

Fuzzy-Grey 모형을 이용한 유입량 예측

Inflow Forecasting Using Fuzzy-Grey Model

김용* / 이충성** / 김형수*** / 심명필****

Kim, Yong / Yi, Choong Sung / Kim, Hung Soo / Shim, Myung Pil

요 지

본 연구는 Deng(1989)이 제시한 Grey 모형을 이용하여 섬진강댐의 월유입량을 예측하였고 그 방법을 제시하였다. Grey 모형은 시계열모형이나 다른 모형에 비해 비교적 적은 수의 자료를 이용하고, 간단한 수식으로 구성되어 있는 장점이 있으나, 적은 수의 자료로 인해 입력자료가 가지는 증감의 경향(trend)으로 오차가 발생하기 쉽다. 그러므로 예측오차를 극복하기 위해서 Fuzzy 시스템을 결합한 Fuzzy-Grey 모형을 구성하였고 Fuzzy 시스템에 필요한 매개변수를 추정하기 위해 최적화기법인 유전자 알고리즘(GA; Genetic Algorithm)을 이용하였다.

Grey 모형과 결합된 Fuzzy 시스템은 현재의 입력자료가 가지는 패턴과 가장 유사한 패턴의 과거자료를 이용하여 현재의 입력자료의 예측오차를 추론해내는 기능을 가졌다. 오차를 추론하기 위해서 과거 월유입량 자료 중 현재 입력 자료와 유사한 패턴을 Grey 상관도를 이용하여 검색하고, 보다 높은 유사성을 가지는 패턴을 선별하고자 노름(norm)을 사용하였고, 유전자 알고리즘의 탐색공간을 제한하였다.

이렇게 구성한 Fuzzy-Grey 모형을 이용하여 전국적인 가뭄년도였던 1982년, 1988년, 2001년에 대해 섬진강댐의 월유입량을 예측하였다. 오차는 1982년, 2001년, 1988년 순으로 비슷한 크기의 오차가 발생하였는데 결과를 분석하여 보면, 급격한 월유입량의 변화가 있었던 경우에 오차가 크게 발생하였으나 가뭄년도에 대해 월유입량의 불확실성이 큼에도 불구하고 비교적 월유입량의 추세를 잘 예측한 것으로 판단된다.

본 연구에서 적용한 Fuzzy-Grey 모형은 적은 수의 자료를 이용하여 예측하고 예측결과를 다시 입력자료로 사용하는 업데이트 방식을 사용하기 때문에 예측결과의 오차가 완전하게 보정되지 않으면 다음 결과에 역시 오차를 주게 되어 오차보정이 상당히 중요하다는 것을 알 수 있었다. 오차를 보다 효과적으로 보정하기 위해서는 퍼지제어에 사용되는 퍼지규칙의 수를 늘리고, 유입량에 직접적인 영향을 주는 강우량과 연계한 2변수의 Fuzzy-Grey 모형을 이용한다면 보다 정확한 유입량 예측이 가능할 것으로 사료된다.

핵심용어 : Grey 모형, 퍼지제어, 유전자 알고리즘, Fuzzy-Grey 모형

1. 서 론

본 연구는 5~8년 빈도로 발생하는 가뭄을 대비하고(건설교통부, 2001), 수자원의 효율적인 이용을 위해 중국을 중심으로 연구되고 있는 Grey모형을 활용하여 가뭄연도의 월유입량을 예측하였다. Grey 모형은 Deng(1989)에 의해 제시되어 전력, 전기 실시간 수요예측, 통합회로의 연수요와 판매량 예측 등에 적용되었다. 수자원분야에는 강우량 예측을 위해 Yu 등(2000)이 Grey 모형의 매개변수를 Fuzzy 회귀방법으로 추정하여 실시간 강우예측에 적용하였고, Yu 등(2001)이 실시간 홍수량 예측에 적용하였다.

* 정회원 · (주)한국종합기술개발공사 수자원부 · E-mail : civlky@kecc.co.kr
** 정회원 · 인하대학교 환경토목공학부 박사과정 · E-mail : g2021533@inhavision.inha.ac.kr
*** 정회원 · 인하대학교 환경토목공학부 조교수 · E-mail : sookim@inha.ac.kr
**** 정회원 · 인하대학교 환경토목공학부 교수 · E-mail : shim@inha.ac.kr

Grey 모형은 4개의 이전 자료를 사용한다는 점에서 확정론적 모형(deterministic model)과 추계학적 모형(stochastic model)과는 크게 다르며, 간단한 수식으로 구성되어 있는 장점이 있다. 그러나 적은 수의 자료로 인해 입력자료가 가지는 증감의 경향성(trend)의 영향을 받아 예측오차가 발생하게 된다. 본 연구에서는 이러한 Grey 모형의 예측오차를 보정하기 위해 퍼지제어를 결합하여 Fuzzy-Grey 모형을 구성하였다. 또한 예측오차에 대한 사전 지식과 정보가 없으므로 퍼지제어에 필요한 소속함수를 결정하기 위해 유전자 알고리즘(GA; Genetic Algorithm)을 이용하였다.

