

댐 수위의 확률분포 결정방법에 대한 분석

문 영 일(서울시립대학교 토목공학과 교수)

제 1 장 서 론

제 2 장 본 론

1절 기존 수위 확률분포 결정방법의 문제점

2절 비매개변수적 방법의 개념

3절 적용결과

3. 결 론

참 고 문 헌

제 1 장 서 론

현재 댐 위험도 분석 등의 사용되는 Monte Carlo Simulation 방법을 이용하는데 있어서 가장 어려운 문제이면서 약점으로 지적되어 온 것이 수리·수문학적 불확실성 변수들에 대해서 정확한 확률분포를 정의할 수 없다는 것이다. 즉 기존 매개변수적 방법들은 모집단 분포의 형태를 가정하고, 분포의 매개변수에 관한 통계적 분석을 하는 방법이다.

조사대상이 되는 모집단 분포에 관한 정보가 부족하기 때문에 어떤 가정이 곤란하거나 부정확한 가정을 전제로 하지 않고, 관측자료로 부터 통계량을 분석하여 통계적 추론을 하는 방법이 필요하게 되며, 이를 비매개변수적 방법(nonparametric methods)이라고 한다. 전통적인 매개변수적 방법에서 가장 어려운 점은 최종적으로 하나의 확률분포함수의 객관적인 선택이다. 또한 짧은 기록의 자료와 왜곡된 자료에 대한 매개변수의 신뢰성 결여와 그 외에 여러 가지 원인으로 인해서 복합분포(mixed distribution)의 밀도함수를 가진 자료가 나타낼 때 기존의 매개변수적 방법으론 어려움이 따른다.

특히 저수지의 수위와 같은 일정한 경계를 갖는(bounded) 변량에 대해서 정규분포 등과 같은 경계를 갖지 않는 확률분포형(unbounded probability distribution)을 적용하는 경우에는 댐의 높이를 초과하는 비현실적인 수위를 추정하는 오류를 범하게 되며 또한 기존의 경계를 갖는 확률분포형인 삼각형분포(triangle distribution), 균등분포(uniform distribution), 베타분포(beta distribution)등은 실제 변량의 확률분포를 왜곡되게 추정하게 되는 오류가 발생하게 된다.

이런 경우 상황에 따라 특성에 맞는 핵밀도함수(kernel density function)방법을 사용하면 상당히 좋은 결과를 얻을 수 있다. 비매개변수적 핵밀도함수 해석방법은 어떤 분포의 가정이 필요 없이 관측자료 자체에서 분포형을 유도할 수 있기 때문에 분포형 선정의 어려움을 해소할 수 있고 또한 관측자료의 적절한 분포형을 선정할 수 있다.

따라서 본 연구에서는 댐 위험도 분석시에 가장 민감한 변량으로 작용하는 저수지 수위에 대해서 기존의 경계를 갖는 확률분포형을 이용하는 매개변수적 방법의 문제점을 검토하고 이를 보완할 수 있는 경계를 갖는 핵밀도함수(boundary kernel density function)를 이용한 비매개변수적 방법을 적용하여 저수지의 적정 확률분포형을 산정하였다.

제 2 장 본 론

1절 기존 수위 확률분포 결정방법의 문제점

수리·수문학적 변량들은 각각 변량 자체의 고유한 통계학적인 특징을 가지고 있다. 그러나 기존의 위험도 분석에서 다루어지는 변량들에 대해서 통계적인 특징을 적절하게 반영하지 못하고 변량들을 기존의 확률분포형에 그대로 적용시켜 왔다. 즉, 수리·수문학적 변량인 강수량, 유출량, 풍속, 초기수위 등에 정규분포와 같은 경계를 갖지 않는 확률분포형(unbounded probability distribution)을 적용시켜 위험도 분석을 실시하였다. 일반적으로 정규분포나 경계를 갖지 않는 확률분포형(unbounded probability distribution)이 현실적으로 통계적인 특성을 반영하지 못함에도 불구하고 계속 사용되는 이유는 상대적으로 간단하고 소프트웨어로 많이 개발되어 사용자로 하여금 쉽게 사용이 가능하기 때문인 것으로 알려지고 있다(G. Lombardi, 2002).

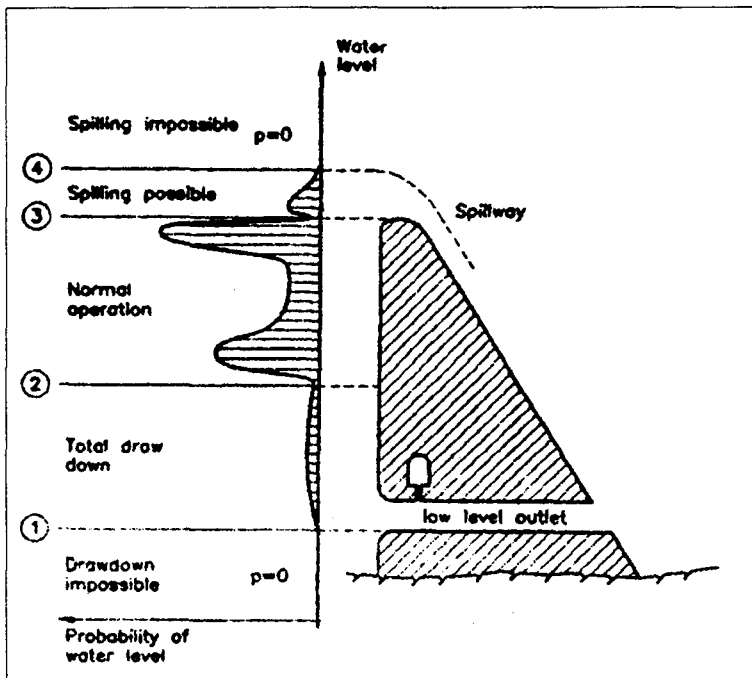


그림 1. 특정 저수지 수위의 확률분포 모식도

위의 그림 1은 어떤 특정댐의 일정한 경계값을 가지는 수위에 대한 확률분포를 나타내고 있으며 댐 위험도 해석시에 수리·수문학적 매개변수들의 확률분포형을 정규분포와 같은 경계를 갖지 않는 일반적인 확률분포형에 꼬리(tail) 부분의 영향 때문에 실제와는 너무나 다른 불확실한 위험도 해석이 이루어진다고 지적되었다(G. Lombardi, 2002).

