

최우도 모형을 이용한 수위-유량곡선식 개발

김경훈^{1*} · 박준일¹ · 신찬기²

국립경상대학교 토목공학과¹ · 국립환경과학원 낙동강물환경연구소²

Development of Rating Curves Using a Maximum Likelihood Model

Gyeong-hoon Kim^{1*} · Jun-il Park¹ · Chan-ki Shin²

¹Department of Civil Engineering, Gyeongsang National University

²Nakdong River Environment Research Center National Institute of Environmental Research
Pyeong-ri, Dasan-myeon, Goryeong-gun, Gyeongsangbuk-do, 717-873, Korea

Abstract

The non-linear least squares model(NLSM) has long been the standard technique used by hydrologists for constructing rating curves. The reasons for its adaptation are vague, and its appropriateness as a method of describing discharge measurement uncertainty has not been well investigated. It is shown in this paper that the classical method of NLSM can model only a very limited class of variance heterogeneity. Furthermore, this lack of flexibility often leads to unaccounted heteroscedasticity, resulting in dubious values for the rating curve parameters and estimated discharge.

By introducing a heteroscedastic maximum likelihood model(HMLM), the variance heterogeneity is treated more generally. The maximum likelihood model stabilises the variance better than the NLSM approach, and thus is a more robust and appropriate way to fit a rating curve to a set of discharge measurements.

Key words : Rating curve, Heteroscedasticity, Non-linear least square model, Maximum likelihood model.

* Corresponding author E-mail : dahong@me.go.kr

I. 서론

수자원의 확보 및 효율적인 관리는 수자원의 정확한 조사에서 시작된다. 이를 위해 여러 나라에서 강수, 하천 수위 및 유량, 유사량, 증발과 증산, 토양수분 등의 정확한 측정에 많은 노력을 기울이고 있다(김원 등, 2004). 이중 하천유량 자료는 수문모형 설계 시 필수적인 기초자료이다. 특히 홍수와 가뭄관리, 2004년 수량과 수질을 연계한 오염총량관리제도 실시로 오염부하량 산정을 위한 기준유량설정 등과 같은 공공성에 기반을 둔 각종 국가사업과 기초연구에 반드시 필요한 기초자료이다. 따라서 적절

한 수위-유량곡선식을 개발하여 하천유량을 추정하는 것은 정확한 수문자료 생산이라는 측면에서 매우 중요하다(건교부; 2005, 김경훈 등; 2005).

일반적으로 하천의 수위관측소에서 수위와 유량사이의 관계는 $Q=a(w+c)^b$ 이며, 여기서 Q 는 유량, w 는 기준수위 그리고 a , b 그리고 c 는 매개변수이다. 이러한 수리학적 모티브를 제공하고 실무에 적용한 사람은 Lambie(1978)이고, Ackers et al.(1978)은 이러한 관계가 이론적으로 성립된다는 것을 증명하였다. 또한 수위-유량간의 여러 모형들도 Gawne and Simonovic(1994) 그리고 Yu(2000)들에 의해 연구되었다.

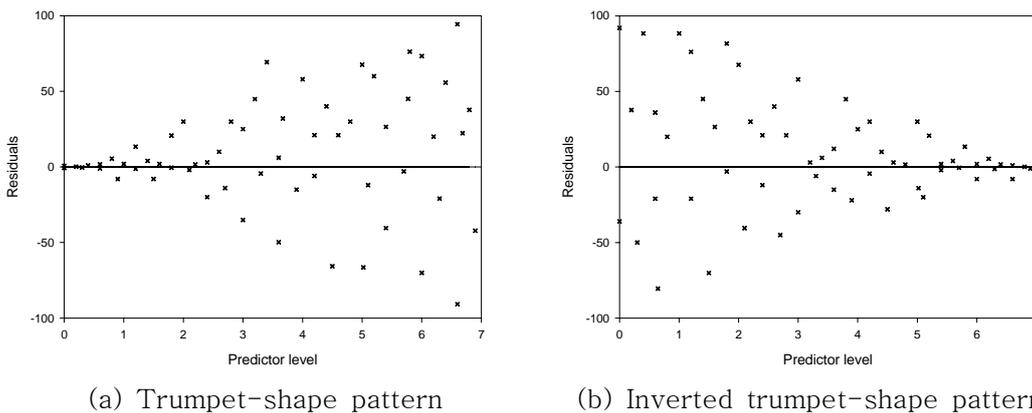


Fig. 1. Example of regression residuals by heteroscedasticity

기존의 수위-유량곡선식 개발은 가용한 기준수위별 실측 유량자료를 이용, 특정 수위범위 내에서 적용 가능한 곡선식을 비선형최소제곱모형(non-linear least squares model ; NL SM)을 이용하여 개발하고 있다. 추정된 곡선식에서 발생하는 오차항은 등분산성(homo-scedasticity)이라고 가정하고 모든 측정치들은 곡선 개발 시 동일한 가중치를 부여하여 개발하게 된다. 현재까지 개발된 지점별 수위-유량곡선식이 이러한 가정에 대한 충분한 검토 없이 결정계수

(coefficient of determination; r^2)값이 1에 가장 근접하는 회귀곡선식을 선정 이용하고 있는 실정이다.

