

遺伝的プログラミングを活用した洪水予測技術の開発 ～草木ダム及び佐波川におけるケーススタディ～

DEVELOPMENT OF FLOOD FORECASTING TECHNOLOGY USING GENETIC PROGRAMMING ~CASE STUDY OF KUSAKI DAM AND SABA RIVER~

渡辺直樹¹・山根総一郎¹・今村仁紀²・深見和彦³

Naoki WATANABE, Soichiro YAMANE, Hitoki IMAMURA, Kazuhiko FUKAMI

¹正会員 工修 JFEエンジニアリング株式会社 産業機械エンジニアリング事業部 ITソリューション部
(〒230-8611 神奈川県横浜市鶴見区末広町2-1)

²正会員 独立行政法人土木研究所 水災害・リスクマネジメント国際センター 水文チーム
(〒305-8516 茨城県つくば市南原1-6)

³正会員 工修 独立行政法人土木研究所 水災害・リスクマネジメント国際センター 水文チーム
(〒305-8516 茨城県つくば市南原1-6)

Promotion of flood forecasting system is crucial for reducing flood disasters urgently and effectively, especially in medium- and small-scale rivers. The authors propose a new method of flood forecasting using Genetic Programming (GP) and GMDH. Traditional method based on physical model takes time to analyze the basin and river channel characteristics, but the new method has potential to make a rainfall-runoff (or rainfall-water level) model automatically by learning the past data of river water level or dam inflow and rainfall.

Case studies were conducted for the dam inflow prediction at the Kusaki Dam and the water-level prediction at Shinbashi point of the Saba River. The results from both the case studies were encouraging to promote the new method, because the inflow or water-level predictions with 6-hour lead time were relatively good. Some results suggested the necessity of the direct incorporation of spatial distribution of rainfall in the upstream area for better accuracy.

Key Words : flood forecasting, water level, dam inflow, Genetic Programming, GMDH

1. はじめに

近年、全国各地で多発する局地的な集中豪雨による河川氾濫災害が頻発しており、国土交通省や都道府県では中小河川を含めた総合的な治水対策を推進している。その中で、洪水予警報システムの整備は喫緊の課題の一つであり、その整備を早急に進める上で、迅速かつ効率的なシステム構築を可能とする手法の確立が必要と考えられる。

本研究では人工知能技術の一つである遺伝的プログラミングを活用し、洪水予測モデルを迅速かつ効率的に構築できる手法を試みた。従来の貯留関数法等の流出解析モデルでは、長期にわたる流域内の雨量分布や河川流量のデータベースが不可欠であるとともに、モデル定数決定に高度なノウハウと多大な時間・労力を要することが問題となっているが、人工知能を活用することで物理的なモデルを想定することなく自動的にモデルを構築することが可能になれば、洪水予測システム構築の人的・

時間的コストを大幅に削減できると期待される。適用検証事例として、利根川水系渡良瀬川上流に位置する草木ダム流入量予測と、山口県にある佐波川の水位予測の研究成果を紹介する。