즉 매개변수적인 모의발생기법을 통한 위험도 해석시에, 분포형의 꼬리(tail)부분의 영향으로 재현기간 10,000년, 100,000년의 댐 수위와 같은 아주 작은 확률을 가지는 값의 모의발생으로 결과적으로 계산되는 파괴확률은 비현실적인 경우가 많다. 그러나 경계를 갖는 확률분포형(bounded probability distribution)인 삼각형, 사각형, Beta분포인 경우에도 그림 2.A에서 그림 2.D 까지에서 보듯이 수리·수문학적 변량들의 특성을 정의하기에는 무리가 따른다. 즉 기존 매개변수적 방법을 사용하여 댐 위험도 해석을 실시하게 되면 2가지 가장 큰 문제점이 발생한다. 첫째는 댐 수위와 같이 일정한 경계를 갖는 변량에 대해서 매개변수적 방법을 사용함으로써 외삽시에 경계를 벗어나는 값을 발생시킬 수 있으며, 둘째로 강우량이나 홍수량, 풍속 등과 같은 변량에 특정 분포형을 가정함으로써 실제 변량에 특성을 정확히 반영하기 힘들다는 것이다. 또한, 수리·수문학적 변량들이 여러 가지 불확실한 원인으로 인하여 복합분포(mixed distribution) 형태를 가질 때, 기존 매개변수적 해석방법은 bimodal을 갖는 확률밀도함수를 해석하는데는 여러 가지 어려움이 따른다. 또한 짧은 기록의 자료와 왜곡된 자료가 포함된 경우에도 매개변수의 신뢰도가 떨어져 관측자료의 분포특성을 정확히 설명하는데 어려움이 따른다. 예를 들면 아래 그림 2.A ~ 그림 2.D에서 실제 관측된 자료가 그림 2.A와 같을 때 그림 2.B ~ 그림 2.D와 같이 삼각형분포·정규분포·균등분포 등과 같이 가정을 하면 관측자료의 통계적인 특징을 적절하게 반영하지 못한다 할 수 있다.

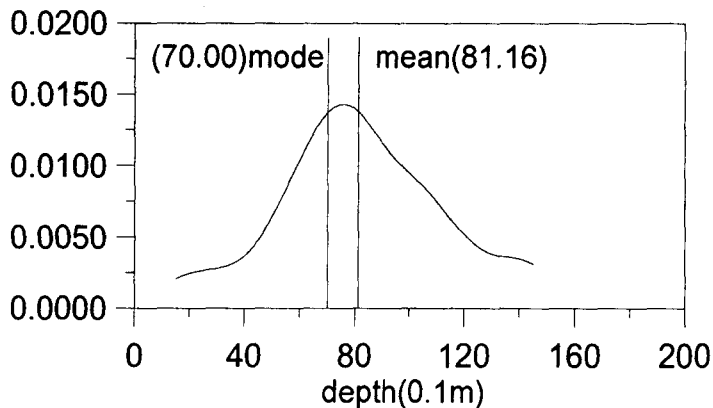


그림 2.A. 실측자료

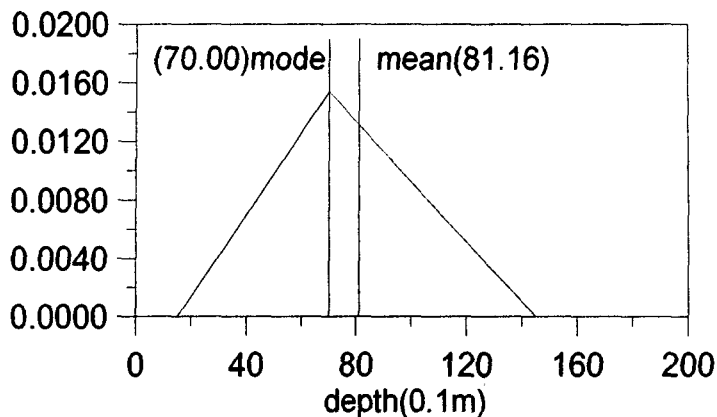


그림 2.B. 삼각형 분포(triangle distribution)

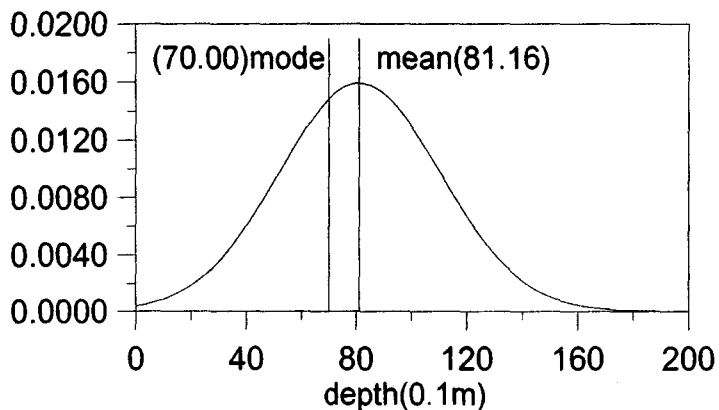


그림 2.C. 정규분포(normal distribution)

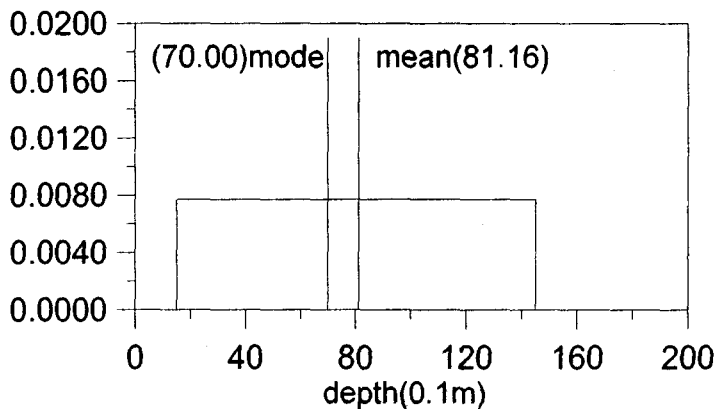


그림 2.D. 균등분포(uniform distribution)

2절 비매개변수적 방법의 개념

수많은 수문학자(hydrologists)와 통계학자들은 홍수나 강우 자료에 완전하게 부합되는 확률 분포를 찾으려고 하고 있다. 그러나 불행하게도 자연사상에 대해서 정확한 확률분포를 찾는다는 것은 불가능하며 단지 자연사상을 표현하기 위한 수단으로 인간이 만들어낸 확률분포형에 적용하고 있으며, 따라서 적용되고 있는 모든 분포형은 High Extreme Tail에 면에서 모든 잘못된 것이다(Yen, 1999). 적용된 분포형은 재현기간을 가지는 무한대의 값을 나타내지만 지구상의 자연현상은 유한한 값을 가진다.