이분산성은 보통 증가 또는 감소하는 형태로 나타나는데, 증가형(trumpet shape)은 추정치가 적은 모수들에 의해 개발되었을 경우 발생하며, 감소형(inverted trumpet shape)은 너무 많은 모수들에 의해 추정치가 개발되었을 경우 발생한다(Fig. 1(a), (b))(Asgeir, 2004). 실질적으로 분석에 자료군이 너무 적어 Fig. 1과 같이 뚜렷한 이

분산성을 발견하기가 현실적으로 불가능한 경우도 발생하겠지만, 모형 개발 시 분석의 효율성과 결과의 신뢰성 향상을 위해 이분산성에 대한 사전 검토가 우선되어야 할 것으로 생각된다.

대개 통계학자들은 분산의 이질성을 가진 자료군을 적절한 변환 없이 회귀분석을 실시하였을 경우, 개발된 곡선식에 부적절한 영향을 미치므로 이를 적절히 제거하여 분석할 것을 추천하고 있다. 그러나 대부분의 수문학자들이 이에 대한 전문적인 통계학적 지식이 부족할 뿐만 아니라, 수문관측 지침서의 곡선 개발 시 이 부분이 주요 검토 요인임에도 불구하고 이에 대한 설명이 자세하게 되어 있지 못한 것도 분석 시 오류를 범하는 한 요인으로 작용하고 있다. 따라서 이를 유량측정 시 기준수위에 비례하여 발생하는 측정오차(measurement error)로 간주하거나 측정 시 발생하는 불확실성에 의해 각 지점별로 다르게 발생하는 것으로 통상 분석을 실시하는 경우가 대부분이다. 따라서 이분산성은 측정 자료를 이용한 수위-유량곡선 개발 시 적절하게 처리되지 못하고 있는 실정이며, 이와 같은 조건에서 개발된 기존 NLSM 곡선식들은 품질이 저하되어 부정확한 추정치를 제공하는 결과를 초래하게 될 것이다.

Lambie(1978)은 NLSM 모형을 이용한 곡선추정방법에 대해서, Herschy(1978)는 추정된 유량의 신뢰구간 추정에 대한 연구를, 국제표준기구(ISO, 1998)에서는 NLSM 모형을 이용한 수위-유량곡선방법을 추천하고 있는데, 이 모든 연구들의 기본가정은 오차항의 분산이 동질하다는 가정을 성립할 경우에만 사용할 수 있는 것이고, 오차항의 분산이 이질적인 경우 사용하게 되면 오류를 범한다는 것에 대한 충분한 부연설명이 되어있지 않은 경우가 대부분이었다.

기타 연구들로는 Gawne and Simonovic (1994)은 이분산성을 고려하여 분산의 안정화를 위해 자료변환을 실시한

후 곡선개발을 실시하였으며, Clake(1999)와 Clake and Mendiondo et al.(2000)은 NLSM 모형으로 개발된 추정유량의 불확실도 분석을 실시하였다.

본 연구에서 다른 분산의 이질성에 대한 추천 기법으로는 선형 케이스일 경우는 Huet et al.(1996)은 가중최소제곱법(weighted least squares)을 추천하고 있는데, 이 경우는 수위-유량곡선 개발과 같은 비선형 추정에는 부적절한 방법이다. 따라서 수위-유량곡선과 같이 비선형 케이스일 경우에는 이분산성을 고려한 최우도모형(heteroscedastic maximum likelihood Model; HMLM)을 추천하고 있다(Asgeir, 2004; 서혜선 등, 2003).

그 외 수위-유량곡선식 개발에 적용된 다른 기법들로 GLSM(generalised least squares model)과 대안적 기법을 적용한 3단계 평균제곱 기법(the three-step alternate mean squares method)이 있다(Seber and Wild; 1989, Huet et al.; 1996). 이런 여러 가지 기법들이 있음에도 본 연구에서 HMLM 모형을 사용한 이유는 첫 번째 단순하고 실용적인 기법이고, 두 번째 추정된 모수추정 결과들이 참값에 대한 접근값을 확인할 수 있는 장점을 지니고 있기 때문이다.

본 연구의 목적은 이분산성이 내재된 자료에서 NLSM 모형으로 개발된 기존 곡선식의 제한성을 확인하고, HMLM 모형으로 개발된 곡선식이 NLSM 모형과 비교하여 얼마나 만족스러운 추정결과를 제공하는가를 객관적 검정기법 이용하여 비교·분석하는데 있다.