2. GM(1,1)모형

보통 1변수에 대한 1차 미분방정식을 사용한 Grey 모형을 GM(1,1)모형이라 하며, N개의 변수를 가지는 Grey 모형의 경우는 GM(1,N)으로 나타낸다. 본 연구는 월유입량을 변수로 하는 GM(1,1)모형을 적용하였으며, 모형의 구조는 다음과 같다.

관측자료의 계열이 $X^{(0)} = [x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(i), \dots, x^{(0)}(n)]$ (여기서, $x^{(0)}(i)$ 는 i 번째 관측 자료)로 주어졌을 때 새로운 시계열 $X^{(1)} = [x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)]$ 은 원시계열로부터 1차 AGO(Accumulated Generating Operator)로써 발생되어진다. 여기서 $x^{(1)}(k)$ 는 식 (1)과 같이 정의된다.

$$AGO X^{(0)} : x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i) , i = 1, 2, 3, \dots, k \in n \quad (1)$$

AGO 계열은 미분방정식에 의해 조절되며 편의성을 위해 1차 미분방정식이 사용된다.

$$\frac{dX^{(1)}}{dt} + \otimes a X^{(1)} = \otimes b \quad (2)$$

여기서, $\otimes a$, $\otimes b$ 는 오차계열에 대한 전개, 조절계수이다. 식 (2)에서 $\otimes a$, $\otimes b$ 는 Grey 매개변수이며, Grey 시스템 이론에서 단위시간을 가지는 이산화 자료(discrete data)의 Grey 미분계수의 whitening은 식 (3)과 같다.

$$\left. \frac{dX^{(1)}}{dt} \right|_{t=k} = \lim_{\Delta t \rightarrow 0} \frac{x^{(1)}(k) - x^{(1)}(k-\Delta t)}{\Delta t} = x^{(1)}(k) - x^{(1)}(k-1) = x^{(0)}(k) \quad (3)$$

새로운 변수 $z^{(1)}(k)$ (background value)를 정의하고, 이 변수에 대한 $X^{(1)}|_{t=k}$ 의 whitening value는 식 (4)와 같다.

$$X^{(1)}|_{t=k} \cong z^{(1)}(k) \equiv \beta [x^{(1)}(k) + x^{(1)}(k-1)] : \forall k = 2, 3, \dots, n \quad (4)$$

여기서, β 는 보통 0.5를 사용하며 식 (4)는 연속된 누적자료의 평균값을 의미하게 된다. 식 (3), (4)를 식 (2)에 대입하고 이산미분방정식으로 식을 쓰면 식 (5)를 얻을 수 있다.

$$x^{(0)}(k) + \otimes a z^{(1)}(k) = \otimes b : \forall k = 2, 3, \dots, n \quad (5)$$

식 (5)를 $Y = U\theta$ 와 같은 행렬로 구성하면, 각각을 식 (6)과 같이 행렬로 변형할 수 있다.

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \quad U = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ -z^{(1)}(4) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad \theta = \begin{bmatrix} \otimes a \\ \otimes b \end{bmatrix} \quad (6)$$

Y와 U는 주어진 whitening 행렬이므로 Grey 매개변수 벡터 Θ 의 whitening value는 최소자승법에 의해 구해지며 식 (7)과 같이 나타낸다.

$$\Theta = \begin{bmatrix} \otimes a \\ \otimes b \end{bmatrix} = (U^T U)^{-1} U^T Y \quad (7)$$

Grey 파라미터가 구해지면 식 (2), (5)의 해는 식 (8)에 의해 주어진다.

$$\hat{x}^{(1)}(k+1) = \otimes c e^{-\otimes a k} + \otimes d \quad (8)$$

여기서 $\otimes c = x^{(0)}(1) - \otimes d$, $\otimes d = \frac{\otimes b}{\otimes a}$ 이다. 이와 같이, $\hat{x}^{(1)}(k+1)$ 의 역을 적용함으로써 시간 k+1 에서의 예측값 $\hat{x}^{(0)}(k+1)$ 은 식 (9)와 같이 주어진다.