조사대상이 되는 모집단 분포에 관한 정보가 부족하기 때문에 어떤 가정이 곤란하거나, 부정확한 가정을 전제로 하지 않고 통계량을 분석하여 통계적 추론을 하는 방법이 필요하게 되며, 이를 비매개변수적 방법(nonparametric methods)이라고 한다. 전통적인 모수적 방법에서 가장 어려운 점은 최종적으로 하나의 확률분포함수의 객관적인 선택이다. 또한 짧은 기록의 자료와 왜곡된 자료에 대한 매개변수의 신뢰성 결여와 그 외에 여러 가지 원인으로 인해서 복합분포(mixed distribution)의 밀도함수를 가진 자료가 나타낼 때 기존의 모수적 방법으론 어려움이 따른다. 이런 경우 비매개변수적 핵밀도함수(kernel density function)방법을 사용하면 상당히 좋은 결과를 얻을 수 있다는 연구결과가 10 여년 전부터 많이 발표되었다(Lall 등, 1993; Moon 등, 1993; Moon과 Lall, 1994; Adamowski, 1996; Moon, 2000; Moon 등, 2001). 비매개변수적 핵밀도함수 해석방법은 어떤 분포의 가정이 필요 없이 관측자료 자체에서 분포형을 유도할 수 있기 때문에 분포형 선정의 어려움을 해소할 수 있고 또한 관측자료의 적절한 분포형을 선정할 수 있다.

비매개변수적 확률밀도함수 추정법의 기본 개념은 히스토그램에서 출발했다 할 수 있다. 히스토그램은 가장 오래되고 쉽게 사용되는 확률밀도함수 추정 법으로 알려지고 있다. 그러나 히스토그램의 단점은 계급구간이 변화하는 점에서 불연속적이고 구간간격과 시작점의 선택에 따라 확률밀도함수의 모양이 달라지는 것이다. 이와 같은 히스토그램의 단점을 보완하여 비매개변수적확률밀도함수 추정식을 모든 실수 x 에 대하여 식 1과 같이 정의하였다(Silverman, 1986).

$$f(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} K\left(\frac{x - X_i}{h}\right) \quad (\text{식 1})$$

여기에서 X_1, X_2, \dots, X_n 은 독립적으로 동일하게 분포된 실 관측치를 나타내며, $K(\cdot)$ 는 핵

함수(kernel function)이고 h 는 n 이 무한대로 갈 때 0으로 접근하는 값을 갖는 양의 광역폭 (bandwidth)이다. 자료가 발생된 위치에 히스토그램의 box 대신 핵함수를 자료의 개수만큼 위치하도록 하여 구간을 이동시킬 수 있는 이동 히스토그램이 핵밀도함수 추정식의 기본개념이다. 즉, 각 관측치에 폭 $2h$ 인 핵함수 $K(\cdot)$ 를 box대신 각 관측치에 놓고, 이 들을 합산하므로 위 식이 만들어진다. 대표적인 핵함수는 아래 표 1과 같고 일반적으로 핵함수는 다음조건을 만족한다.

$$\int K(t)dt = 1, \text{ 여기서 } t = \frac{x - X_i}{h} \quad (\text{식 2})$$

그림 3은 관측된 자료위치에 핵함수를 각각 위치하여 합산한 후 확률밀도함수로 표현하였다. 그림 3에서 관측자료의 개수 10, 광역폭 h 가 1인 핵함수를 사용한 핵밀도함수를 보이고 있다. 그림 4는 여러 가지 핵함수의 형태를 나타내고 있다.

표 1. 여러 핵함수의 종류와 식

Kernel	$K(t)$
Rectangular	$\frac{1}{2}$ for $ t < 1$ if otherwise 0
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$
Epanechnikov	$\frac{3}{4} \left(1 - \frac{1}{5}t^2\right) / \sqrt{5}$ for $ t < \sqrt{5}$
Rajagopalan	$\frac{3h}{(1-4h^2)} (1-t^2)$, 여기서 $ t \leq 1$
Cauchy	$\frac{1}{\pi(1+t^2)}$

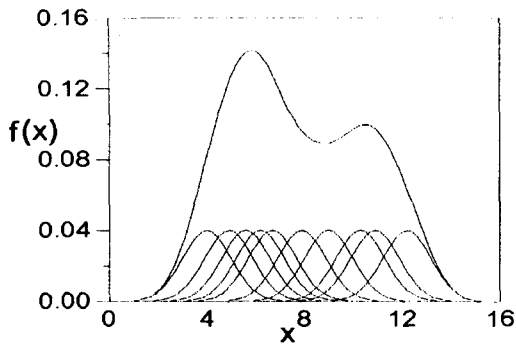


그림 3. 핵밀도함수(kernel density function)

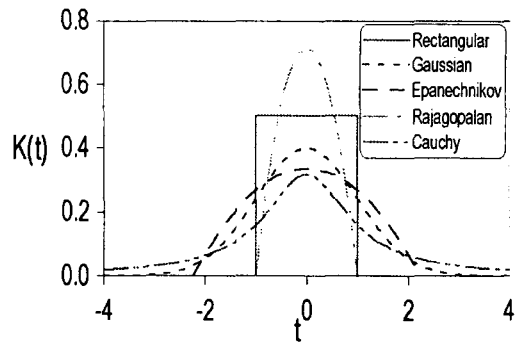


그림 4. 여러 가지 핵함수

비매개변수적 핵밀도 함수법에서 광역폭(bandwidth) h 의 선택은 매우 중요한 문제로 다루어져 왔다. 지금까지의 광역폭을 선택하는 방법은 maximum likelihood, least squares cross validation 또는 Adamowski Criterion, Breiman method(Adamowski, 1985; Lall 등, 1993; Moon 등, 1994)등이 있다.

