

観測水位のみを用いた時系列解析による 洪水予測手法の適用性に関する研究

小山直紀*, 山田正**

A Study on Applicability of Flood Forecasting Method Based on Time Series Analysis Using Observed Water Level

Naoki KOYAMA* and Tadashi YAMADA**

Abstract

Forecast information of water level is important information that encourages residents to evacuate. Generally, flood forecasting is conducted by using runoff analysis. However, in developing countries, there are not enough rainfall observation stations in a basin. Therefore, this study assumes where poor hydrologic data basin and evaluates it through reproducibility and prediction by using time series analysis which statistical model with the only water level data. This method is possible to reproduce hydrograph, if the observation stations exist several points in the basin. And using the estimated parameters from past flood events, we can apply this method to predict the water level until the flood concentration time which the reference point and observation station

1 はじめに

近年、日本各地において、毎年のように洪水が発生し、多大な被害が発生している。また、日本に限らず、アジアにおいても、豪雨による洪水災害が発生し、人的・経済的被害が生じている。このような災害対策は、一般的にハード対策とソフト対策の両面から講じられているが、気候変動による豪雨の激甚化、高頻度化が懸念されている。そのため、従来では、「水害は施設整備により発生を防止するもの」、つまりハード対策による災害対策が考えられてきたが、今後は、「施設能力には限界があり、防ぎきれない大洪水は必ず発生するもの」と社会全体の意識変革の必要がある[1]とされている。そのため、ハード対策では防ぐことができないほどの大規模な水害時にはソフト対策で補う必要があり、今後ソフト対策の重要性がより一層高まっている。

ソフト対策の中でも水位の予測情報は、人命を守るための住民の避難行動を促す重要な情報である。一般的な予測手法は、流出解析モデルと統計モデルの2種類に分けられる。物理モデルである流出解析モデルは、降雨をインプットデータとするため、降雨の精度が重要である。しかし、新興国のような地域において水文観測の整備が十分にされておらず、必要な精度をもった降雨データが入手できないことが多くある。そのような降雨観測が不十分な地域においては、GSMaP(Global Satellite Mapping of Precipitation)のような衛星全球降水マップを用いることもある。しかし、GSMaPは、降雨データの無い地域においては非常に有効なものであるが、降雨が過小評価される傾向がある[2]ということが分かっており、依然として地上雨量計が降雨データの代表的なものとなっている。著者ら[3]は、降雨流出解析をする際に降雨データの空間解像度がピーク流量に及ぼす影響を定量的に評価しており、洪水ピーク流量を10%以内にするには雨量計1つ当たりの支配面積(流域面積/雨量計の個数)を10km²に1つの雨量計データが必要という結論を示しており、また、雨量計1つ当たりの支配面積が約100km²の場合には、流域面積により5%~50%程度の洪水ピーク流量に幅があることを示している。ここで、新興国の1つであるベトナムを例に、地上雨量計による観

* 中央大学大学院理工学研究科都市人間環境学専攻 〒112-8551 東京都文京区春日1-13-27

** 中央大学理工学部都市環境学科 〒112-8551 東京都文京区春日1-13-27

測整備状況をみると、国土面積(約362, 215km²)に対して、雨量計が755台[4]であり、雨量計1つ当たりの支配面積は約437km²/個となっている。流域により雨量計の個数は異なるが、算出される流域平均雨量には大きな不確実性があることが予想される。流域内の雨量観測整備状況に依存する降雨データに対し、その地点を代表とする河川水位は比較的精度良く観測が行うことができ、代表性の高いものである。また、物理モデルによる流出解析では、出力結果が流量で出てくるため、H-Q曲線により水位に変換する際に誤差が生じることもある。

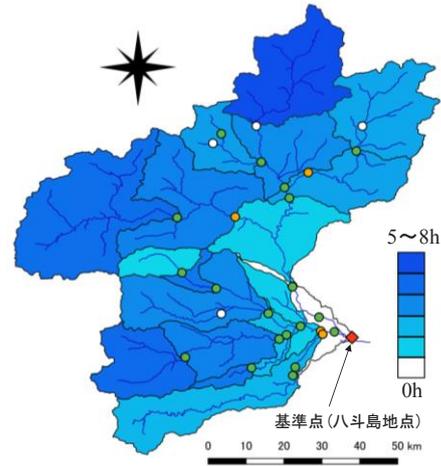


Fig-1 Upper Tone river basin with the location of Water Level Obs. Station, river network and Yattajima points.