II. 수위-유량 곡선식

1. 비선형최소제곱모형(NLSM)

전통적인 수위-유량곡선식은 식 (1)과 같고, 식 (1)로 유도된 곡선식에서 측정오

차는 기대유량(expected discharge)에 선형적으로 비례한다고 가정한다.

$$Q_i = a(w_i + c)^b + E[Q_i] \varepsilon_i \quad (1)$$

여기서 $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ 이고 i 는 총 측정치 n 의 계열치를 의미한다. 먼저 전통적인 기법 적용 시 기본적으로 다음과 같은 가정을 만족한다는 전제하에서 분석을 실시하게 된다. 첫째로 오차항의 불확실도를 정규분포(normal distribution)라고 가정한다는 것이다. 왜 정규분포인가? 일반적으로 총 불확실도는 유량 측정치에 내포된 여러 다른 요인들(횡단면 형상, 유속분포, 측정장비 등)에 의해서 발생하는 오차들의 합으로 구성된다. 이와 같이 각 오차들은 서로 다른 요인들에 의해서 발생하므로 각 오차들의 분포 또한 서로 다르게 발생할 것으로 생각되지만 오차의 중앙집중성향(central limit theorem)에 따라 일반적으로 정규분포의 경향을 가질 것으로 가정한다.

둘째로 기준수위(w)는 항상 정확하게 관측된 자료값이라고 가정한다. 즉 수위 w 는 항상 예측변수이고 유량 Q 는 반응변수가 되는 것이다. Seber and Wild(1989)은 식 (1)을 다음과 같이 변환하여 표시하였다.

$$Q_i = a(w_i + c)^b(1 + \varepsilon_i) \quad (2)$$

식 (2)의 양변에 대수변환을 실시하면

$$\log Q_i = \log a + b \log(w_i + c) + \log(1 + \varepsilon_i^*) \quad (3)$$

이 된다.

여기서 $|\varepsilon_i| \ll 1$ 이고 테일러의 접근(Taylor approximation)을 이용하여 식 (4)와 같이 변환할 수 있다.

$$\log Q_i = \log a + b \log(w_i + c) + \varepsilon_i^* \quad (4)$$

식 (4)에서 잔차분포는 정확하지는 않지만 대략적으로 정규분포를 따를 것이라고 가정한다.

모든 수문관측지침서 및 설계기준들을 살펴보면 전통적인 비선형최소제곱법 모형으로 식 (4)을 추천하고 있으며, 단지 앞에서 언급했듯이 이분산성에 대한 언급만 있을 뿐이며, 식 (1)에서 (3)에 표현한 것과 같은 가정들에 대한 설명은 자세하게 언급되어 있지 않은 경우가 대부분이었다(건교부, 1994; 김원 등 2004; Herschy, 1978; ISO, 1998; Lambie, 1978).

식 (4)에서 매개변수 (a, b, c) 추정은 잔차제곱합 S^2 이 최소가 되는 해를 찾으려 가능하다.

$$S^2 = \sum_{i=1}^n [\log Q_i - \log a - b \log(w_i + c)]^2 \quad (5)$$

또한 식 (1)에 오차항의 표준편차 σ 는 잔차분석을 통하여 추정할 수 있다. 일반적으로 잔차분석은 모멘트법을 사용하는데 매개변수가 3개이므로 자유도는 $n-3$ 을 가지는 식 (6)을 사용하여 표본 표준편차를 이용하여 추정하게 된다.

$$\hat{\sigma} = \sqrt{\frac{1}{n-3} \sum_{i=1}^n (e_i - \bar{e})^2} \quad (6)$$

여기서 e_i 는 표본잔차(observed residual)를 의미한다. 식 (4)와 같은 대수 변환모형은 분산을 안정화시키기 위한 것으로 종속변수와 독립변수에 대한 일반적인 분산 안정화 변환식으로 알려져 있다. 이 방법은 단지 유용한 변환방식일 뿐이며 최소제곱법에 국한되는 기법임을 명심해야 한다. 그러므로 이 기법 이외의 보다 응용된 기법들에서 원자료에 대한 매개변수 a, b, c 로 자료변환이 불가능한 경우가 종종 발생한다고 알려져 있다(서혜선 등, 2003). 그러나 이 변환이 이분산성의 형태를 조절할

수 있고 선형분석을 수행할 수 있는 간편식으로 표현할 수 있는 장점을 가지고 있다.

2. 이분산성을 고려한 최우도모형(HMLM)

이분산성을 정확하게 해석하기 위해서는 먼저 보다 완전한 모형을 채택하여야 하며 매개변수 추정에 있어서도 경제적이고 유동성을 가지고 있어야 한다. 또한 선형 이분산성을 일반화 시킬 수 있어야 한다. 이와 같은 조건을 만족시키는 확장모형을 제안하면 식 (7)과 같다.

$$Q_i = \alpha(w_i + c)^b + E^w[Q_i] \varepsilon_i \quad (7)$$

위 식을 다음과 같이 변환하여 표시하면

$$Q_i = \alpha(w_i + c)^b + (w_i + c)^\beta \varepsilon_i \quad (8)$$

여기서 $\varepsilon_i \sim N(0, \tau^2)$ 이라고 가정한다. 전자의 거동이 정규분포라고 가정하면 대수우도함수(log-likelihood function) l 은 식 (9)와 같이 표현할 수 있다.

$$l = -\frac{n}{2} \log 2\pi - n \log \tau - \beta \sum_{i=1}^n \log(w_i + c) - \sum_{i=1}^n \left[\frac{Q_i - \alpha(w_i + c)^b}{\sqrt{2\tau(w_i + c)^\beta}} \right]^2 \quad (9)$$

매개변수 ($\hat{\alpha}$ \hat{c} \hat{b} $\hat{\tau}$ $\hat{\beta}$)들은 식 (9)을 최대화시키는 값으로 추정할 수 있으며, 추정된 매개변수들은 HMLM 모형의 수위-유량곡선식의 매개변수로 사용된다.

