability via comparison with statistics of observed data and other models already used for low-flow frequency analysis. For this, we carried out a performance analysis through a low-flow frequency analysis of inflow data from the hydroelectric dam and the reappearance capacity assessment of observed data at the Han river watershed. As a result, the hydrological applicability of the Lindley distribution model and its relative qualitative and quantitative excellence compared to the existing model were verified.

Key words : Lindley Distribution, Low-flow Frequency, Hydroelectric Dam, Monthly Inflow

요 지

수자원의 효율적 계획수립을 위해서는 장기간의 유입량 자료는 필수자료이며, 이를 활용한 분석대상 지점의 예측 또는 기대 유입량 분석을 통해 정확한 결과를 산정하는 것이 중요하다. 그러나 적절한 빈도별 기대 유입량의 추정이 어려운 것은 수집된 자료기간이 짧거나 신뢰성 결여, 표본자료의 특징을 대표할 수 있는 분포모형의 부적합성 등에서 발생하는 것으로 나타나고 있다. 본 연구에서는 지수 및 감마 분포모형의 혼합구조 형식을 이루고 있으며 최근 다양한 연구를 통해 그 효율성이 증명되고 있는 Lindley 분포모형을 댐 유입량 자료에 적용하여 가뭄빈도해석을 수행하고 관측자료의 통계량 및 기준에 활용되고 있는 확률분포모형과의 비교를 통해 그 적용성을 살펴보았다. 이를 위해 한강수계내 발전용댐을 선정하여 유입량 자료를 수집하고 가뭄빈도해석을 통해 성능분석 및 관측자료의 재현능력 평가를 수행하였다. 그 결과 Lindley 분포모형의 수문학적 적용성을 확인하였으며, 기존 모형에 비해 정성적, 정량적으로 우수함을 확인하였다.

핵심용어 : Lindley 분포모형, 가뭄빈도, 발전용댐, 월유입량

1. 서 론

가뭄이란 하천유량의 감소로 수자원 관리계통으로부터 기대되는 용수의 공급을 충족시키지 못하는 것을 말하며,

이때의 하천유량을 갈수량(Low flow)이라 한다. 수자원 계획 시 수계 내 시설물 규모를 결정하는 기준으로 갈수량을 이용하고 있으며, 이는 최소한의 하천기능 유지를 위해 필요한 유량으로 우리나라의 경우 1년 중 355일은 이보다 높은

*정회원, 한국수력원자력 중앙연구원 선임연구원(E-mail: h.j.shin@khnp.co.kr)

Member, Senior Researcher, Central Research Institute, Korea Hydro & Nuclear Power

**교신저자, 정회원, 한국환경정책·평가연구원 물국토연구본부 통합물관리연구실 연구원, 공학박사(Tel: +82-44-415-7431, Fax: +82-44-415-7431, E-mail: hjahn@kei.re.kr)

Corresponding Author, Member, Ph.D., Researcher, Division for Integrated Water Management, Water and Land Research Group, Korea Environment Institute

***정회원, 인덕대학교 토목환경설계과 교수(E-mail: csjeong@induk.ac.kr)

Member, Professor, Department of Civil & Environmental Engineering, Induk University

수위를 유지하는 유량으로 제시하고 있다(KWRA, 2009). 장기간의 갈수량 자료는 댐 용량 결정, 수질 예측, 용수공급계획 수립, 가뭄대책 수립, 하천유지유량 결정 등 수자원의 효율적 계획수립에 필요한 자료이며, 이를 위해 해당 지점의 유출량 자료 분석을 통해 정확한 결과를 산정하는 것이 중요하다(MOCT, 2000). 이를 위해 가장 좋은 방법은 신뢰성 있는 유량 자료를 사용하는 것이나, 우리나라의 경우 신뢰성을 갖는 유량 자료의 생성 가능한 관측소의 수가 제한적이고 일부 유역의 경우 그 신뢰성이 낮아 유량 자료의 직접 활용보다는 강우량을 유출량으로 변환하는 방법이 주로 사용되고 있으며, 이를 위해 대표적으로 활용되는 것이 강우 관측자료를 이용한 가뭄빈도해석이다. 그러나 이와같이 적절한 빈도별 가뭄우량의 산정이 어려운 것은 강우기록이 짧거나 신뢰성 결여, 분포모형의 부적합성, 분포모형의 매개변수 추정방법 등에서 발생되는 것으로 나타나고 있다.