近年、統計的手法による洪水予測では、ニューラルネットワークを用いた多くの研究[5]が進められている。この手法では、非線形性を考慮できるが、多くのデータが存在する場合に精度の高いモデルを構築できると一般的に言われている。それに対し、時系列解析は、モデル構造が単純なため、データ量が少ない際においてもある程度の精度が出ることが多いと言われている。そこで本稿では、水文データが乏しい流域にも適用できる洪水予測を目的に、統計モデルの1つである時系列解析から水位のみを用いて、水位予測手法の構築、水位予測精度の検証及び予測値の安定性に関する検証を行った。

2 解析手法

時系列解析は、統計モデルの1つであり、現象を過去の動きと関連付け、現象の予測や制御のために用いられている。最も基本的な時系列解析に用いられている自己回帰モデル式を(1)式に示す。

$$h_n = \sum_{i=1}^N a_i h_{n-i} + \varepsilon_n \quad (1)$$

(1)式は、水位観測値の時系列 h_n を過去の時系列 h_{n-i} にパラメータ a_i を掛け合わせたものとホワイトノイズ ε_i の線形和で表したものである。 N は次数であり、考慮する過去のデータ数である。パラメータ a_{li} は、過去の観測値 h_{n-i} の重みであることから a_{li} の大きさによって、 h_n に対してどの程度過去の観測値が影響しているかを表している。しかし河川の場合、水位はある1地点で決まるものではなく、上流の様々な支川の河川水が集まり、ある水位に決まる。そのため、ある地点における水位は、相互に関係していると考えることができる。そこで、自己回帰モデルを拡張し、それぞれの支川の影響を含めたものを表現した式が(2)式である。

$$\begin{bmatrix} h_n^1 \\ h_n^2 \\ \vdots \\ h_n^p \end{bmatrix} = \sum_{i=1}^N \begin{bmatrix} a_{11}^i & a_{12}^i & \cdots & a_{1p}^i \\ a_{21}^i & a_{22}^i & & a_{2p}^i \\ \vdots & & \ddots & \vdots \\ a_{p1}^i & a_{p2}^i & \cdots & a_{pp}^i \end{bmatrix} \begin{bmatrix} h_{n-i}^1 \\ h_{n-i}^2 \\ \vdots \\ h_{n-i}^p \end{bmatrix} + \begin{bmatrix} \varepsilon_n^1 \\ \varepsilon_n^2 \\ \vdots \\ \varepsilon_n^p \end{bmatrix} \quad (2)$$

(2)式はベクトル形式で表されたものであり、(2)式をテンソル形式で表したものが(3)式である。

$$\mathbf{h}_n^p = \sum_{i=1}^N \sum_{l=1}^P \mathbf{a}_{pl}^i \mathbf{h}_{n-i}^l + \varepsilon_n^p \quad (3)$$

(3)式は(1)式をより一般化したものであり、多変量自己回帰モデルやベクトル自己回帰モデルと呼ばれているものである。ここに、 N :モデルの次数、 P :水位観測所数、 \mathbf{q}_n^p : n 時刻における p 地点の水位、 \mathbf{a}_{pl}^i : i 時刻前における p 地点観測値に対する l 地点観測値の影響である。この多変量自己回帰モデルは、工学分野において赤池・中川[6]が、セメントキルプロセスの制御において先駆的に用いている。