Maximum Likelihood cross-validation은 Habbema 등(1974) 과 Duin(1976)은 우도(likelihood)의 개념을 사용하여 광역폭 h 를 선택하는 방법을 제시하였다. 제시된 maximum likelihood cross-validation 법은 다음 식 3의 $CV(h)$ 가 최대로 될 때의 h 값을 선택하는 것이다.

$$CV(h) = \frac{1}{n} \sum_{i=1}^n \log \hat{f}_{-i}(X_i) \tag{식 3}$$

여기서 $\hat{f}_{-i}(x_i)$ 는 관측자료 X_i 가 생략된 관측자료로부터의 밀도추정값이다.

$$\hat{f}_{-i}(x) = \frac{1}{(n-1)h} \sum_{j \neq i} K\left(\frac{x - X_j}{h}\right) \tag{식 4}$$

Least squares cross-validation은 Rudemo (1982)와 Bowman (1984)에 의해 제안된 방법으로 다음과 같이 정의되는 함수 R 을 최소로 하는 광역폭 h 가 최적의 광역폭이라 정의하였다.

$$R(f) = \int \hat{f}^2 - 2 \int f \hat{f} \tag{식 5}$$

주어진 임의의 확률밀도함수를 f , 이에 대한 추정치를 \hat{f} 라고 하면 제곱오차의 합은 다음과 같이 표현된다.

Breiman 등(1977)은 변동 핵밀도함수 방법을 다음과 같이 제안하였다.

$$\hat{f}(x) = \frac{1}{n} \sum_{j=1}^n \frac{1}{h_k d_{j,k}} K\left(\frac{x - X_j}{h_k d_{j,k}}\right) \quad (\text{식 6})$$

여기서, $d_{j,k}$ 는 한 개의 자료 x_j 에서 그 나머지 자료 $(n-1)$ 개 중에서 k 번째로 가까운 지점에 있는 자료까지의 거리이고, h_k 는 광역폭으로서 상수이다. 따라서, 고정 핵밀도함수 방법에서 고정광역폭 h 는 국부적인 자료의 밀도값에 의존하여 조정되어지는 변동광역폭의 $h_k d_{j,k}$ 로 대체되어지는데, Breiman 등(1977)은 다음과 같은 분산값이 최소가 되는 수치적인 방법을 제안하였다.

$$S = \sum_{j=1}^n \left(W_j - \frac{1}{n}\right)^2 \quad (\text{식 7})$$

여기서 $W_j = \exp(-\hat{f}(x_j) n d_{j,1})$ $j=1, 2, \dots, n$ 을 나타낸다.

Adamowski Criterion은 Adamowski(1985)는 광역폭 h 의 수치적 근사값을 다음과 같은 MSE의 식 8을 이용하여 이 값이 최소가 되는 광역폭 선택방법을 제시하였다.

$$\sum_{j=1}^n [F_j(x) - F(x_j)]^2 \quad (\text{식 8})$$

여기서 미지의 확률 $F_j(x)$ 는 확률도시공식 $F(x_j)$ 로 추정될 수 있다.

비매개변수적방법을 이용하여 해당 변량에 대한 발생확률에 따른 분위값(quantile)을 추정하기 위해서는 핵밀도함수(kernel probability density function)를 직접 적분하여 누가핵분포함수(cumulative kernel distribution function)의 분위값을 직접 구하는 방법과 관측값과 이에 대응하는 경험적인 발생확률을 이용하여 회귀추정식을 이용하는 두 가지 방법이 있다.

2.1 직접적분법

비매개변수적 확률밀도함수 추정식을 모든 실수 x 에 대하여 식 1과 같이 정의될 수 있으며, 누가분포함수 $F(\cdot)$ 에 상응하는 분위값(quantile)은 다음 식 10과 같이 확률밀도함수를 직접 적분하여 계산할 수 있다.

$$F(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (\text{식 9})$$

여기서,

$$K(t) = \int_{-\infty}^t k(u) du \quad (\text{식 10})$$

만약 p 를 누가분포함수 $F(\cdot)$ 에 대응하는 확률이라 나타내고, $F^{-1}(\cdot)$ 를 누가분포함수 $F(\cdot)$ 의 역함수라 하면 모든 실수 x 에 대해서 분위값을 다음 (식 11)과 같이 나타낼 수 있다.

$$x = F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\}, \quad p \in (0, 1) \quad (\text{식 11})$$

$F^{-1}(p)$ 의 비매개변수적 추정자(estimator)의 근사적인 분산(variance)이 일반적으로 $\sigma^2 = (F^{-1})'(p)p(1-p)$ 로 주어지기 때문에 확률 p 의 분위값 $F^{-1}(p)$ 값은 quantile-밀도함수 $\partial F^{-1}(p)/\partial p = (F^{-1})'(p)$ 와 밀접하게 관련을 가지고 있다.

F_n 을 표본크기 n 을 가지는 경험적인(empirical) 분포함수라 정의하면, 표본 p -분위값 $F_n^{-1}(p)$ 에 대응하는 분위값 x 에 대한 Kernel Quantile 추정자 \hat{x}_n 은 다음 식 12와 같이 나타낼 수 있다(Falk, 1986).

$$\hat{x}_n(p) = \int_0^1 F_n^{-1}(x) \alpha_n^{-1} / k((p-x)/\alpha_n) dx \quad (\text{식 12})$$

여기서, $k(\cdot)$ 는 핵함수(kernel function)를 α_n 은 광역폭(bandwidth)을 나타낸다.

2.2 회귀 추정식

회귀추정식의 기본이 되는 함수는 관측 자료의 경험적 발생확률값 p_i 에 대응되는 관측값 y_i 로 결정되어지며 추정되는 함수의 분위값(quantile) $x(p)$ 는 Gasser-Muller(1984)가 식 13과 같이 제시한 핵함수를 이용한 회귀추정식으로부터 추정할 수 있다. 이는 하나의 가중함수인 핵함수를 이용하여 경험적 빈도해석 함수의 회전형(convolution)을 고려한 것이다(Moon과 Lall, 1994).

$$\hat{x}(p) = \sum_{i=1}^n \frac{1}{h} \int_{s_{i-1}}^{s_i} y_i k\left(\frac{p-u}{h}\right) du = \sum_{i=1}^n \frac{1}{h} y_i \int_{s_{i-1}}^{s_i} k\left(\frac{p-u}{h}\right) du \quad (\text{식 13})$$

여기서 h 는 점 p 와 관련된 대역폭(bandwidth), $k(\cdot)$ 는 핵함수(kernel function)이며, $s_i = (p_i + p_{i+1})/2$, ($i = 1, \dots, n-1$), $s_0 = 0$, $s_n = 1$ 이다. 이 때 p 는 초과확률을 나타내는 구간 $[0,1]$ 에서의 임의의 값이다.