최근 가뭄빈도해석에 대한 연구는 꾸준히 진행되고 있으며, 그 중에서 Kwak et al. (2012)은 수문학적 가뭄 측면에서 평창강과 남한강 등 한강의 상류유역을 대상으로 Copula 기법을 이용하여 수문학적 가뭄의 결합 확률분포모형을 유도한 바 있으며, Cho (2013)와 Lee et al. (2015)은 표준강수 지수를 이용하여 우리나라의 과거 가뭄사상을 규모적 관점의 정량적 평가를 실시하여 우리나라에 발생했던 가뭄의 재현기간 및 규모에 따른 평가를 수행하였다. 또한 Yu et al. (2017)은 가뭄심도와 지속기간을 동시에 고려할 수 있는 이변량 가뭄빈도해석을 5개 다목적댐에 수행하여 특정 가뭄상황에서의 용수공급능력을 평가하였으며, Jeong et al. (2020)은 관측자료를 기반으로 하여 과거 가뭄 특성 및 기후변화 시나리오를 활용한 장래 발생 가능한 극치 가뭄에 대한 장기전망을 수행한 바 있다.

가뭄빈도해석 기법 자체에 관한 연구로 Ryoo et al. (2009)은 이수기 유입량 빈도분석을 위해 사용한 누가차분법이 빈도값의 차이를 통해 계산을 수행하여 부분적으로 음수에 해당하는 값이 추정되고 가뭄빈도가 증가함에도 수문량이 증가하는 이상 현상이 발생하며 분석 초기에는 극한 가뭄이 오고 분석말기에는 평년수준의 유입량이 유입되는 것으로 분석되어 새로운 가뭄빈도 유입량의 분해방법을 제시한 바 있다. Jung et al. (2012)은 갈수량을 산정하기 위하여 화천댐과 춘천댐의 적정 확률분포모형을 선정한 후, 재현기간에 따른 월별 유입량 계열로 산정한 가뭄빈도해석 방법과 월 단위 누가유입량 계열의 차로 산정한 가뭄빈도해석 방법의 비교 연구를 수행한 바 있다.

이러한 통계적 기법 기반의 가뭄빈도해석을 통해 최적의 수자원 관리를 수행하기 위해서는 분석하고자 하는 자료를 대표할 수 있는 확률분포모형을 찾는 것이 매우 중요하다. 최적의 확률분포모형은 관측 또는 수집된 데이터의 특성에 따라 각기 다를 수 있기 때문에 다양한 확률분포모형을 적용하고 평가해보는 과정이 필요하다(Heo, 2016). 또한 타 분야에서 널리 활용되고 있는 확률분포모형을 새롭게 도입하여 그 적용성을 검토하고 이를 극대화하기 위해 모형의 매개변수 확대 등 새로 도입된 모형의 수정 및 보완을 통해 정확성을 높이는 연구도 활발히 진행 중이다(Shao et al., 2004; Ganora and Laio, 2015; Seo et al., 2017; Heo et al., 2019).

따라서 본 연구에서는 지수 및 감마 분포모형의 혼합구조 형식을 이루고 있으며 최근 다양한 연구를 통해 그 효율성이 증명되고 있는 Lindley 분포모형을 댐 유입량 자료에 적용하여 가뭄빈도해석을 수행하고 관측자료의 통계량 및 기준에 활용되고 있는 확률분포모형과의 비교를 통해 그 적용성을 살펴보고자 하였다.

2. 적용 이론

2.1 댐 유입량 가뭄빈도해석

가뭄빈도해석 방법은 일반적인 확률강우량 추정 또는 홍수빈도해석의 전체적인 방법과 절차에 있어서 동일하지만, 가장 큰 차이점은 확률분포모형의 좌측 꼬리(Left tail)에 해당하는 부분을 활용하여 확률수문량을 추정한다는 것이다. 기존의 확률강우량 추정 및 홍수빈도해석에서 비초과확률(q)은 1에서 재현기간(T)의 역수를 빼준 것과 같지만, 가뭄빈도해석의 경우 재현기간(T)의 역수로만 표현되며, 이는 결국 수문량이 작아질수록 가뭄의 빈도가 증가하는 것을 의미한다.