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) : \forall k = 1, 2, 3, \dots, n \in N \quad (9)$$

3. 유전자 알고리즘을 이용한 퍼지학습

퍼지제어를 학습시키기 위한 학습자료가 결정되면 입력-출력 쌍을 만들고 추론규칙을 적용하고(Huang, 1996), 그림 1과 같이 유전자 알고리즘을 이용하여 삼각형 소속함수를 결정한다.

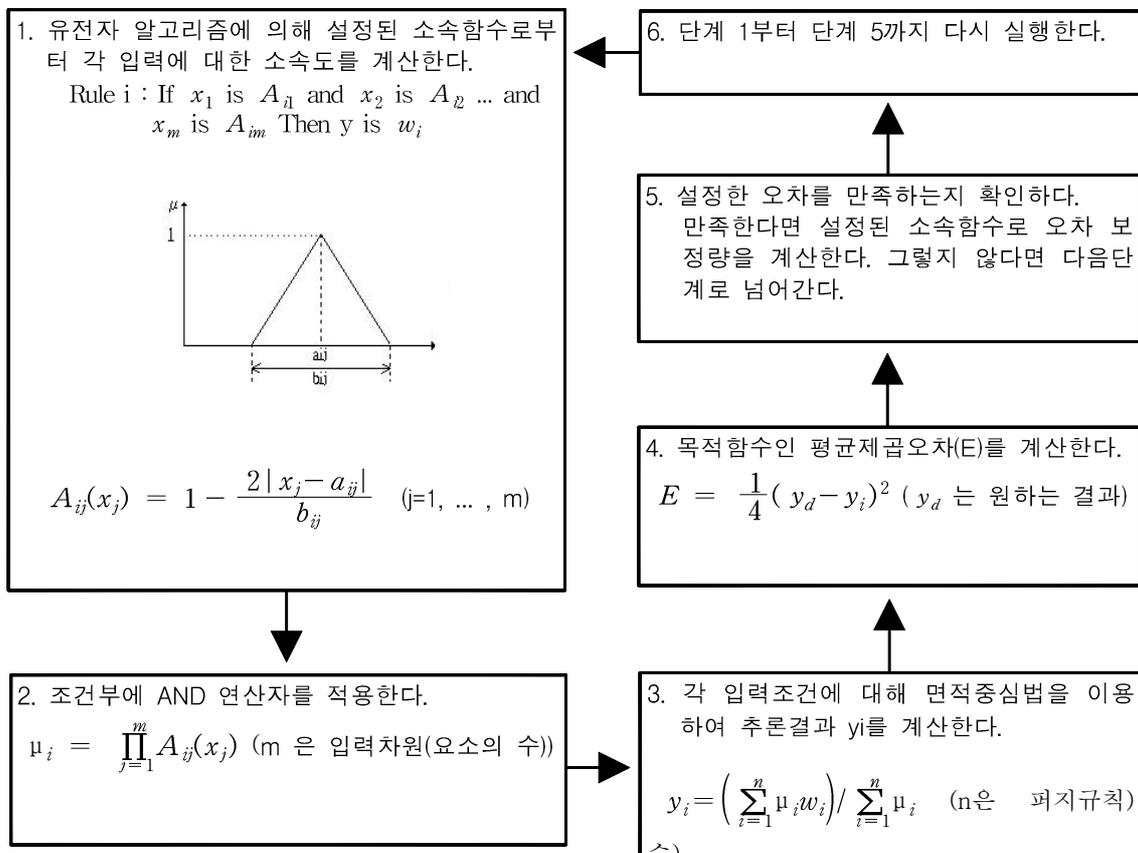


그림 1. 유전자 알고리즘을 이용한 퍼지제어 학습과정

4. Fuzzy-Grey 모형을 이용한 월유입량 예측

4.1 대상유역 및 학습자료 구성

본 연구의 적용대상 유역은 관측자료가 비교적 풍부하고, 각 수계별 다목적댐인 소양강댐, 안동댐, 남강댐, 대청댐, 충주댐에 비해 유역이 작은 섬진강유역을 선정하였다. 적용연도는 5~6년 빈도로 최저 유입량을 보이는 1982년, 1988년, 2001년을 대상으로 하였으며, 1982년 월유입량 예측부터 순차적으로 적용하였다.