또는, Sheather-Marron(1990)은 실관측치를 크기 순으로 나열한 자료($X_1 \leq X_2, \dots \leq X_n$)를 이용하여 근사적으로 다음 4가지의 분위값 x_p 의 산정식을 제시하였다.

$$\hat{x}_{p,1} = \sum_{i=1}^n [n^{-1}k_h(i/n - p)]X_{(i)} \quad (\text{식 14})$$

$$\hat{x}_{p,2} = \sum_{i=1}^n [n^{-1}k_h((i-1/2)/n - p)]X_{(i)} \quad (\text{식 15})$$

$$\hat{x}_{p,3} = \sum_{i=1}^n [n^{-1}k_h(i/(n+1) - p)]X_{(i)} \quad (\text{식 16})$$

$$\hat{x}_{p,4} = \sum_{i=1}^n k_h\left(\frac{i-\frac{1}{2}}{n} - p\right)X_{(i)} / \sum_{j=1}^n k_h\left(\frac{j-\frac{1}{2}}{n} - p\right) \quad (\text{식 17})$$

여기서, $k_h(\cdot) = h^{-1}k(\cdot/h)$ 를 나타낸다.

3절 적용결과

특정댐이 수위에 확률밀도함수를 비교하기 위해서 댐의 내구년수가 길고 비교적 수문자료가 잘 갖추어진 소양강댐 연최대치 수위자료를 바탕으로 모의를 실시하였다. 표 2는 소양강댐의 수위를 나타내며 표 3은 수위자료에 대한 통계치를 나타낸다.

표 2. 소양강댐 연최대수위 및 최저수위 (단위 : EL.m)

연 도	최고수위	최저수위
1974	178.92	160.98
1975	185.38	153.87
1976	182.92	157.42
1977	173.19	159.45
1978	188.41	151.93
1979	189.59	172.84
1980	190.26	170.19
1981	194.03	173.76
1982	185.31	165.36
1983	185.56	169.02
1984	197.79	160.56
1985	186.98	165.93
1986	187.56	168.42
1987	193.02	174.41
1988	182.24	164.42
1989	186.41	162.15
1990	197.99	173.60
1991	184.66	164.39
1992	185.14	168.62
1993	185.90	171.37
1994	176.30	158.90
1995	194.66	160.34
1996	177.83	160.39
1997	177.37	156.10
1998	189.96	164.93
1999	191.43	161.48
2000	187.08	160.74

소양강댐의 수위는 평균값이 186.5 EL.m이며 최소값은 173.19 EL.m 최대값은 197.99 EL.m를 나타내고 있다. 연최대치자료의 평균값은 186.5 EL.m는 댐의 홍수기 제한수위를 밀도는 값이며 대략 여수로 높이에 해당하는 수위를 나타내며 비교적 낮은 수위를 나타내고 있다.

표 3. 소양강댐 연최대수위 통계치

댐 수위	MEAN	STANDARD DEV.	COEFF. VARIATION	COEFF. SKEWNESS	MIN	MAX
Annual Peak	186.5	6.29	0.034	- 0.157	173.19	197.99

그림 5는 소양강댐의 제원을 나타낸다. 그림 5에서와 같이 댐의 수위는 저수위 150 EL.m부터 댐의 마루(crest)까지의 높이 203 EL.m 범위를 갖고 있다.

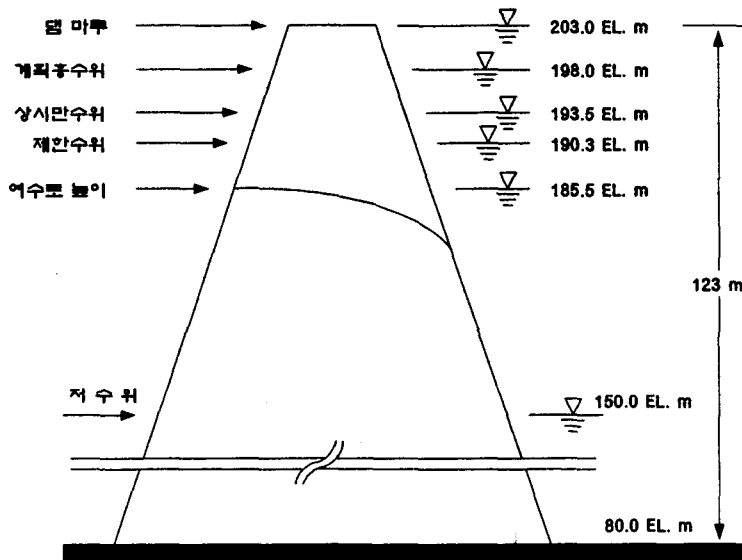


그림 5. 소양강댐의 제원

표 2에 연최대치 자료를 바탕으로 경계를 갖지 않는 대표적인 확률분포형을 적용시켜 보았다. 적용결과는 그림 6과 같다. 아래 그림 6에서 보듯이 매개변수적 방법은 매개변수 추정방법인 모멘트법을 이용한 확률밀도함수를 나타내며, 분포형마다 모두 다른 결과를 보여주고 있다. 댐 수위에 경우에 댐 정상적인 운영 조건하에서 일정한 범위가 주어지 있지만 기존의 매개변수적 방법을 사용하여 수위값을 추정할 경우에 이 범위를 초과하는 값을 나타낼 수 있기 때문에 불합리하다. 그림 5에서 보듯이 소양강댐은 댐의 저수위 150 EL.m를 하한치로 해서 댐 정부(crest)의 높이 203 EL.m를 상한치를 갖는 값을 나타내야 한다. 하지만 그림 6에서 보듯이 경계를 갖지 않는 분포형을 댐의 수위에 적용시킨 결과 이 범위를 초과하는 비현실적인 값들을 나타내게 된다.