유입량 자료를 활용한 가뭄빈도해석의 방법으로는 크게 두 가지 접근방법이 있다. 첫 번째는 월별 유입량 자료를 활용한 가뭄빈도해석과 두 번째는 누가유입량 자료를 활용한 가뭄빈도해석이다. 누가유입량 자료를 활용한 방법을 이용할 경우 수집된 자료특성에 따라 재현기간 또는 비초과확률이 증가함에 따라 확률수문량이 증가하는 통계적 오류가 발생할 수 있다(Ryoo et al., 2009). 따라서 본 연구에서는 월별 댐 유입량 자료를 활용한 방법을 이용하여 가뭄빈도해석을 수행하였으며, Table 1에 홍수빈도해석과 가뭄빈도해석시 수문량 산정방법의 차이를 정리하였다.

Table 1. Comparison of the Difference between Monthly Inflow Method and Difference in Cumulative Sum Method

Method	Monthly inflow	Difference in cumulative sum
	$Q_i^T = \left\{ \sum_{i=1}^N inflow(i) \right\}$	$Q_i^T = q_{j,i}^T - q_{j-1,i+1}^T$

여기서, i 는 해당 월을 의미하며, j 는 누가기간, N 은 자료 기간을 의미하고, Q_i^T 는 i 월에 대한 재현기간(T)의 빈도값 $q_{j,i}^T$, 누가기간(j)에 대한 시작 월(i)의 재현기간(T)에 해당하는 빈도값이다.

2.2 Lindley 분포모형

Lindley 분포모형은 Lindley (1958)에 의해 소개된 확률 분포모형으로 지수 분포모형과 감마 분포모형의 혼합 (Mixture) 구조 형식을 이루고 있다. 주로 장비와 생산 및 운영의 수명 주기를 설명하는 분야에서 주로 이용되고 있으며 공학 뿐만 아니라 의학 및 생물학 등 여러 분야에서 사용되고 있다. Ghitany et al. (2011)과 Adhikari and Srivastava (2014)의 연구에 의하면 기존의 지수 분포모형 (Exponential distribution)보다 효율적인 모형이며, 기존의 지수 분포모형과 감마 혼합 분포모형을 활용한 모의실험을 통해 Lindley 분포모형이 지수 분포모형보다 다양한 조건에서 유연한 모형이라는 것을 보인 바 있다(Ghitany et al., 2008). Lindley 분포모형의 확률밀도함수 및 누적분포함수는 Eqs. (1), (2)와 같다(Tomy, 2018).

$$f(x) = \frac{\theta^2}{\theta + 1} (1 + x) e^{-\theta x}, \quad x > 0, \theta > 0 \quad (1)$$

$$F(x) = 1 - \frac{\theta + 1 + \theta x}{\theta + 1} e^{-\theta x}, \quad x > 0, \theta > 0 \quad (2)$$

여기서, x 는 변수($x > 0$), θ 는 scale parameter ($\theta > 0$)를 나타낸다.

2.3 베이지안 정보 지수(Bayesian Information Criterion, BIC)

BIC 방법은 Bayesian 이론에서 우도함수와 사전분포

(Prior probability distribution)를 통해 계산된 사후분포 (Posterior probability distribution)를 근사화하여 유도된 통계량을 이용하여 주어진 모형 중에서 최적의 모형을 선택하는 기준을 마련한다(Schwarz, 1978). 우도함수를 이용하여 최적의 모형을 선정하기 때문에 BIC의 값이 가장 낮은 모델이 최적의 모형을 의미하게 된다. 일반적으로 모형을 구성하는 매개변수가 많을 경우 우도가 증가하여 적합한 모형을 구성할 수 있지만, 매개변수가 많을수록 모형을 추정하기가 까다로워지거나 추정이 된 모형에서 때때로 과적합(Overfitting) 문제가 발생하기도 한다. 따라서 BIC에서는 모델의 매개변수의 수에 대한 페널티 항을 고려함으로써 최소한의 매개변수를 활용한 최적 분포모형 선정에 대한 기준을 마련해주며, 다음 Eq. (3)과 같다.

$$BIC = \ln(n)k - 2\ln(L) \quad (3)$$

여기서 n 은 적용된 데이터의 개수, k 는 매개변수의 개수, L 은 최대우도를 나타낸다.