GM(1,1)모형이 가지는 오차를 보정하기 위해서 1981년 9월부터 1982년 11월까지 연속된 4개월의 입력자료와 유사한 패턴을 가지는 1975년부터 해당 입력연도까지의 과거자료를 Grey 상관도를 통해 관측값에서 찾은 후 유전자 알고리즘의 탐색공간을 제한하고 오차의 방향이 일정한 패턴을 선별하기 위해 현재 입력 자료에 대한 과거 패턴 자료의 놈(norm)을 구하여 퍼지제어 학습을 위한 자료를 표 1과 같이 구성하였다.

표 1. 월별 학습자료(1982년)

	1월	2월	3월	4월
패턴1	79. 09~79. 12	79. 10~80. 01	76. 11~77. 02	79. 12~80. 03
패턴2	80. 09~80. 12	76. 10~77. 01	77. 11~78. 02	76. 12~77. 03
패턴3	75. 09~75. 12	75. 10~76. 01	78. 11~79. 02	78. 12~79. 03
	5월	6월	7월	8월
패턴1	78. 01~78. 04	78. 02~78. 05	81. 03~79. 06	76. 04~76. 07
패턴2	79. 01~79. 04	77. 02~77. 05	76. 03~76. 06	77. 04~77. 07
패턴3	76. 01~76. 04	81. 02~81. 05	77. 03~77. 06	79. 04~79. 07
	9월	10월	11월	12월
패턴1	76. 05~76. 08	76. 06~76. 09	77. 07~77. 10	78. 08~78. 11
패턴2	81. 05~81. 08	78. 06~78. 09	76. 07~76. 10	76. 08~76. 11
패턴3	78. 05~78. 08	75. 06~75. 09	78. 07~78. 10	75. 08~75. 11

4.2 적용결과 및 오차분석

예측결과의 정확성은 두 가지 방법으로 평가되었으며, 첫 번째 방법은 상대백분율오차(RPE; Relative Percentage Error)로서 식 (10)으로 정의된다. 두 번째 방법은 평균제곱근오차(RMSE; Root Mean Square Error)가 사용되었으며 식 (11)과 같다.

$$RPE = \frac{|x^{(0)}(k) - \hat{x}_r^{(0)}(k)|}{x^{(0)}(k)} \times 100 (\%) \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x^{(0)}(k) - \hat{x}_r^{(0)}(k))^2}{n}} \quad (11)$$

여기서 $x^{(0)}(k)$ 은 관측값, $\hat{x}_r^{(0)}(k)$ 은 예측값이며 n은 자료의 수이다. 본 연구에서 설명한 이론과 구성한 적용절차를 1982년 이외의 가뭄연도인 1988년, 2001년에 적용하여 같은 방법을 통해 표 2와 같이 월유입량을 예측하였으며 예측결과의 오차를 표 3에 나타내었다.

표 2. Fuzzy-GM(1,1) 예측결과 및 비교 (단위 : CMS)

구 분		1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
1982년	관측값	2.10	1.40	4.80	9.20	5.00	3.70	7.80	44.70	3.00	0.9	4.2	13.6
	GM(1,1)	5.47	2.58	5.28	7.33	27.44	30.84	11.04	27.78	86.75	47.78	5.86	1.80
	Fuzzy-GM(1,1)	2.70	4.04	3.94	20.02	8.10	7.74	23.44	41.06	6.02	0.54	2.86	5.48
1988년	관측값	2.10	2.40	6.40	4.70	7.70	5.80	32.90	5.40	3.60	0.60	0.50	0.90
	GM(1,1)	10.16	12.67	15.01	14.86	15.12	10.23	21.15	57.98	46.13	5.42	7.36	6.72
	Fuzzy-GM(1,1)	9.43	7.48	10.46	7.25	5.94	16.63	40.19	3.44	6.18	5.42	2.36	6.17
2001년	관측값	4.40	12.80	14.4	1.70	1.6	27.7	59.50	3.60	2.00	3.40	1.30	2.60
	GM(1,1)	2.48	5.07	20.12	37.29	30.11	20.82	33.08	101.2	64.21	16.4	8.98	7.52
	Fuzzy-GM(1,1)	4.32	13.92	16.59	16.72	7.52	29.15	49.76	14.64	7.99	5.75	4.52	1.72

표 3. 예측결과에 대한 RPE와 RMSE

RPE	1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
1982년	28.57	188.57	17.92	117.61	62.00	109.19	200.51	8.14	100.67	40.00	31.90	59.71
1988년	349.05	211.67	63.44	54.26	22.86	186.72	22.16	36.30	71.67	803.33	372.00	585.56
2001년	1.82	8.75	15.21	883.53	370.00	5.23	16.37	306.67	299.5	69.12	247.69	33.85
평균	126.48	136.33	32.19	351.80	151.62	100.38	79.68	117.04	157.28	304.15	217.20	226.37
RMSE	4.25	3.37	2.71	10.79	3.99	6.73	11.44	6.81	4.15	3.10	2.28	5.61