なお、本成果は独立行政法人土木研究所とJFEエンジニアリング株式会社の共同研究によるものである。

2. 遺伝的プログラミングによるモデル構築手法

の概要

本論文では物理的なモデルを想定することなく、降水量とダム流入量あるいは水位の実績データを学習することにより自動的にその因果関係を記述するモデルを同定する。

従来の学習手法としては、ニューラルネットワークを用いた流出解析手法¹⁾⁻²⁾が知られているが、ネットワーク構造の設計に人的試行錯誤が必要であることに加え、

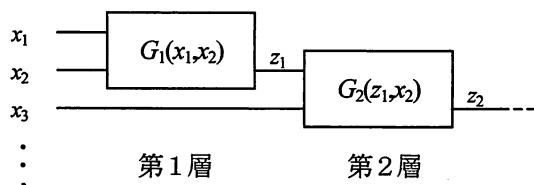


図-1 GMDHの重回帰ユニット構造

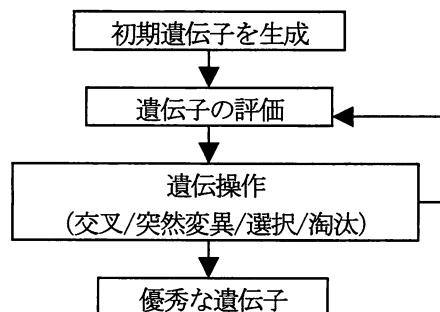


図-2 遺伝的プログラミングの計算フロー

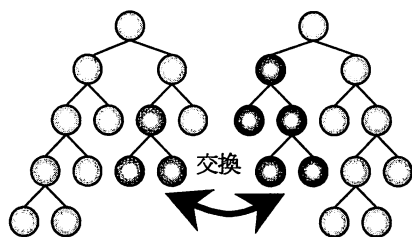


図-3 遺伝操作の一例 (交叉)

パラメータの決定に非線形最適化手法を用いているため計算の収束に時間がかかり、収束した場合も局所解であることが多いといった課題がある。

このような従来手法の課題を解決するモデル構築手法として、本論文では伊庭ら³⁾による、GMDH (Group Method of Data Handling) の演算構造を遺伝的プログラミング (Genetic Programming: GP) と呼ばれる生物進化を模擬した手法を用いて決定する方法を採用した。

GMDHは、重回帰ユニットを多層に構造化することを基礎とする非線形システム同定手法である。図-1にGMDHの重回帰ユニット構造を示す。ここで、 x は初期に与えられた入力要素で、 z は重回帰ユニットの出力である。1層目は入力要素 x のみでユニットが構成され、2層目以降は下層ユニットの出力 z あるいは入力要素 x によりユニットが構成されている。

重回帰ユニットの構成式は式(1)に示すような二次式で表され、各項の係数 $a_1 \sim a_6$ は各ユニットの入力データと教師データから最小二乗法により決定する。

$$G(x, y) = a_1 + a_2x + a_3x^2 + a_4xy + a_5y + a_6y^2 \quad (1)$$

重回帰ユニットの構造化手法として、遺伝的プログラミング (Genetic Programming: GP) を採用した。遺伝的プログラミングはネットワーク構造を遺伝子個体として表現し、重回帰ユニットの一つ一つを部分ノードとみなすことを基本とする最適化手法である。

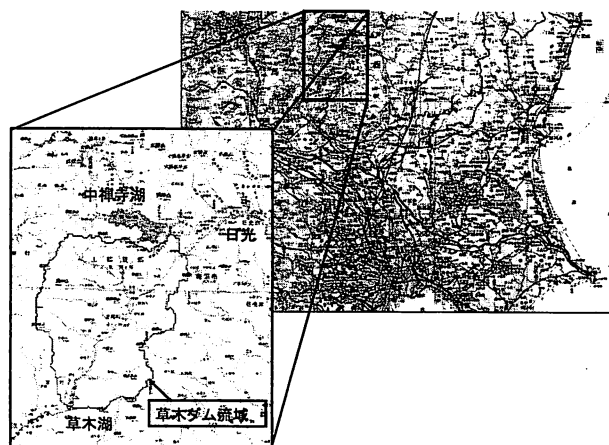


図-4 草木ダム流域位置図

図-2に遺伝的プログラミングの計算フローを示す。ランダムに生成された初期の遺伝子個体に対して交叉、突然変異等の遺伝操作を繰り返し行うことで優秀な遺伝子個体すなわち最適なGMDHのネットワーク構造を探ることができる。また、図-3に遺伝操作の一つである交叉の例を示す。これは優秀な遺伝子個体間で遺伝子ノードの一部を相互に交換することで、より優秀な遺伝子個体を生成することを狙ったものである。

3. 草木ダム流入量予測への適用

遺伝的プログラミングによる流出解析モデル構築手法を草木ダムの流入量予測に適用した例を示す。

(1) 流域概要

草木ダムの流域は、利根川水系渡良瀬川の上流に位置し、流域面積は254km²である。解析対象地点を図-4に示す。

(2) 流入量予測モデルの構築

解析モデルを式(2)のように定義し、学習によりモデル関数 f を同定した。なお、雨量の過去参照時間は水位と雨量の相互相関が最大となる時間差とした。

$$Q(t+1) = f(Q(t), R(t+1), \dots, R(t-8)) \quad (2)$$

ここで、 Q :流入量、 R :流域平均雨量、 t :時間である。

ここで求まる1時間後の予測流量を次の時間ステップの予測流量算出のための $Q(t)$ として繰り返し用いて計算することで、雨量予測値さえ与えれば原理的には何時間先までも予測できる。そこで、式(2)を連続的に適用し、6時間先の流入量まで予測した上でそこでの予測誤差を評価することにする。ここで雨量予測値としては実績雨量を用いており完全予測を仮定するが、この評価手法により遺伝的プログラミングにより最適化された流出解析モデルそのものの精度をより厳しい条件でチェックする