3. 이분산성 조건에서 회귀추정, 진단 및 추론

Seber and Wild(1989)와 서혜선 등(2003)에 따르면 이분산성에 대한 적절한 변환 없이 비선형 모형을 개발하게 되면 매

개변수 추정에 오류를 범할 수 있다고 경고한다. 즉 이분산성이 적절하게 계산되지 않은 자료를 이용하여 최소제곱 추정을 실시하였을 경우, 추정치는 분석에 사용된 자료 중 최고 또는 최저 관측치에 지배적인 영향(=전체제곱합의 크기를 좌우함)을 받게 되므로 적절치 못한 추정결과를 산출하게 된다. 따라서 모형 개발 시 이분산성을 내포한 자료를 등분산이라고 가정하면 극치값들에 대한 영향을 크게 받는 회귀분석결과를 초래할 것이다. 본 연구에서는 실제 측정치의 교차-타당성(cross-validation)을 확인하기 위해서 robustness 검정을 실시하였으며 일반적인 절차는 다음과 같다.

- (1) n 개의 원자료에서 무작위로 $n/2$ 표본을 추출한다.
- (2) $n/2$ 표본에 대한 매개변수를 추정하고 추정된 회귀식을 이용하여 w^* 에 대한 Q^* 을 계산한다.
- (3) 500번 표본 추출을 실시하고 이 표본에 대한 평균 \hat{Q} 을 계산한다.
- (4) robustness 검정을 비교하기 위해 계산된 각각의 \hat{Q} 에 대한 변동계수 $CV[\hat{Q}]$ 을 계산한다. 분석결과 원자료의 저항성이 강하다는 것은 각 표본자료의 영향에 덜 민감하다는 것으로, 계산된 변동계수 $CV[\hat{Q}]$ 가 적게 발생하게 될 것이다.

분산의 이질성을 적절하게 변환하지 않았다고 해서 매개변수 추정에 편의(bias)를 유발시킬 것이라고 단정 지을 수는 없지만, 10~50개 내의 측정자료를 이용하여 분석할 경우 곡선의 신뢰도를 점근적으로 저하시키는 결과를 초래하므로 robustness 검정결과는 매우 중요하다.

일반적으로 곡선식의 적합도검정(goodness-of-fit)으로 주로 사용하는 통계치로 결정계수 r^2 을 사용한다. 이 계수는 서로 다른 지점별 적합도 검정을 비교하기

나 이분산성이 적절하게 계산되지 않은 특정 지점의 서로 다른 표본 비교에 사용할 수 없다. 이와 같은 단점으로 인하여 추계학적 기법에서 결정계수 r^2 을 일반적인 곡선 추정식의 회귀검정에 사용하는 것에 의문을 제기하고 있다(Asgeir, 2004).

수위-유량곡선식에서 추정된 유량은 개발된 곡선식의 가변성으로 인하여 불확실성을 내포하고 있다. Clarke(1999)에 의해 적용된 매개변수적 붓트스트랩(parametric bootstrap)과 점근적 기법들(asymptotic methods)과 같이 등분산성을 기초로 한 추계학적 이론들은 현 모형에 이분산성이 존재할 경우 부정확한 결과를 초래하게 되고 신뢰구간 산정에도 문제를 발생시킬 것이라고 주장하였다. 대안적 방법으로 비매개변수적 붓트스트랩(non-parametric bootstrap)을 사용할 수 있을 것이며, 추출된 샘플링 표본자료에 대한 회귀분석결과 비교시 매개변수적 방법보다는 비매개변수적 방법들이 추정모형에 덜 민감한 것으로 연구되었다(Efron and Tibshirani, 1993).

III. 적용 및 곡선 추정

1. 이분산성 발생

본 연구에 적용된 자료는 남강댐 상류유역에 위치하는 산청, 신안 그리고 창촌 수위관측소의 2004년 실측유량자료를 활용하였으며, 이분산성 발생을 확인하기 위해서 먼저 각 지점별로 개발되는 곡선식은 안정하상으로 단일곡선식으로 개발 가능하다고 가정하였다. 이와 같은 가정은 각 지점별 단면변화 및 하상변동 등으로 인하여 두 개 이상의 실제 수위-유량곡선식이 개발되는 것이 일반적이나, 두 분석기법(NLSM과 HMLM 모형)의 비교 및 HMLM 모형의 적용성 및 우수성을 확인하는 차원과 기호 및 검정의 간편성을 위해 단일 수위-유량곡선식에 대해서 우선 비교·분석을 실시하였다. 이후 분석결과들을 살펴보면 HMLM으로 개발된 모형의 우수성이 객관적인 검정법에 의해 입증되므로, 각 지점별로 실제 수위-유량곡선식을 개발하였으며 그 적용성에 대해 비교·검토하였다.