3. 적용대상 및 자료분석

3.1 적용 대상 및 분석방법

본 연구에서는 댐 운영에 필요한 유입량 자료를 대상으로 가뭄빈도해석을 통해 Lindley 분포모형의 수문학적 적용성을 검토해보고자 하였다. 이를 위해 한강수계에 위치하고 있는 발전용댐, 홍수조절용댐, 용수전용댐, 다목적댐 등 다양한 목적의 댐(Table 2) 중 발전용댐을 대상으로 하였다. 또한, Fig. 1에서는 한강수계에 분포되어 있는 주요 댐들의 분포현황을 모식화하였다. Fig. 1에서 검은색으로 표시된 댐은 발전용댐이며 회색으로 표시된 댐은 홍수조절용댐, 흰색으로 표시된 댐은 다목적댐이다. 본 연구에서는 한강수계의 댐 중 발전용댐인 화천댐(HC), 춘천댐(CC),

Table 2. Summary Specification of Dams According to the Purposes of Dams

Purpose	Name	High (m)	Flood water level (El.m)	Total storage (Million m ³)	Basin Area (km ²)
Hydropower	Hwacheon	81.5	183.0	1,018	3,901
	Chuncheon	40.0	104.9	150	4,736
	Uiam	23.0	73.36	80	7,709
	Cheongpyeong	31.0	52.0	186	9,921
	Paldang	29.0	27.00	244	23,800
Multi-purpose	Soyang	123.0	198.0	2,900	2,703
	Chungju	97.5	145.0	2,750	6,648
	Hoengseong	48.5	180.0	87	209
Flood control	Peace	125.0	264.6	2,630	3,227
Water supply	Gwangdong	39.5	675.3	11,000	125

Fig. 1. Dam Locations in Hangang River Basin

청평댐(CP), 괴산댐(GS)을 대상 댐으로 선정하였다. 상대적으로 상류 및 지류를 통해 유입되는 유입량의 영향이 크지 않은 지점들을 적용 대상으로 선택하였다. 각 댐별로 월별 변화량 및 사용수량 등을 고려한 월 평균 자기유입량을 산정하여 가뭄빈도해석을 위한 입력자료로 활용하였다. 가뭄빈도해석 방법으로는 월별 유입량 자료를 활용하여 수행하는 방법을 채택하였는데, 누가 유입량 자료를 활용한 가뭄빈도해석 방법으로 분석을 수행할 경우 재현기간이 증가함에도 불구하고 추정량이 증가하는 역전현상이 발생할 수 있기 때문이다.

3.2 유입량 자료

Fig. 2는 HC (1953~2019), CC (1971~2019), CP (1971~2019), GS (1981~2019)의 댐별 자기유입량을 Box-Whisker 도식화한 것으로, HC, GS와 달리 CC, CP의 겨울철~봄철 자기 유입량이 0에 근접할 정도로 작게 나타나는 것을 볼 수 있다. 또한 각 댐별로 이수기에 비해 상대적으로 홍수기 유입량의 분산이 크게 나타나는데, 이는 년도별 강우 정도의 차이에서 발생한 것으로 특히 수계 최상단에 위치한 HC와 GS가 상대적으로 유입량의 변동폭이 큰 것으로 나타났다.

Fig. 2. Box-whisker Plot of Monthly Dam Inflow

4. 적용결과

4.1 Lindley 분포모형 적용결과

먼저 HC, CC, CP, GS의 유입량 자료를 이용하여 각 댐의 월별 Lindley 분포모형의 매개변수를 추정하였다. 매개변수 추정방법은 널리 이용되고 있으며, 표본크기가 일정량 이상 확보될 경우 가장 정확한 추정이 가능한 최우도법(maximum likelihood estimation method, MLE)을 적용하였다. MLE 방법을 통해 댐별, 월별 Lindley 분포모형의 매개변수 추정결과는 Table 3과 같다.

Fig. 3에서는 구축된 모형의 확률밀도함수(5월)와 유입량 자료의 히스토그램(5월)을 도시하였다. 이때 기존 가뭄빈도 해석시 널리 사용되고 있는 Weibull 분포모형 및 Gumbel 분포모형을 추가하여 함께 비교하였다. Weibull 분포모형의 경우 국내외에서 저수지 또는 댐 유입량 자료 등 다양한 자료를 활용하여 가뭄빈도해석에 적용된 바 있다(Nathan and McMahon, 1990; Adamowski, 1996). 특히 Kang and

Seoh (1995)는 저수지 규모를 결정할 때 기준이 되는 설계 갈수량 추정시 Weibull 분포모형을 적정 모형으로 선택한 바 있다. 또한 Maeng et al. (2009)에 의하면 저수지 유입량 산정 시 최적의 도시위치공식(plotting position formula) 중에서 Weibull 공식이 Cunnane, Gringorten 공식 등 다른 공식보다도 오차가 작은 것으로 밝힌 바 있다. 따라서 본 연구에서는 Lindley 분포모형의 적용성 검토를 위해서 발전용댐 운영 및 홍수기 댐 수위운영 등에 널리 사용되고 있는 Weibull 분포모형과 극치수문사상에 적절하다고 알려져 국내외에서 널리 사용되고 있는 Gumbel 분포모형을 이용하여 비교분석하였다.