表-1 草木ダムの出水イベント

No.	出水日時	総雨量 (mm)
学 習	1998-09-16 7:00 &1998-09-21 23:00	273.6 196.7
	1990-08-10 15:00	328.2
	1985-07-01 5:00	200.92
	1991-08-20 23:00	307.2
	1990-09-20 4:00	212.4
	1983-08-17 12:00	386.24
	1997-06-20 16:00	166
	1994-09-30 6:00	296.6
	1989-08-27 17:00	238.1
	1991-08-31 10:00	176.9
	1991-09-19 14:00	295.3
	1990-11-30 22:00	162.3
	1986-09-03 8:00	235.66
	1998-08-30 14:00	434.7
	1985-08-05 2:00	149.24
	1993-11-14 2:00	138.6
	1999-06-30 12:00	218.9
	1994-05-27 7:00	177.6
	1998-10-18 9:00	154.1
	1994-09-17 19:00	287.4
未 知	2001-08-22 11:00	348.2
	2001-09-10 22:00	554.5
	2002-07-11 3:00	413.7
	2003-08-09 16:00	261
	2002-10-01 21:00	149.1
	1999-08-14 22:00	296.1
	1999-07-12 17:00 &1999-07-14 5:00	156.8 189.6
	2004-06-21 20:00	85.4
	2005-08-26 1:00	204
	2000-09-12 14:00	227.6
	2004-09-30 5:00	107.1
	2005-07-26 22:00	177
	2004-10-09 20:00	288.1
	2004-10-21 0:00	118.9
	1983-09-29 0:00	117.38
	2001-10-11 1:00	123.3

ことができる。

解析に用いたデータは1983年から2005年までに観測されたピーク流入量 $200\text{m}^3/\text{s}$ 以上の出水36ケースとし、各ケースは流域平均雨量とダム流入量の時間毎データで構成される。検証では、まず前半1983年～1998年の20ケースを学習して解析モデルを構築し、次にモデルにとっては未知となる後半1999年～2005年の16ケースを予測した。表-1に解析に用いた出水イベントを示す。各ケースNo.はピーク流量の高い順に表示した。