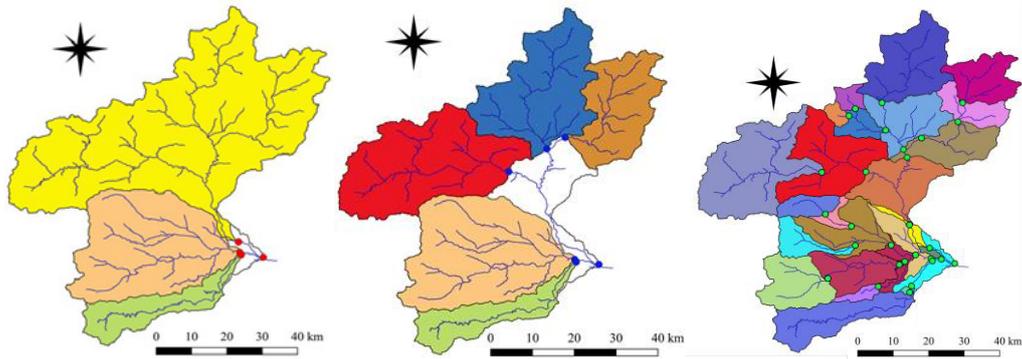


Fig-2 Basin map divided by water level stations(4, 6, 21 points)

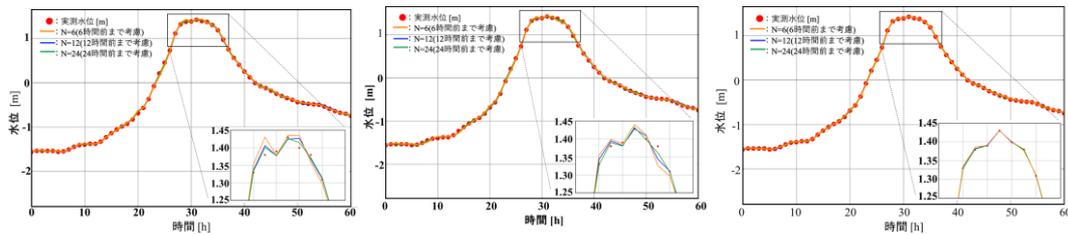


Fig-3 Reproduction results of flood events(Water level) (4, 6, 21 point)

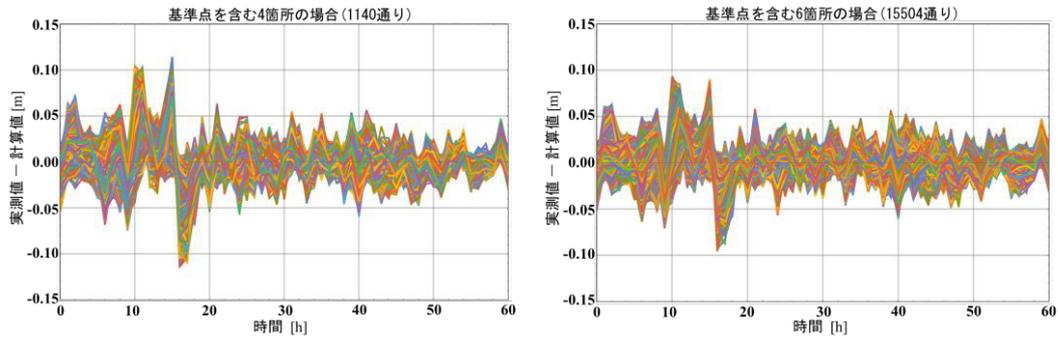


Fig-4 Differences between observed and calculated water level
(Left: using 4 points including the reference point,
Right :6 points including the reference point)

3 対象流域

対象流域は、利根川八斗島地点を下流端した利根川上流域とした。流域面積は5100km²であり、河口から181.45kmの地点に基準点である八斗島がある。Fig-1は、流域図であり、それぞれの点は水位観測所を表しており、全21箇所である。また、サブ流域ごとのそれぞれの色は、観測所ごとに流域分割し、過去の洪水イベントデータから求めた八斗島地点とそれぞれの観測地点のピーク水位の時間差を表している。パラメータの推定に用いた洪水イベントごと時間差の平均値で色分けしており、濃青色ほど時間差があることを示しており、基準点から最も遠い地点で5時間~8時間程度である。また、それぞれの水位観測データは1時間ごとのデータである。

4 再現計算によるモデルの検証

本手法による洪水予測を行う前に過去の洪水イベントにおいて再現計算を行い、本手法の妥当性の検討を行う。(3)式によるモデルを構築するためには、使用する観測所数 P と過去の観測データをどこまで考慮する