4.2 적합성 검토결과

4개 댐에 대하여 월별로 구축된 총 48개 모형에 대하여 적합성 검토를 위하여 BIC 분석을 수행하였다. Table 4는 HC, CC, CP, GS의 1월부터 12월까지 48개의 모형에 대한 BIC 결과값을 보여주고 있다. 회색 음영으로 표시된 모형은

Table 3. Scale Parameters of Lindley Distribution Models for Each Dam

Month	1	2	3	4	5	6	7	8	9	10	11	12
HC	0.1156	0.1079	0.0568	0.0295	0.0358	0.0321	0.0076	0.0076	0.0135	0.0532	0.0647	0.0942
CC	0.8812	0.4891	0.3097	0.1636	0.1193	0.1031	0.0241	0.0266	0.0625	0.3059	0.3508	0.1385
CP	0.3488	0.1958	0.1281	0.0755	0.0579	0.0502	0.0099	0.0140	0.0241	0.1165	0.1667	0.2184
GS	0.6495	0.4164	0.2833	0.2301	0.2252	0.1528	0.0423	0.0520	0.0823	0.2634	0.4160	0.4650

Fig. 3. Histogram and Probability Density Function by Each Dam with Inflow Data in May

Table 4. The BIC Test Results for Various Distribution Models for Each Month

Month	HC			CC			CP			GS		
	LIN	WBL	GUM	LIN	WBL	GUM	LIN	WBL	GUM	LIN	WBL	GUM
1	488.21	496.98	477.80	68.56	71.16	76.47	169.18	173.24	177.41	139.42	129.87	120.42
2	493.44	487.56	479.09	101.77	104.15	110.10	225.61	226.92	234.75	186.86	189.54	197.54
3	598.83	607.31	610.77	202.56	205.96	210.87	335.45	339.18	343.99	207.90	203.90	205.95
4	684.00	691.15	698.92	321.86	323.12	330.27	397.00	396.98	402.84	234.27	237.27	238.92
5	662.28	668.75	674.65	326.01	323.48	336.63	413.10	412.69	424.94	244.44	246.66	255.87
6	699.75	693.59	717.34	296.72	278.21	302.23	440.24	414.53	448.71	290.71	279.84	302.87
7	875.59	879.21	887.80	514.38	505.06	518.06	582.55	583.50	591.08	375.65	378.27	383.41
8	869.73	875.57	879.75	488.56	491.58	499.42	553.04	552.36	559.80	364.95	365.75	373.77
9	810.92	809.21	826.82	436.65	413.77	442.69	520.84	504.90	527.22	332.97	330.56	342.68
10	597.53	605.68	599.06	217.43	220.60	227.83	355.19	355.73	367.10	227.03	231.17	235.47
11	574.80	586.60	578.55	222.46	213.74	237.53	310.75	313.38	321.96	179.99	169.64	169.04
12	513.33	504.34	505.42	255.31	174.96	248.83	259.50	262.23	270.22	171.92	165.21	164.49

Lindley (LIN), Weibull (WBL), Gumbel (GUM) 분포모형 중에서 월별로 가장 적합한 모형을 나타낸다. 결과에 따른 분포모형별 최적분포모형의 개수와 점유율은 전체 49개 모형 중 LIN 분포모형이 26개(54%), WBL 분포모형이 17개(35%), GUM 분포모형이 5개(10%)로 조사되어 LIN 분포모형이 적합정도가 가장 높은 것으로 나타났다.