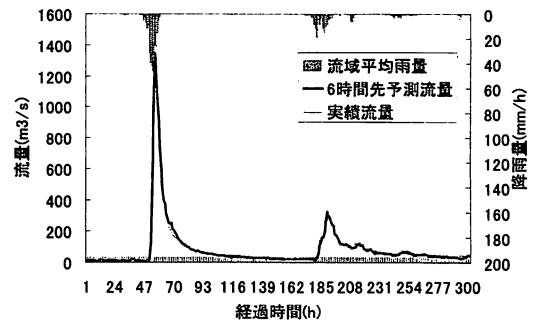


図-5 草木ダムの学習トレース結果（学習ケースNo. 1）

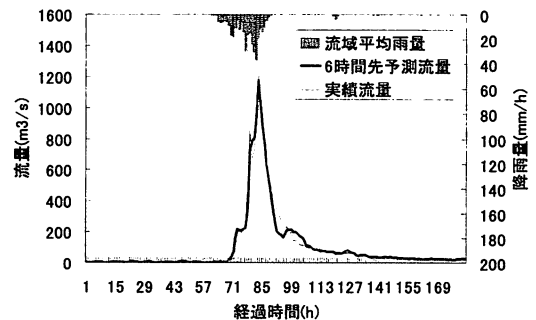


図-6 草木ダムの未知予測結果（未知ケースNo. 1）

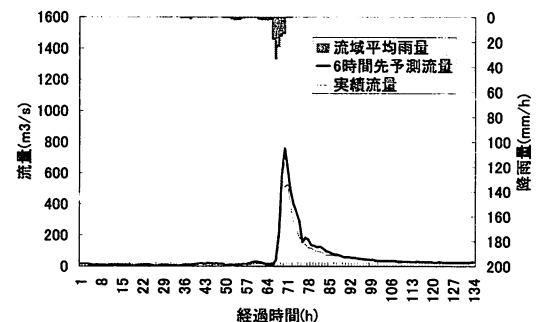


図-7 草木ダムの学習トレース結果（学習ケースNo. 8）

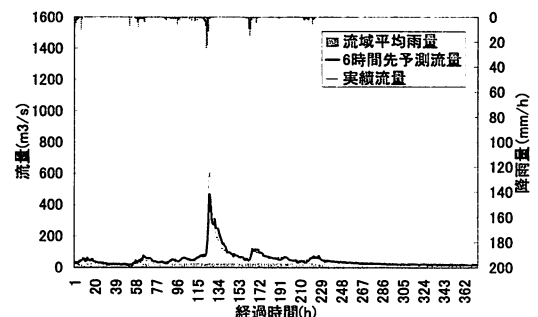


図-8 草木ダムの未知予測結果（未知ケースNo. 5）

(3) 解析結果

学習データを再現計算した学習トレース結果と、未知データを予測した結果のうち最大出水の計算結果を図-5、図-6に示す。結果は6時間先の予測流入量と実績流入量とを併記している。なお6時間先までの入力雨量は実績値を用いた。図-5の学習トレースの例ではピークの実績流量 $1347.14\text{m}^3/\text{s}$ に対し予測流量が $1349.6\text{m}^3/\text{s}$ と誤差は0.2%に止まり、図-6の未知予測ではピークの実績流量

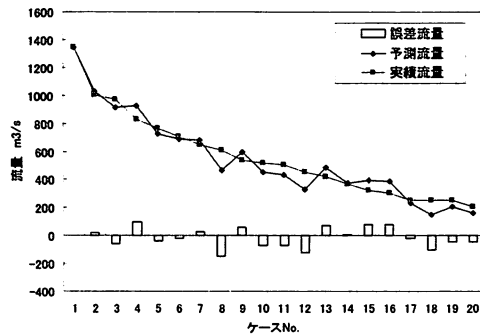


図-9 草木ダムピーク予測精度評価（学習トレース）

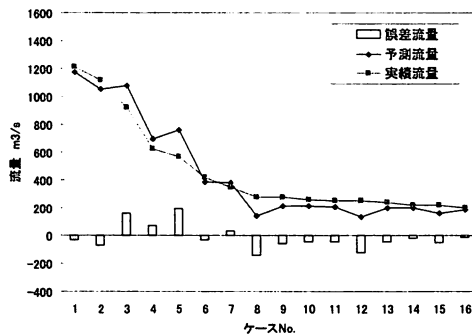


図-10 草木ダムピーク予測精度評価（未知予測）

1210.13m³/sに対し予測流量が1176.5m³/sと誤差は2.7%に止まっている。いずれの計算結果も、良好な精度で予測できた事例である。

次に誤差が最も大きくなった事例を図-7、図-8に示す。学習トレースの例ではピークの実績流量611.2m³/sに対し予測流量が465.3m³/sと誤差は23%、未知予測ではピークの実績流量565m³/sに対し予測流量が761.2m³/sと35%の誤差が生じている。本解析では流域平均雨量を用いているため、局地的な雨量の影響を捉えられないケースで比較的大きな誤差が発生したと考えられる。

全ケースの解析結果について、洪水ピーク流量の予測値と実績値を比較した結果を図-9、図-10にまとめて示す。学習トレースの一部（学習ケースNo. 8, 12）、未知データの一部（未知ケースNo. 3, 5, 8, 12）を除く全てのケースで100m³/s以下のピーク誤差に収まった。

図-9、図-10の結果を基にピーク誤差のRMSE（平均二乗誤差）を調査した。なお、RMSEは式(3)により定義される。

$$RMSE = \sqrt{\frac{1}{n} \sum_i (\hat{x}_i - x_i)^2} \quad (3)$$

ここで、 $\hat{x}$ ：予測値、 x ：実績値、 i ：データ番号、 n ：データ総数である。

学習トレースのRMSEは69.6m³/s、未知予測のRMSEは87.6m³/sとなった。未知データの誤差が若干増加しているが、大幅な悪化ではなく、本解析モデルは十分な汎化能力を有していると考えられる。