か次数 N を決定する必要がある。そこで、はじめに使用する観測所数による再現精度の確認を行った。使用した観測所数は、対象地域である利根川上流域の下流端である八斗島を含む4地点(大きな支川のみを考慮する場合)、6地点(中程度の支川も考慮する場合)、21地点(全ての観測所)の3つのパターンで計算を行った。合流部下流や合流前、ある1つの河川であればある程度長い区間を受け持った場合に水位計が設置してあることからこのような計算条件にした。それぞれ使用した観測所数で流域分割をしたものがFig-2である。次に、3つのパターンにおいて、洪水イベントの再現精度の検証を行った。

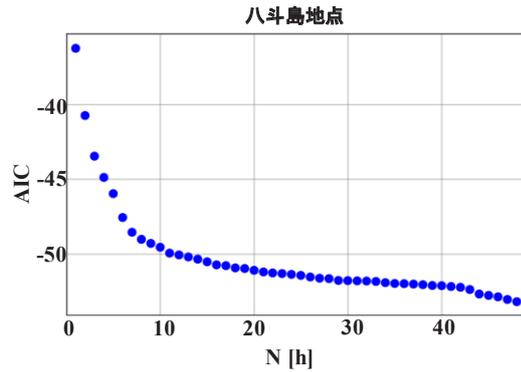


Fig-5 Relation of N and AIC

再現計算の対象にしたイベントは、2015年9月に起きた関東・東北豪雨のものを用いた。本イベントは、八斗島地点において水防団待機水位を超えており、ピーク流量は約4000m³/sである。そのイベントを含む1ヶ月分のデータをパラメータ推定のデータとした。Fig-3が推定したパラメータを用いて再現計算をした結果である。また、Fig-3のそれぞれにおける青線、緑線、赤線は、次数 N を変更した結果である。使用した観測所数が4,6,21地点すべての場合において、どの次数においても、観測値とそれぞれの線が重なっており、ほぼ傾向が捉えられていることから、観測所数の数に関わらず十分に再現できることが分かった。また、Fig-3のそれぞれに洪水ピーク水位付近の拡大したものを載せている。観測所数が4, 6において、ピーク付近で観測値と外れている点もあるが、数cmであり、十分に再現できると言える。また、全ての観測所を用いた場合には、用いた観測所数4, 6の場合と比較し、精度が向上していることが分かる。

これまでの計算は、流域の大・中の支川ごとにある観測値を用いた。しかし、水文データの整備されていない流域では、対象流域と同等のデータが得られない可能性もあり、ある支川には水位データがない等が考えられる。そこで、基準点である八斗島地点を含む4, 6地点をそれぞれ21地点ある観測所から重複のないように選択し、それぞれ1140, 15504通りの再現計算を行った。その結果がFig-4である。この図より、基準点を含む4,6において、ピーク水位約1.4mに対し、最大で計算値と観測値の差が11cm, 9cmであることから、それぞれピーク付近において、約8%, 6%であり良好な再現が得られている。このことから、対象流域の複数の観測地点を用いれば、精度良く再現できることが分かった。用いる観測所数の4, 6を比較すると微小ではあるが観測所数6を用いると精度が高い。そのため、今後の解析では、6を P として用いる。

次に(3)式における次数 N の決定は、赤池情報量基準 (AIC: Akaike's Information Criterion) [6]を用いて決定した。AICは(4)式で表される。

$$AIC = -2 \log(L) + 2K \quad (4)$$

ここに、 L : 最大対数尤度、 K : モデルのパラメータ数である。Fig-5が次数とAICの関係を示したものである。これより、10時間以上前を考慮しても、AICの値は大きく変化しないことがわかる。そこで、今後の解析では、時間軸を考慮して、半日である12を N とした。

5 時系列解析による洪水予測式の導出

前章において、過去の洪水イベントにおいて、再現計算を行うことでモデルの検証を行った。その結果よりどの観測所を用いても十分に再現できることが分かった。(3)式は、過去の値を用い、現在の値を求める式となっている。水位データはリアルタイムの値が知ることができると、現在の値を含めた式にすることにより、1時間先の値を求めることが可能である。1時間先を予測した式が(5)式である。

$$\hat{h}_{n+1}^p = \sum_{i=1}^N \sum_{l=1}^P \mathbf{a}_{pl}^i h_{n+1-i}^l + \varepsilon_n^p \quad (5)$$