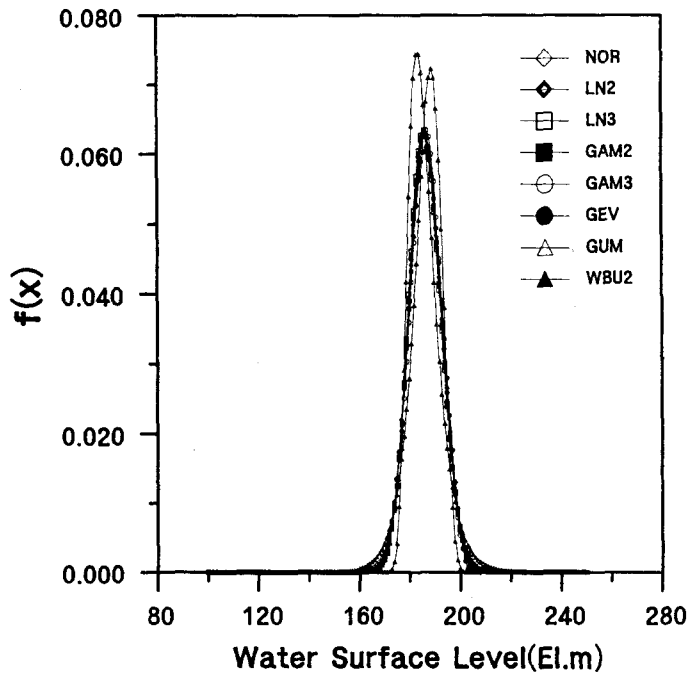


그림 6. 소양댐 수위의 매개변수적 확률밀도함수 적용결과

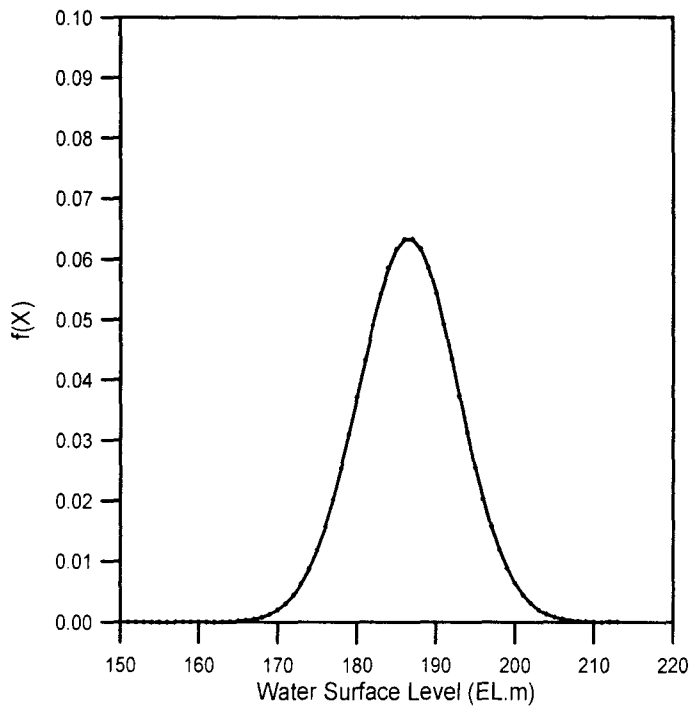


그림 7. 정규분포를 적용한 결과

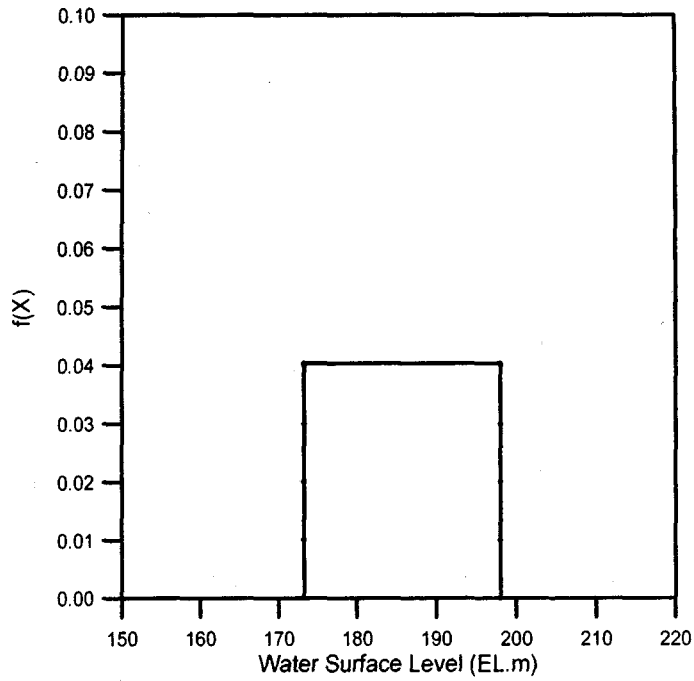


그림 8. 균등분포를 적용한 결과

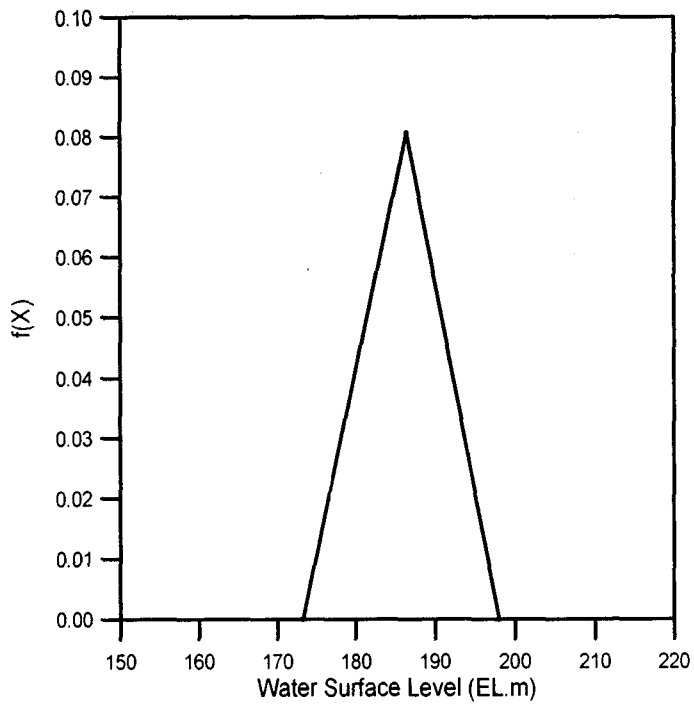


그림 9. 삼각형분포 적용 결과

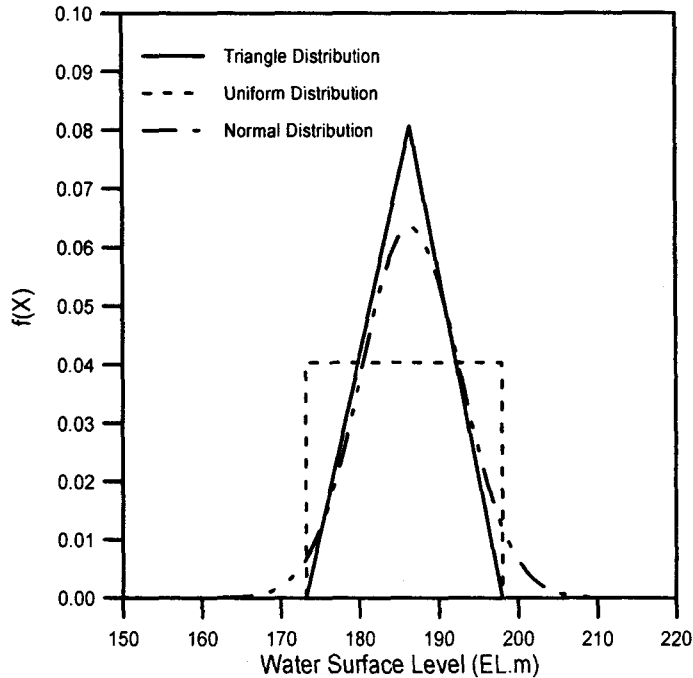


그림 10. 정규분포, 균등분포, 삼각형분포 적용 결과

그림 7에서 그림 10까지는 일반적으로 수위에 확률밀도함수 추정시에 가장 많이 사용되는 정규분포, 균등분포, 삼각형분포를 소양강 수위에 대해서 적용시킨 결과를 나타낸다.

그림 7에서 보듯이 정규분포는 댐 수위의 한계를 넘어서는 값의 확률값이 부여되는 불합리한 결과를 나타내었고, 그림 8과 그림 9에서 보듯이 균등분포와 삼각형분포는 기존 댐수위에 최대값과 최소값을 기준으로 경계를 가지고 확률분포를 나타내지만 너무 이상화시키는 경향이 있어 정확한 확률분포를 표현할 수 없다.

그림 11은 소양강댐 지점 수위에 대해서 비매개변수적 방법을 적용한 확률밀도함수를 나타내고 있다. 비매개변수적 방법의 경우에는 일정한 범위 내에서 변량의 통계특성을 유지하면서 댐수위에 범위를 초과하지 않는 범위내에서 변량에 대한 확률값이 부여되고 있다.