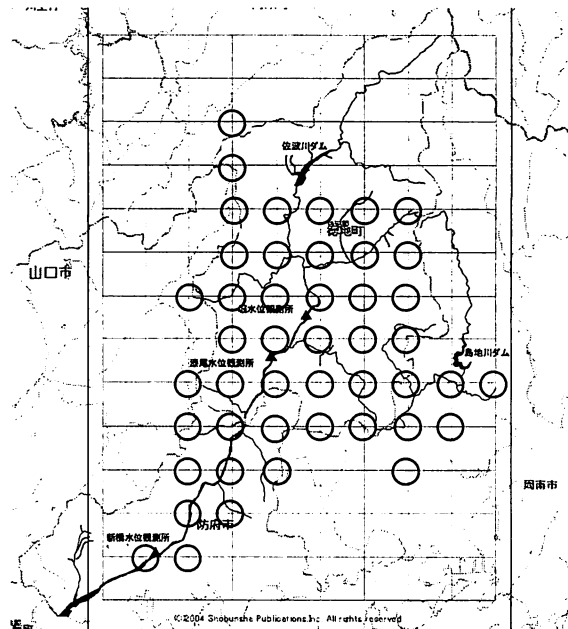


図-11 佐波川新橋観測所水位予測モデルにおける雨量入力メッシュ（レーダアメダス解析雨量）

4. 佐波川水位予測への適用

次に、佐波川の水位予測に適用した例を示す。従来の水位予測では流量評価を介する必要があるが、流量観測所の必要性や水位流量関係式の精度管理など、適用可能な河川を制限する要因となっている。遺伝的プログラミング手法を用いることにより流量を介さずに雨量から直接水位を予測するモデルを構築することが可能となれば、中小河川における洪水予測システムの普及に拍車をかけることができると期待される。

(1) 流域概要

佐波川は、山口県内を流れる流域面積423km²、幹川流路延長56kmの一級河川である。解析対象地点は下流に位置する新橋水位観測所とし、入力する雨量観測値としてはレーダアメダス解析雨量値（2.5kmメッシュ）を用いる。解析に用いた雨量メッシュを図-11に示す。流域上流には佐波川ダムと島地川ダムがあるため、本解析ではダムの放流量は既知のものとして与えた。

(2) 水位予測モデルの構築

解析モデルを式(4)のように定義し、学習によりモデル関数 f を同定した。なお、雨量及びダム放流量の過去参照時間は、水位と雨量、水位とダム放流量の相互相関が最大となる時間差とした。

$$H(t+1) = f \left(\begin{matrix} H(t), \dots, H(t-14), \\ R(t+1), \dots, R(t-7), \\ D(t+1), \dots, D(t-6) \end{matrix} \right) \quad (4)$$

ここで、 H ：新橋観測所水位、 R ：雨量、 D ：ダム放流量、

表-2 佐波川の出水イベント

	No.	出水日時	総雨量 (mm)	備考
学 習	1	2005-09-07 0:00	264	台風第14号
	2	2004-08-30 20:00	135	台風第16号
	3	2005-07-03 14:00	275	梅雨前線
	4	2004-08-01 17:00	132	台風第10号
	5	2004-06-26 11:00	172	梅雨前線
	6	2005-07-10 20:00	130	梅雨前線
	7	2004-09-07 18:00	104	台風第18号
	8	2004-08-02 21:00	50	熱帯低気圧
	9	2004-05-17 6:00	100	梅雨前線
	10	2004-05-13 13:00	81	梅雨前線
	11	2004-10-20 23:00	129	台風第23号
	12	2004-09-24 7:00	69	秋雨前線
	13	2004-09-29 18:00	53	台風第21号
未 知	1	2006-06-26 9:00	166	梅雨前線
	2	2006-04-11 10:00	149	前線
	3	2006-05-10 21:00	84	前線
	4	2006-06-23 7:00	146	梅雨前線
	5	2006-07-01 9:00	94	梅雨前線
	6	2006-07-02 11:00	34	
	7	2006-07-05 4:00	67	梅雨前線
	8	2006-07-20 18:00	105	梅雨前線
	9	2006-06-15 11:00	77	梅雨前線
	10	2006-08-19 12:00	99	台風第10号
	11	2006-09-10 1:00	88	秋雨前線