Table 1. Information and rating curve variance heterogeneity classification to the gauging stations included in the study.

Station	Area(km ²)	Number of measurements	Measurement methods ¹⁾	NLSM classification ²⁾	HMLM classification ²⁾
Sancheong	1,158.6	24	EOCFM, MWSCM	B	A
Shinan	390.5	24	EOCFM, MWSCM	A	A
Changchon	333.9	24	EOCFM, MWSCM	B	A

1) Electromagnetic open channel flow meter(EOCFM), Microwave water surface current meter(MWSCM)
 2) Rating curves are classified as non-heteroscedastic (A) and heteroscedastic (B) after fitting the rating curves using both the non-linear least square and the heteroscedastic maximum likelihood models.

각 수위관측소별 유역면적, 지점별 실측 자료수 및 측정방법에 대한 일반적인 사항들을 정리해 보면 Table 1과 같다. 또한 Table 1에는 잔차분석을 통해 이분산성에 대한 정리를 함께 수록하였다. 여기서 이분산성의 증거는 다음의 정의에 따라 조사하여 표시하였다.

- A : 이분산성이 발생하지 않거나 미약할 경우
- B : 이분산성이 강력하게 발생할 경우

이분산성을 확인하기 위해서 식 (4) (NLS M 경우)와 식 (8) (HMLM 경우)을 이용하여 오차항에 대한 정규화된 표본잔차 플롯을 실시한다. 분석결과 반응변수에 잔

차들이 무작위적으로 발생하는 경우는 등분산성이라는 기본가정을 만족하게 되고, 반응변수에 임의패턴(증가 또는 감소형)을 보일 경우 이분산성을 의심해 보아야 할 것이다. 특히 수위-유량자료들은 홍수기 및 저갈수기 유량측정 시 측정오차가 심하게 발생하는 개연성을 가지므로 증가형 또는 감소형의 이분산성이 많이 발생할 것으로 추정된다. Fig. 2는 남강댐 분석 지점 중 신안과 산청에 대한 잔차플롯 결과로 신안지점(Fig. 2(a))은 오차의 등분산성을 가정할 수 있지만, 산청지점(Fig. 2(b))의 경우 분산이 예측변수에 따라 증가하는 형태를 보이고 있는 것을 확인할 수 있었다.

Table 1의 분석결과 NLSM의 경우에는 신안지점의 경우를 제외한 산청과 창촌지점에서 이분산성이 현저히 발생하였고, HMLM의 경우에는 이분산성이 발생한 지점에 대해서 적절한 변환함수를 찾아서 분석함으로써 분산에 대한 이질성이 모두 제거된 것을 확인할 수 있었다. 위 결과에서 알 수 있듯이 NLSM으로 개발된 모형에서는 식 (7)과 같은 이분산성을 제거할 수 있는 가중함수의 변환함수를 적절하게 찾을 수 없기 때문에 이분산성을 제거할 수 없는 단점을 지니므로 이와 같은 자료에 NLSM을 사용하는 것은 추정치의 신뢰도를 떨어뜨리는 결과를 초래하게 된다.

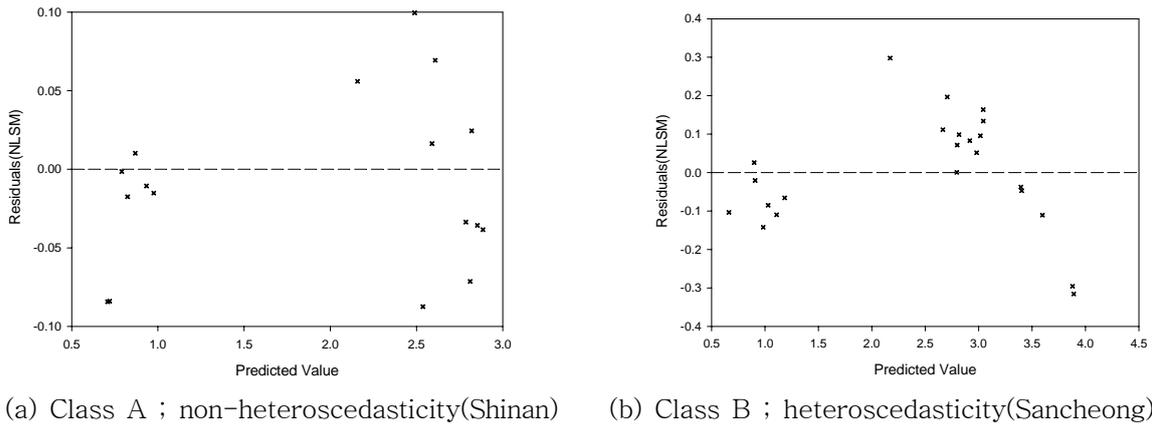


Fig. 2. Examples of residuals in a rating curve fit using the non-linear least square method.