이는 이상홍수위를 Critical Level로 하고 Bounded Kernel을 사용한 결과 댐의 높이를 초과하지 않는 범위 내에서 모의가 가능하였다. 또한 그림 ?에서 보듯이 비매개변수적 방법은 bimodal의 형태를 갖는 변량에 대해서도 비교적 좋은 결과를 보여주고 있으며 하나의 분포형으로 제시하고 있다. 원자료와 모의된 자료의 확률분포를 비교하기 위해서 원자료를 바탕으로 비매개변수적으로 Monte Carlo Simulation을 실시하여 모의 발생된 자료계열의 확률밀도함수를 <그림 8.13>에 나타내었으며 원자료와 확률분포와 범위(range)가 유사한 확률분포를 나타내고 있다.

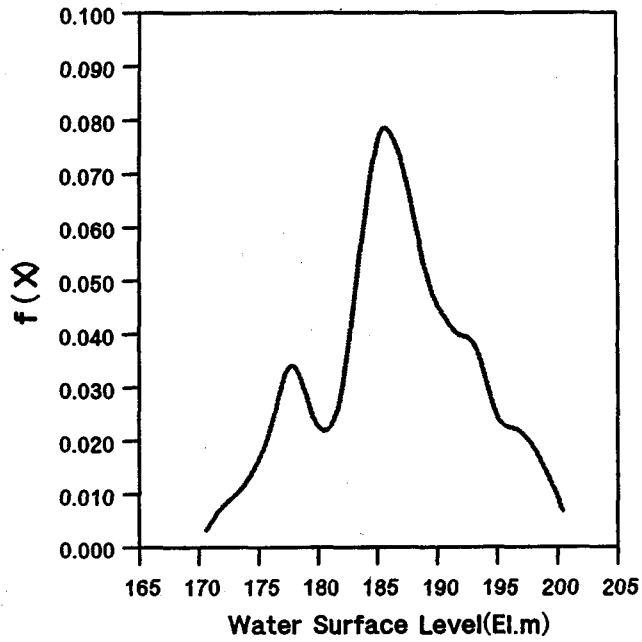


그림 11. 비매개변수적 방법의 적용 결과(원자료)

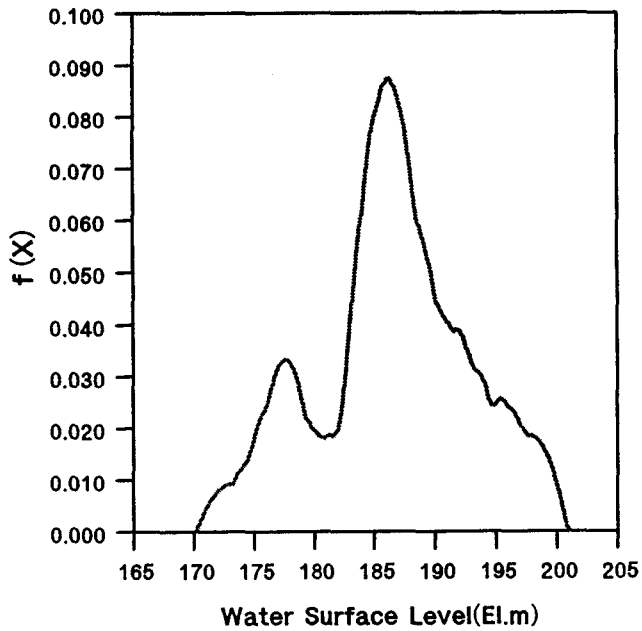


그림 12. 비매개변수적 적용 결과(모의결과)

그림 13은 경계를 갖는 확률분포형인 삼각형분포, 균등분포와 경계를 갖지 않는 분포형중 가장 많이 사용되는 정규분포와 boundary 핵밀도함수를 이용한 비매개변수적 방법을 나타내며 그림에서 보듯이 실제 소양강댐 수위와는 모두 왜곡된 확률밀도 함수를 보여주고 있다.