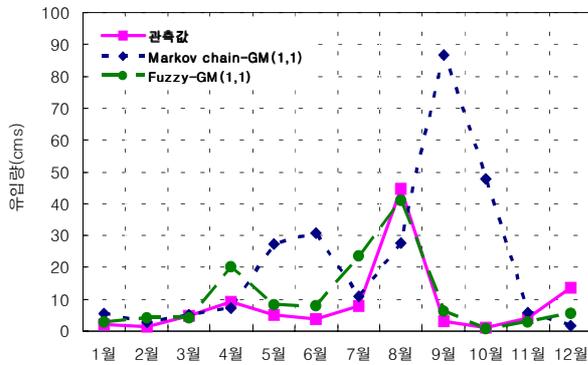


그림 2. Fuzzy-GM(1,1) 예측결과(1982년)

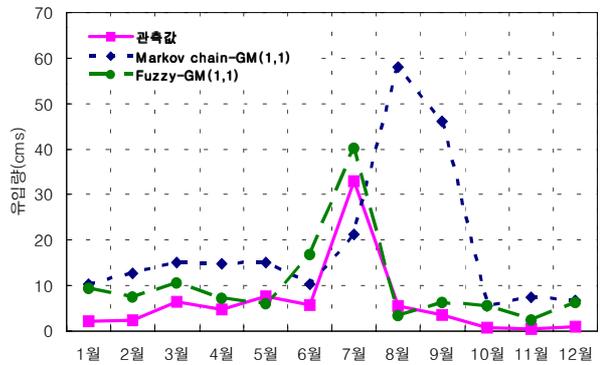


그림 3. Fuzzy-GM(1,1) 예측결과(1988년)

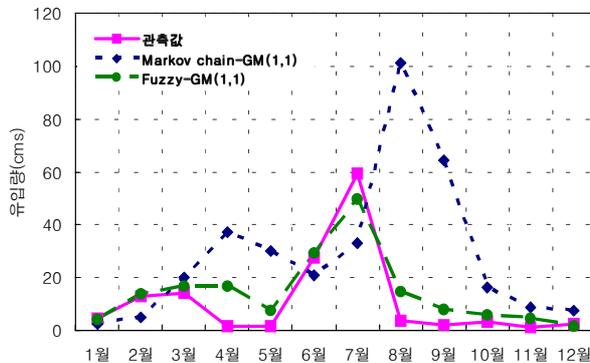


그림 4. Fuzzy-GM(1,1) 예측결과(2001년)

5. 요약 및 결론

본 연구에서는 Grey 모형이 시간에 대한 1차미분식으로 구성되어 있어 연속된 4개의 자료의 증감, 즉 1-AGO에 의한 누적자료가 가지는 기울기에 영향을 크게 받기 때문에 월별 유입량 변동이 상대적으로 적은 가뭄연도에만 적용하였다. 적용결과에서 알 수 있듯이 급격한 유입량 변화를 일으키는 경우 오차가 크게 발생하였으나 비교적 전체적인 월유입량 추세를 잘 예측하고 있는 것으로 판단된다. 앞으로 많은 자료가 구축되면 보다 많은 패턴자료의 확보가 가능하므로 충분한 오차보정이 가능할 것이며, 유입량에 직접적인 영향을 주는 강우량과 연계한 2변수의 Fuzzy -Grey 모형을 이용한다면 보다 정확한 월유입량 예측이 가능할 것으로 사료된다.

참 고 문 헌

1. 건설교통부 (2002). 2001년 가뭄기록조사 보고서.
2. Deng, J. L. (1989). "Introduction to Grey System Theory." *The Journal of Grey System*, Vol. 1, pp. 1-24.
3. Huang, Y. P., Huang, C. C. (1996). "The integration and application of fuzzy and grey modeling methods." *Fuzzy sets and Systems*, Vol.78, pp. 107-109.
4. Yu, P. S., Chen, C. J., Chen, S. J., Lin, S. C. (2001). "Application of Grey Model toward runoff forecasting." *Journal of the American Water Resources Association*, Vol. 37, No. 1, pp. 151-166.
5. Yu, P. S., Chen, C. J., Chen, S. J. (2000). "Application of Gray and Fuzzy Methods For Rainfall Forecasting." *Journal of Hydrologic Engineering*, Vol. 5, No. 4, pp. 339-345.