Table 2. Robustness analysis¹⁾ of the NLSM and HMLM.

Station	Area (km ²)	N	NLSM			HMLM				
			class	CV[\widehat{Q}_{hd}^*]	CV[\widehat{Q}_{md}^*]	CV[\widehat{Q}_{lud}^*]	class	CV[\widehat{Q}_{hd}^*]	CV[\widehat{Q}_{md}^*]	CV[\widehat{Q}_{lud}^*]
Sancheong	1,158.6	24	B	0.076	0.013	0.017	A	0.047	0.012	0.014
Shinan	390.5	24	A	0.026	0.011	0.012	A	0.025	0.012	0.013
Changchon	333.9	24	B	0.037	0.020	0.026	A	0.014	0.020	0.026

1) The rating curves are estimated with a random selection of n/2 measurements, and the discharge related to the lowest water level, median water level and the highest water level in each set of measurement is calculated. This procedure is repeated 500 times and the coefficient of variation is then calculated.

Fig 3은 이분산성이 존재하는 산청지점의 실측자료에 대해 HMLM의 모형 개발시 발생된 잔차분석 결과를 도시한 것이다. 그림에서 알 수 있듯이 모형에서 발생하는 오차항에 이분산성에 대한 가중 변환함수를 이용하여 적절한 가중치를 부여함으로써 개발된 모형에서 발생한 잔차플롯이 등분산성이라는 가정에 만족하는 것을 확인할 수 있다.

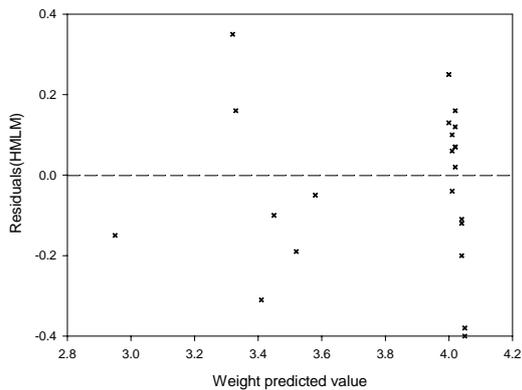


Fig. 3. HMLM of residual plot(sancheong).

2. Robustness 검정

추출된 표본자료가 원자료의 특성을 얼마나 잘 반영해줄 수 있는가를 검정하는 기법으로 robustness 검정이 있다. 본 연구에서는 2.3절에서 설명한 검정절차에 따라 남강댐 상류 3개 지점에 대해 검정결과를 정리하면 Table 2와 같다. Table 2는 원자료에 $n/2$ 표본을 500번 무작위 표본추출을 실시하여 산출된 결과이며, 관측지점별 실측유량자료의 측정수위범위 내의 최저수위 (w_{mi}^*), 중앙수위(w_{mi}^*) 그리고 최고수위

(w_{mi}^*)에 대한 추출 표본의 추정유량에 대한 변동계수 $CV[\widehat{Q}_{mi}^*]$, $CV[\widehat{Q}_{mi}^*]$ 그리고 $CV[\widehat{Q}_{mi}^*]$ 를 각각 계산하였다.

Table 2의 분석결과를 살펴보면 NLSM 분석결과 등분산성이라는 가정을 위배하지 않은 경우(Class A)에는 HMLM와 비교하여 각 수위별 유량 변동계수 값이 같거나 조금 작은 변동치를 보이지만, NLSM 분석결과 이분산성이 발생한 경우(Class B)는 HMLM의 분석결과와 비교하여 그 변동계수의 변동치가 상대적으로 크게 발생하는 것 확인할 수 있었다. 또한 HMLM 분석결과가 최고/최저수위와 같은 극치값에 대한 변동치가 상당히 감소되는 것으로 분석되었다. 이는 HMLM 모형이 전통적인 NLSM 모형보다 최고수위 및 최저수위와 같은 극치부분의 실측 유량값을 잘 반영하는 것을 입증해 주는 것이다. 따라서 이분산성이 존재하는 자료에 대해서는 NLSM보다 HMLM이 원자료의 특성을 잘 반영하여 적절한 수위-유량곡선식을 제공해 줄 수 있을 것으로 판단된다.

3. 실제 수위-유량곡선식 개발 및 검토

위 연구결과로부터 오차항의 기본 가정을 위배하는 경우에 NLSM보다 HMLM 모형이 우수하다는 것을 확인할 수 있었다. 따라서 남강댐 상류 산청, 신안 그리고 창촌 수위관측소에 대한 가용한 유량측정성과를 활용하여 실제 수위-유량곡선식을 NLSM과 HMLM 모형별로 각각 개발하였다.

Table 3. Development of rating curves for NLSM and HMLM.