t: 時間である。

草木ダム流入量予測検討と同様に、雨量は完全予測を仮定し、1時間毎の予測を6回連続実施し6時間先予測値としての評価を行う。まず2004年から2005年のデータを学習して解析モデルを構築し、同モデルにとっては未知となる2006年のハイドログラフを予測した。表-2に解析に用いた出水イベントを示す。各ケースNo. はピーク水位の高い順に表示した。

雨量データ入力手法を比較検討するため、草木ダムの事例と同様に図-11の領域の平均雨量を与えた場合と、レーダメッシュ雨量を直接入力する場合の2手法で水位予測モデルの構築を行った。

(3) 解析結果

学習データを再現計算した学習トレース結果と、未知データを予測した結果のうち最大出水の計算結果を図-12、図-13に示す。結果は6時間先までの予測水位と実績水位を併記している。なお、6時間先までの入力雨量は実績値を用いた。

全ケースの解析結果について、メッシュ雨量/平均雨量別、予測先行時間別のRMSEを図-14、図-15にまとめた。RMSEは水位レベルごとに算出し、最低水位から0.35mま

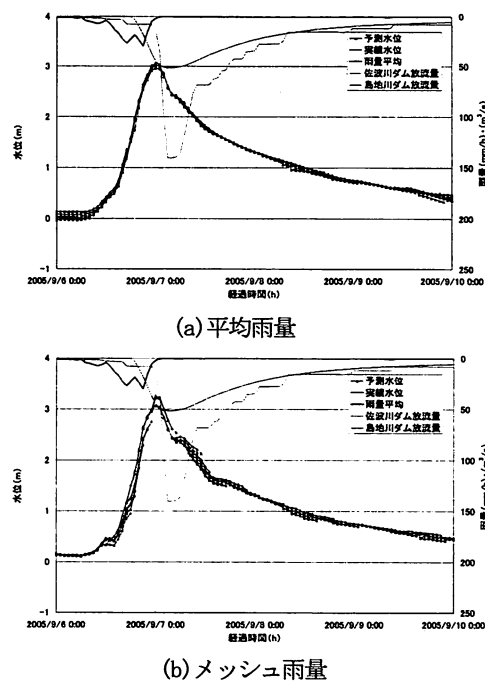


図-12 佐波川の学習トレース結果 (学習ケースNo. 1)

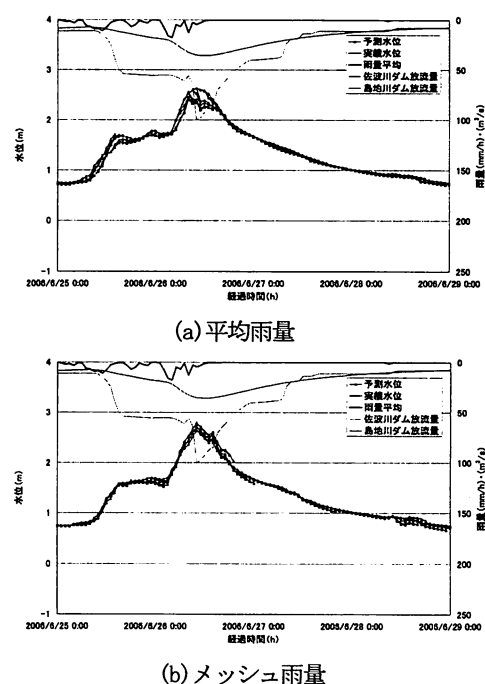


図-13 佐波川の未知予測結果 (未知ケースNo. 1)

でを低水位、0.35mから2.14mを中水位、2.14m以上を高水位として分類した。結果では、予測時間が大きくなるほどRMSEが大きくなることが分かる。また、高水位ほどRMSEが大きくなることが分かる。更に、学習トレースでは全ての水位レベルにおいて平均雨量の方がRMSEが小さいが、未知予測においては逆にメッシュ雨量の方がRMSEが小さくなっている。これは、学習データでは平均雨量を用いて十分な学習ができたものの、未知データの予測では雨量の情報量が不足したために予測精度が低下したためと考えられる。