본 연구에서는 분기 또는 계절에 따른 최적 분포모형의 채택률을 살펴보기 위하여, 전체 대상 댐에 대하여 각 모형별로 월별 최적 분포모형의 채택률을 일정 기간에 따라 분석해 보았다(Fig. 4). Fig. 4에서 MAM은 3월(March), 4월(April), 5월(May)을 의미하며 우리나라의 계절상 봄에 해당한다. JJA는 6월(June), 7월(July), 8월(August)로 우리나라의 여름에 해당한다. SON과 DJF 역시 마찬가지로 각각 9월~11월(Sep~Nov), 12월~2월(Dec~Feb)까지를 의미하며 계절로는 각각 가을과 겨울을 의미한다. 또한 본 연구에서는 계절적 관점을 고려한 최적 분포모형에 대한 분석뿐만 아니라 우리나라 기상학의 특성을 반영한 분석을 수행하였다. 이를 위해 1년을 2개의 기간으로 구분하였다. WET은 홍수기에 해당하는 기간으로 우리나라에서 본격적인 장마 또는 태풍의 영향을 많이 받는 6월부터 9월까지로 정의하였다. DRY는 갈수기로 정의하였으며 상대적으로 WET 기간에 비하여 강수량이 적은 기간으로 10월부터 5월까지의 기간을 의미한다. Fig. 4를 살펴보면 구분된 기간에 따른 최적 분포모형 선정결과를 포함하고 있다. 우선 MAM을 살펴보면 LIN 분포모형이 67%로 최적 분포모형으로 채택되고 있다. 이어 JJA에서는 LIN 분포모형과 WBL 분포모형이 50%씩 각각 차지하고 있으며, SON에서는 LIN 분포모형의 최적 분포모형의 채택 비율은 변함이 없으나 GUM 분포모형이 8% 증가하였다. 마지막으로 DJF에서는 GUM 분포모형의 채택률이 33%로 증가하였고 WBL은 17%로 감소하였다. 다음으로 WET 시

기에는 LIN 분포모형이 38%밖에 채택되지 않은 반면 WBL 분포모형은 63%가 최적 분포모형으로 선정되었으며, DRY 시기에는 LIN 분포모형이 63%의 비율로 최적 분포모형으로 선정되었다. WBL 분포모형과 GUM 분포모형은 각각 22%와 16%의 비율로 선정되었다. 즉, 홍수기에는 WBL 분포모형의 적합도가 가장 뛰어나고 갈수기에는 LIN 분포모형이 가장 뛰어나다고 할 수 있으나, 가뭄빈도해석의 특성을 고려하면 홍수기보다는 갈수기에 적합한 모형이 더 적절하다고 판단된다.

Fig. 4. Summary of the Goodness-of-fit Test Results according to Seasons for Each Distribution by BIC

다음 Fig. 5는 댐별 분포모형별 연간 RMSE 합계, Fig. 6은 분포모형별 Box-Whisker plot을 나타낸 그림으로, 댐에 관계없이 LIN 분포모형의 오차 및 분산도가 가장 작은 것으로 나타났다. 일반적으로 가뭄빈도해석에서 많이 적용되어 오던 WBL 분포모형보다 더 작은 RMSE 및 분산을 나타내어 LIN 분포의 적용가능성이 충분히 확인된다고 판단된다.

Fig. 5. Comparison of RMSE between Observation Data Mean and Theoretical Mean of Each Distribution for Each Dam

Fig. 6. Boxplots of the RMSE with Whole Models by Each Distribution Model

4.3 관측유입량과 분포모형별 월 평균 유입량 비교

댐을 이용한 수자원 관리 측면에서 평균 유입량 자료는 가장 기본이 되는 동시에 매우 중요한 의미를 갖는 자료로, 만약 댐 유입량을 비교적 정확히 예측 또는 분석할 수 있다면 효율적인 댐 운영을 통해 보다 안정적인 수자원 활용이 가능하다. 따라서 본 연구에서는 구축된 월별 분포모형으로부터 기대되는 평균 유입량과 실제 각 댐별로 관측된 월별 평균 유입량 자료를 비교해보았다. 비교 방법으로는 실제 댐별로 관측된 월별 유입량 자료의 평균값을 기준값 100%로 가정하고 구축된 각 분포모형으로부터 통계적으로 기대할 수 있는 이론적 평균값과의 차이를 살펴보고 그 결과는 Fig. 7과 같다. Fig. 7에서 붉은색 원 기호와 실선은 기존 관측 유입량 대비 LIN 분포모형의 유입량의 비율을 나타낸다. 마찬가지로 삼각형 모양의 기호와 보라색 실선은 WBL