Station	Range of stage	Rating curve	r^2	Remark
Sancheong	$1.30 \leq H \leq 3.42$	$Q=0.1835(H+0.40)^{5.4322}$	0.984	NLSM
		$Q=73.2324(H-1.50)^{1.9809}$	0.964	
	$3.42 < H \leq 9.21$ $H=h+1$	$Q=0.1372(H+0.40)^{5.7220}$	0.998	HMLM
		$Q=77.3736(H-1.50)^{1.9454}$	0.967	
Shinan	$1.95 \leq H \leq 4.01$	$Q=0.3068(H-0.30)^{5.1644}$	0.997	NLSM
		$Q=144.8143(H-2.50)^{1.4758}$	0.891	
	$4.01 < H \leq 5.72$ $H=h+1$	$Q=0.3022(H-0.30)^{5.1857}$	0.995	HMLM
		$Q=147.4416(H-2.50)^{1.4548}$	0.871	
Changchon	$0.46 \leq h \leq 1.36$	$Q=0.6811(h+1.00)^{6.8893}$	0.996	NLSM
		$Q=298.9474(h-0.47)^{1.4912}$	0.948	
	$1.36 < h \leq 2.88$	$Q=0.6973(h+1.00)^{6.8519}$	0.999	HMLM
		$Q=44.0827(h+0.55)^{2.6877}$	0.964	

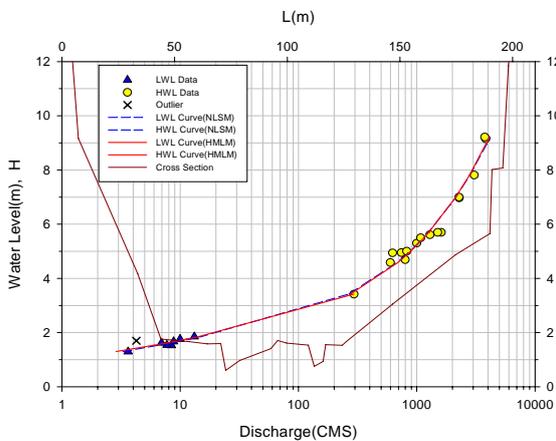


Fig. 4. Estimated rating curves and actual data(sancheong)

Fig. 4에는 두 기법을 적용하여 개발된 산청 수위관측소의 실제 개발된 수위-유량 곡선식을 측정자료, 기준 횡단면도와 함께 작성한 것이며, Table 3은 남강댐 상류 3개 수위관측소에서 두 기법으로 개발된 수위-유량곡선식을 결정계수 r^2 와 함께 제시한 것이다. Table 3의 분석결과 각 지점별

적합도 검정결과인 결정계수 r^2 은 이분산성이 발생한 산청과 창촌지점의 경우 NLSM보다 HMLM 값이 조금 향상된 것을 확인할 수 있지만, 이분산성이 발생하지 않은 신안지점의 경우 HMLM의 결정계수 r^2 은 조금 나빠진 것을 확인할 수 있다. 이 결과로 오차항의 기본가정을 위배하는 이분산성 자료의 경우는 HMLM이 보다 향상된 곡선식을 제공해 주는 것을 기존 검정법으로도 확인할 수 있었다.

또한 개발된 곡선식의 타당성을 검토하기 위해 남강댐 유입량과 각 지점별 일평균 유량에 대한 유황분석을 실시하였으며, 그 분석결과는 Table 4와 같다. 남강댐 유입량과 각 지점별 유출량 분석결과를 살펴보면 이분산성이 발생한 산청과 창촌지점의 차이(%)가 크게 발생하는 것을 알 수 있다. 특히 산청의 연유출량을 살펴보면 NLSM이 HMLM가 비교하여 0.22%(1.85백만톤) 과대 산정되는 것으로 분석되었으며, 신안과

창춘은 각각 -0.12% (0.48백만톤)과 -0.07% (0.36백만톤) 과소 산정되는 것으로 분석되었다.

유황분석결과 Table 4에서 알 수 있듯이 자료에 이분산성이 존재하는 경우의 유황값은 등분산성이 가정을 만족하는 유황값 자료들과 비교하여 그 변동폭이 심하게 발생하는 것을 알 수 있으며, 결국 각 지점별 유출량 산정 시 오류를 범할 확률이 그만큼 증가할 수 있는 개연성을 가지고 있는 것으로 분석되었다. 따라서 곡선 개발 시 이분산성이 존재하는 자료에서는 HMLM 모형이 보다 타당하지 않을까 생각한다.

IV. 결 론

유량은 시공간적 변화가 매우 심한 자연 현상으로 그 참값을 정확하게 알 수 없는 인자이다. 하지만 이러한 자연현상이라도 신뢰성 있는 자료관측기법과 분석기법을 활용하여 참값에 가까운 값을 규명하고자 노력하여야 하는 것도 수문학자들의 몫이라고 하겠다.