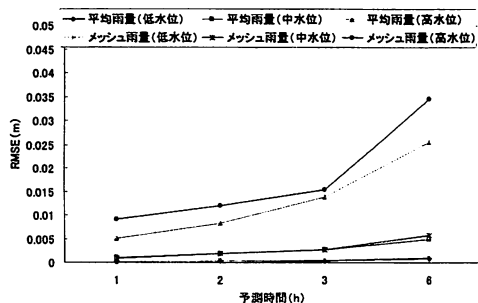


図-14 佐波川学習トレースのRMSE

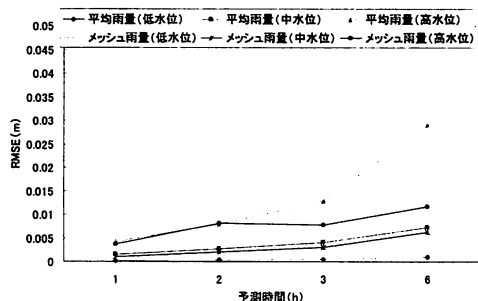


図-15 佐波川未知予測のRMSE

次に未知データの上位10出水について、ピーク水位の予測値と実績値を比較した結果を図-16に示す。メッシュ雨量を用いたケースでは±0.25m以内のピーク誤差に収まり、平均雨量を用いたケースでは±0.5m以内のピーク誤差に収まった。図-16を基にピーク誤差のRMSEを調査したところ、メッシュ雨量で0.14m、平均雨量で0.26mと、平均雨量の方が誤差が若干大きくなった。これは、新橋水位観測所の流域面積は400km²を超えており、メッシュ雨量を平均化したことにより雨量分布の違いが新橋地点の流量・水位波形の違いに与える情報が失われているためと考えられる。

5. おわりに

本報では、遺伝的プログラミングを用いて草木ダム流入量及び佐波川水位を事例とした予測モデルの構築・検証を試みた。その結果、比較的良好な精度でダム流入量及び水位を予測可能なモデルを構築できることを確認した。また、佐波川の事例において、雨量入力条件として

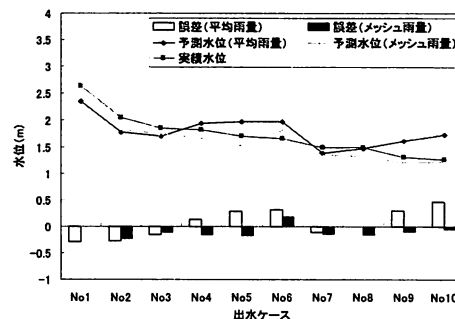


図-16 佐波川ピーク精度評価（未知予測）

平均雨量とメッシュ雨量を与えたケースを比較した結果、平均雨量では、一定の分布を前提とした上での学習がより確実に実施できる利点があるものの、雨量分布の情報が失われるため未知データにおける予測精度が低下する可能性があることを確認した。今後、雨量情報の最適な入力方法や外挿条件と言えるような極值的洪水事例における安定性等について検討していく予定である。

なお、本報告での解析はJFEエンジニアリング製の人工知能ソフトウェアWinmuSe^{®4})を用いて実施した。

謝辞：貴重な資料を提供して頂いた山口河川国道事務所、及び独立行政法人水資源機構草木ダム管理所に感謝の意を表します。

参考文献

- 1) 四俣正俊：ニューラルネットワークを用いた流出計算，水文・水資源研究のためのAI技術利用に関するシンポジウム論文集，水文・水資源学会，1992. 11.
- 2) 藤田睦博：ニューラルネットを用いた流出予測，水文・水資源研究のためのAI技術利用に関するシンポジウム論文集，水文・水資源学会，1992. 11.
- 3) 伊庭斉志，佐藤泰介：システム同定アプローチに基づく遺伝的プログラミング，人工知能学会誌，Vol. 10，No. 4，pp. 590-600，Jul. 1995.
- 4) JFEエンジ，” muSeでプラントの最適運転を支援”，エンジニアリング・ビジネス，Vol. 24 (22)，pp6-7，2004.

(2008. 4. 3受付)