분포모형을, X 모양의 기호와 파란색 실선은 GUM 분포모형의 값을 나타낸다. Fig. 7에서 모형으로부터 예측된 값(기호 또는 실선)이 전체적으로 원형의 모형을 갖추어 1월부터 12월까지 안정적인 유입량을 재현해낸다고 볼 수 있다. LIN 분포모형의 경우는 Fig. 7(b)의 1월을 제외하고는 전체적으로 실제 관측값의 90%~100% 사이에 모든 값이 분포되고 있다. 즉, 관측 유입량 값을 10% 이내의 범위에서 재현하고 있다고 볼 수 있다. 반면 WBL 분포모형과 GUM 분포모형의 경우 일부 구간에서는 LIN 분포모형보다 다소 높은 정확도로 관측된 유입량 값을 재현하고 있지만, 5월, 6월, 12월 등 특정 기간에서는 관측된 유입량 자료와 비교하여 과대 또는 과소 등 매우 큰 차이가 나는 것을 볼 수 있다. 따라서 상대적으로 LIN 분포모형의 기댓값이 WBL 분포모형이나 GUM 분포모형보다 안정적으로 추정될 수 있음을 확인할 수 있었다.

Fig. 7. The Relative Error between the Monthly Mean Inflow of the Observation Data and the Theoretical Mean of Each Model for HC, CC, CP, and GS

5. 결 론

본 연구에서는 최근 그 사용이 증가하고 있는 Lindley 분포모형의 수문자료에 대한 적용성 검토를 수행하였다. 이를 위하여 한강수계에 위치하는 발전용댐 4곳을 선정하여 유입량 자료를 수집하고 가뭄빈도해석을 수행하였으며, 성능분석을 위해 기존에 사용되고 있는 분포모형들과 비교분석을 수행하였다. 또한 모형의 안정성 및 실용성을 검토하기 위하여 관측자료의 재현능력 평가를 수행하였다. 그 결과 Lindley 분포모형을 이용하여 수문자료(댐 유입량)를 활용한 가뭄빈도해석 등의 수문학적 적용성 검토를 확인하였으며, PDF, BIC, RMSE 등의 비교를 통해 기존 모형에 비해 정성적, 정량적으로 우수함을 확인하였다. 또한 우리나라 강우현상의 특징을 고려한 임의기간과 계절 등 시간적 개념에서 모형의 적용성 검토를 수행한 결과 기상현상에 따라 최적 분포모형이 달라질 수 있음을 확인하였으며, 실제 관측된 유입량 자료와의 비교를 통한 모형의 적합성 여부 검토결과, 기존의 분포모형보다 안정적인 결과를 보이는 것을 확인하는 동시에 모형에 따라 특정 시기별로 큰 오차를 보일 수 있으므로 공학적인 판단이 매우 중요하다는 점을 확인하였다. 이러한 결과가 나온 이유로 댐 유입량 자료는 해당 유역에서 내린 강우자료와의 상관성이 높다고 판단되어 지는데 일 강우자료의 경우 지수 분포모형과 감마 분포모형이 적절한 것으로 알려져 있다. 또한 기존 문헌조사에 따르면 우리나라 댐 유입량 자료에 대한 적용성 검토를 분포모형으로 감마 분포모형이 선정된 바가 있다(Jung et al., 2012). Lindley 분포모형의 경우 지수 분포모형과 감마 분포 모형의 혼합 구조 형식을 이루고 있기 때문에 기존의 다른 분포 모형들에 비해 댐 유입량 자료에 대해 좀 더 유연하게 적용되어 상대적으로 낮은 오차를 보이고 적합도가 높았던 것으로 추정된다. 이를 통해 앞으로 댐 유입량 자료에 대한 분석 수행 시 다양한 확률분포모형의 적용 및 여러 확률분포모형의 결합된 혼합분포모형(Mixture distribution)에 대한 적용성 검토가 필요할 것으로 판단된다.