본 연구는 일반적으로 수위-유량곡선식 개발 시 사용되는 비선형최소제곱 모형(NLSM)의 제한성을 제시하고자 하는 것이다. 측정유량 내에는 불확실성을 내포하고 있으며, 이 불확실성은 기대유량에 선형적으로 비례한다고 가정한다. 즉 비선형 회귀분석은 이와 같은 가정에서 출발하게 된다. 하지만 본 연구결과를 살펴보면 모든 자료에서 이와 같은 가정을 만족하지 않는다는 것을 확인할 수 있었다.

본 연구에서는 이와 같은 가정에 만족하지 않는 경우, 이분산성을 고려한 최우도모형(HMLM)을 제안하였다. 이 기법은 원모형에 추가적인 매개변수를 이용하여 분산이 동질한 모형을 구축하는 단지 한 가지 기법에 불과하며 절대적인 방법은 아님을 밝혀둔다.

본 연구의 핵심은 범용성에 의문점을 제

시할 수 있는 비선형회귀모형과 같은 통계적 기법에 대해 그 의문점을 해결할 수 있는 방안을 제시하여 문제를 해결하고자 하는데 초점을 두고 있다. 만일 모형의 가정이 위배된다면 분석결과는 잘못된 결과를 제공하게 될 것이기 때문이다.

수위-유량곡선식은 수문조사 업무에서 주요한 부분이므로 유량측정지침들을 마련하는 기관들은 그 적용성 및 범용성에 보다 신중을 기하여 지침들은 제작해야 할 것이다. 만일 표준 유량측정에 대한 의문점이 발생하여 적용성에 제한을 받는다면 그 지침들에서 기법 적용의 제한성에 대한 실례들을 자세하게 설명해 놓으면 보다 나은 수위-유량곡선식 개발이 이루어 질 수 있을 것이다. 본 논문은 이러한 하나의 시도이며 후속 연구로 본 연구에서 사용한 기법 이외에 보다 나은 통계적 기법들에 대한 연구들이 계속해서 수행되어야 할 것으로 생각된다.

참 고 문 헌

1. 김원, 윤광석 등. 하천유량측정지침, 수자원의 지속적 확보기술개발사업단 기술보고서(TR 2004-01), pp. 73-82, 2004.
2. 건교부 수자원국 하천관리과 수문조사선진화 5개년 계획(하천관리과-1147, 2005. 6. 1), 2005.
3. 건교부 수자원관리기법 개발연구조사 보고서, 1994.
4. 건설교통부 수문조사연보, 수위편, 유량편, 2004.
5. 김경훈, 김문수 등 낙동강수계 T/M유량과 실측유량의 비교분석, *한국물환경학회. 대한상하수도학회 공동춘계학술발표회 논문집*, pp. 629-632, 2005.
6. Lambie, J.C. Measurement of flow-velocity-area methods. In: Herschy, R.W., (Ed.), *Hydrometry: Principles and Practices*, Wiley, Chichester, Chapter

- 1, 1978.
7. Ackers, P., White, W.R., Perkins, J.A., Harrison, J.M. Weirs and Flumes for Flow Measurement, Wiley, Chichester, 1978.
 8. Gawne, K.D., Simonovic, S.P. A computer based system of modelling the stage-discharge relationships in steady state conditions, *Hydrol. Sci. J.* **39**(5), 487-506, 1994.
 9. Yu, B. A systematic over-estimation of flows. *J. Hydrol.* **233**, 258-262, 2000.
 10. Asgeir Petersen-Øverleir Accounting for heteroscedasticity in rating curve estimates, *J. Hydrol.* **292**, 173-181, 2004..
 11. Herschy, R.W. Accuracy. In: Herschy, R.W., (Ed.), *Hydrometry: Principles and Practices*, Wiley, Chichester, Chapter 10, 1978..
 12. ISO Technical Corrigendum 1 to International Standard ISO 1100-2:1998. Prepared by Technical Committee ISO/TC 113. *Hydrometric Determinations*, Sub-committee SC 1. Velocity Area Methods, 1998.
 13. Clarke, R.T. Uncertainty in the estimation of mean annual flood due to rating curve indefintion, *J. Hydrol.* **222**, 185-190, 1999.
 14. Clarke, R.T., Mendiondo, E.M., Brusa, I.C. Uncertainties in mean annual discharge from two large South American rivers due to rating curve variability, *Hydrol. Sci.* **45**, 221-236, 2000.
 15. Huet, S., Bouvier, A., Gruet, M.A., Jolivet, E. *Statistical Tools for Nonlinear Regression*, Springer, London, 1996.
 16. Huet, S., Bouvier, A., Gruet, M.A., Jolivet, E. *Statistical Tools for Nonlinear Regression*. Springer, London, 1996.
 17. Seber, G.A.F., Wild, C.J. *Nonlinear Regression*, Wiley, Chichester, 1989.
 18. 서혜선, 양경숙, 김나영, 김희영, 김미정 SPSS를 활용한 회귀분석, SPSS 아카데미, pp. 233-262, 2003.
 19. Efron, B., Tibshirani, R.J. *An Introduction to the Bootstrap*, Chapman & Hall, London, 1993.