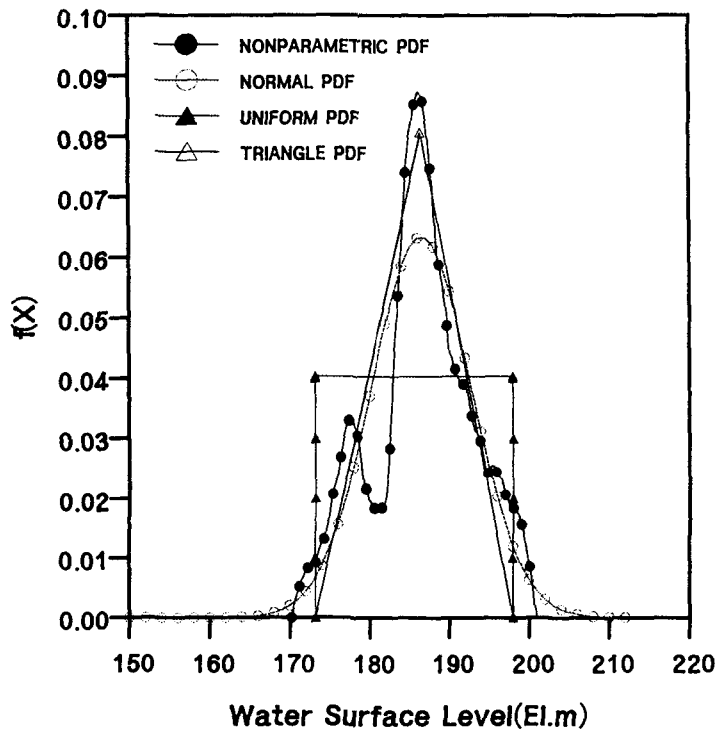


그림 13. 매개변수적 방법과 비매개변수적 방법 적용 결과

제 3 장 결 론

현재 댐 위험도 분석 등의 사용되는 Monte Carlo Simulation 방법을 이용하는데 있어서 가장 어려운 문제이면서 약점으로 지적되어 온 것이 수리·수문학적 불확실성 변수들에 대해서 정확한 확률분포를 정의할 수 없다는 것이다. 특히 댐 수위와 같이 자료의 특성을 충분히 반영하면서 댐마루(dam crest)의 높이를 초과하지 않으면서 경계를 갖는 분포형을 추정하는 것은 무엇보다 중요하다 하다. 그러나 기존의 매개변수적 방법으로는 이러한 문제점을 적절히 반영할 수 없다. 즉 기존의 경계를 갖는 삼각형분포, 균등분포, beta분포 등은 변량에 대한 통계 특성을 반영하지 못하고 이상화시키는 단점이 있으며 경계를 갖지 않는 정규분포와 같은 분포형은 일정 범위를 초과하는 비현실적인 확률분포의 추정하게 되는 문제점이 발생할 수 있다.

이런 경우 경계를 갖는 핵밀도함수(kernel density function)방법을 사용하면 상당히 좋은 결과를 얻을 수 있으며 어떤 분포의 가정이 필요 없이 관측자료 자체에서 분포형을 유도할 수 있기 때문에 분포형 선정의 어려움을 해소할 수 있고 또한 관측자료의 적절한 분포형을 선정할 수 있다.

특히 댐수위와 같은 변량에 대해서는 Epanechnikov 핵함수와 같은 경계를 갖는 핵함수를 설정하면 좋은 결과를 얻을 수 있다.

비매개변수적 방법을 소양강댐 수위에 대해서 적용결과 기존 매개변수적 방법에 비해서 비교적 변량의 통계특성을 잘 반영하며 발생 가능한 댐수위의 범위를 초과하지 않게 확률밀도 함수를 구성할 수 있었다.

참 고 문 헌

- 문영일, 신은우, 김동권, 권현한(2002), "비매개변수적 Monte Carlo Simulation에 의한 댐 위험도 분석", 한국수자원학회 학술발표회.
- 권현한, 문영일(2002), "댐 위험도 분석을 위한 실행함수 설정방안 연구", 대한토목학회 학술발표회.
- 신창건, 권현한, 신은우, 문영일(2003), "댐 위험도 분석과 대책방안" 한국수자원학회 학술발표회.
- 권현한, 문영일, 차영일, 전시영(2003), "비 모수적 모의기법을 이용한 댐의 월류확률 산정에 관한 연구", 대한토목학회 학술발표회.
- Adamowski, K., Nonparametric kernel estimation of flood frequency. *Water Resources Research* 21(11):1585-1590, 1985.
- Bowman, A. W., "A comparative study of some kernel-based nonparametric density estimators", *Journal of Statistical Computing and Simulation*, Vol. 21, pp. 313~327, 1985.
- Breiman, L., W. Meisel, and E. Purcell., "Variable kernel estimates of multivariate densities", *Technometrics*, Vol. 19, No. 2, pp. 135~144, 1977.
- Duin, R. P. W., On the choice of smoothing parameters for parzen estimators of probability density functions. *IEEE Trans. Comput.* C-25:1175-1179, 1976.
- Falk, Michael., On the Estimation of The Quantile Density Function, *Statistics & Probability Letters* 4, pp. 69-73, 1986.
- Gasser, T. and Muller, H. G., Estimating regression functions and their derivatives by the kernel method. *Scandinavian Journal of Statistics* 11, 171-185, 1984.
- Habbema, J. D. F., J. Hermans, and V. D. Broek., A stepwise discrimination program using density estimation, p. 100 - 110. In G. Bruckman (Ed.). *Physical verlag. Compstat Vienna*, 1974.
- Lall, U., Moon, Young-Il, and Bosworth, K., "Kernel flood frequency estimators: bandwidth selection and kernel choice", *Water Resources Research*, Vol. 29, No. 4, pp. 1003-1015, 1993.
- Lombardi, G., Preparatory works and regular maintenance. *Dam Maintenance and Rehabilitation*, 7-15, 2002.
- Moon, Young-Il, Lall, U., and Bosworth, K., "A comparison of tail probability estimators", *Journal of Hydrology*, Vol. 151, pp. 343-363, 1993.
- Moon, Young-Il. and Lall, U., Kernel Quantile Function Estimator for Flood Frequency Analysis, *Water Resources Research* 30(11), pp. 3095-3103, 1994.

- Moon, Young-II. and Kwon, Hyun-Han., Nonparametric Dam Risk Analysis for Dam Rehabilitation in the South Korea. *Dam Maintenance and Rehabilitation*, 119-124, 2002.
- Rudemo, M., Empirical choice of histograms and kernel density estimators. *Scand. J. Statist.* 9:65-78, 1982.
- Sheather, S. J. and Marron, J. S., Kernel Quantile Estimators, *Journal of American Statistical* Vol. 85, pp.410-416, 1990.
- Silverman, B. W., "Density estimation for statistics and data analysis." New York: Chapman and Hall, 1986.
- Yen, B. C, Risk analysis in dam safety assessment proceedings of the International Workshop on Risk Analysis in Dam Safety Assessment, held at National Taiwan University, Taipei, Taiwan, June 14-15, 1999.