References

- Adamowski, K. (1996). Nonparametric estimation of low-flow frequency. *Journal of Hydraulic Engineering*, Vol. 122, No. 1, pp. 46-49.
- Adhikari, T.R., and Srivastava, R.S. (2014). Poisson-size-biased Lindley distribution. *International Journal of Scientific and Research Publications*, Vol. 4, No. 1. Retrieved from <http://www.ijsrp.org/research-paper-0114.php?rp=P252183>
- Cho, K. (2013). *Development of potential drought hazard map of Korean peninsula using drought frequency analysis*. Ph.D. dissertation, Kyung Hee University.
- Ganora, D., and Laio, F. (2015). Hydrological applications of the Burr distribution: Practical method for parameter estimation. *Journal of Hydrological Engineering*, Vol. 20, No. 11, Article No. 04015024. doi:10.1061/(ASCE)HE.1943-5584.0001203
- Ghitany, M.E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, Vol. 78, No. 4, pp. 493-506.
- Ghitany, M.E., Alqallaf, F., Al-Mutairi, D.K., and Husain, H.A. (2011). A two-parameter weighted Lindley distribution and its applications to survival data. *Mathematics and Computers in Simulation*, Vol. 81, No. 6, pp. 1190-1201.
- Heo, J.H. (2016). *Statistical hydrology*. Seoul: Gumibook.
- Heo, J.H., Ahn, H., Shin, J.Y., Kjeldsen, T.R., and Jeong, C. (2019). Probability distributions for a quantile mapping technique for a bias correction of precipitation data: A case study to precipitation data under climate change. *Water*, Vol. 11, No. 7, 1475. doi:10.3390/w11071475
- Jeong, M., Park, S.-Y., Jang, H.-W., and Lee, J.-H. (2020). A study on derivation of drought severity-duration-frequency curve through a non-stationary frequency analysis. *Journal of Korea Water Resources Association*, Vol. 53, No. 2, pp. 107-119.
- Jung, Y., Nam, W.S., Shin, H., and Heo, J.H. (2012). A study on low-flow frequency analysis using dam inflow. *Journal of the Korean Society of Civil Engineering*, Vol. 32, No. 6B, pp. 363-371.
- Kang, K.S., and Seoh, B.H. (1995). Analysis of droughts for hydrological design of reservoirs at dam sites. *Proceeding of 1995 Conference*, Korean Society of Civil Engineering, Vol. 2, pp. 149-152.
- Kwak, J.W., Kim, D.G., Lee, J.S., and Kim, H.S. (2012). Hydrological drought analysis using Copula theory. *Journal of the Korean Society of Civil Engineering*, Vol. 32, No. 3B, pp. 161-168.
- Korea Water Resources Association (KWRA) (Ed.). (2009). *River design standard*. Seoul: KWRA.
- Lee, J.H., Jang, H.W., Kim, J.S., and Kim, T.W. (2015). Quantitative characterization of historical drought events in Korea: Focusing on drought frequency analysis in the five major basins. *Journal of Korea Water Resources Association*, Vol. 48, No. 12, pp. 1011-1021.
- Lindley, D.V. (1958). Fiducial distributions and Bayes'

- theorem. *Journal of the Royal Statistical Society, Series B*, Vol. 20, No. 1, pp. 102-107.
- Maeng, S.J., Hwang, J.H., and Shi, Q. (2009). Estimation of reservoir inflow using frequency analysis. *Journal of the Korean Society of Agricultural Engineers*, Vol. 51, No. 3, pp. 53-62.
- Ministry of Construction and Transportation (MOCT). (2000). *Research on the water resources management and development*.
- Nathan, R.J., and McMahon, T.A. (1990). Practical aspects of low-flow frequency analysis. *Water Resources Research*, Vol. 26, No. 9, pp. 2135-2141.
- Ryoo, K.S., Lee, H.G., Park, J.H., and Hur, Y.T. (2009). Improvement of estimation method on the low flow frequency inflow for the optimal reservoir operation. *Proceeding of 2009 Annual Conference*, Korea Water Resources Association, pp. 1287-1291.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, Vol. 6, No. 2, pp. 461-464.
- Seo, J.H., Shin, H.J., Ahn, H.J., and Heo, J.H. (2017). Applicability of the Burr XII distribution through dimensionless L-moment ratio of rainfall data in South Korea. *Journal of Korea Water Resources Association*, Vol. 50, No. 3, pp. 211-221.
- Shao, Q., Wong, H., Xia, J., and Ip, W. (2004). Models for extremes using the extended three-parameter Burr XII system with application to flood frequency analysis. *Hydrological Sciences Journal*, Vol. 49, No. 4, pp. 685-702.
- Tomy, L. (2018). A retrospective study on Lindley distribution. *Biometrics and Biostatistics International Journal*, Vol. 7, No. 3, pp. 163-169.
- Yu, J.S., Shin, J.Y., Kwon, M., and Kim, T.-W. (2017). Bivariate drought frequency analysis to evaluate water supply capacity of multi-purpose dams. *Journal of the Korean Society of Civil Engineers*, Vol. 37, No. 1, pp. 231-238.

Received	September 17, 2020
Revised	September 17, 2020
Accepted	October 5, 2020