さらに、 x 時間後までの予測を行うとするならば、(6)式となる。

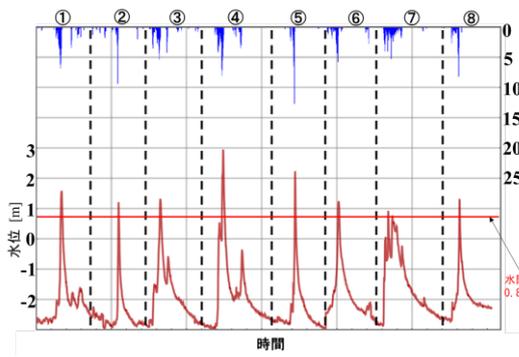


Fig-6 Floods event that exceeded the designated water level at reference point from 2002 to 2018(all 8 events).

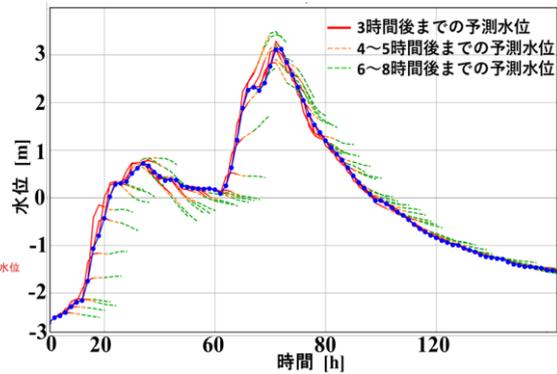


Fig-7 Result of water level forecast in flood event4 at reference point.

$$\hat{h}_{n+x}^p = \sum_{j=1}^{x-1} \sum_{l=1}^P a_{pl}^i \hat{h}_{n+x-j}^l + \sum_{i=1}^N \sum_{l=1}^P a_{pl}^i h_{n+1-i}^l + \varepsilon_n^p \quad (6)$$

(6)式を用いることにより、過去及び現在のデータから x 時間先までの予測を行うことが可能である。(5)式及び(6)式の \hat{h} は予測値であり、 h は観測値を表している。本論文では、洪水時の避難情報に役立てるため、8時間後までの予測を行った。

6 水位のみを用いた時系列解析による水位予測

6-1 水位予測の精度検証

前章において、洪水予測に用いる(6)式を導出した。本章では、水位のみを用いた洪水予測を行う。通常、流出解析では降雨をインプットデータとする。しかし、降雨データは、水文データ整備が乏しい流域においては、広域の代表性、信頼性が高いとは限らない。それに対し、水位データは、観測所より上流において降雨が流出したものであり、堰やダム等の水理構造物がない限り、大部分は下流部に流れるため、上流と下流の観測所における水位の相関は高いと考えられる。そこで、水文データが乏しい流域を想定し、水位のみから水位予測を行った。(6)式による水位予測を行うためには、パラメータ a_{pl}^i を決定させる必要がある。このパラメータは、観測値と(3)式により計算された再現の計算値の差が最も小さくなるように最小二乗法によって決定した。パラメータ推定に用いたデータは、2002年から2018年における出水期である5月から11月までの水位を抽出し、その中から八斗島地点における水防団待機水位を超えたものを1つの洪水イベントとした。その期間中に水防団待機水位を超えた洪水イベントは8回であった。Fig-6は、8洪水イベントを全て並べたものである。図中上の数字はそれぞれの洪水イベントを表している。(6)式による予測精度を評価するため、限られた洪水イベントのデータを全て用いる分割交差検定で行った。分割交差検定の手順は以下の通りである。①全対象洪水イベントをFig-6のように8個のデータセットに分割する。②8個のうち1個のデータセットを検証データとする。③残りの7個のデータからパラメータ推定用のデータセットとする。④設定したデータでパラメータ推定・精度検証を行う。これらの②、③、④を繰り返し、予測精度の評価を行った。Fig-7は、8時間先までの水位予測の結果の代表として最も水位が上昇した洪水イベント4を載せる。使用した水位計は、Fig-1で示すオレンジ色のものである。これは、水位計を設置するとき、支川や人口密集地などの重要な地点に設置されるため、それぞれの支川の代表観測点である観測所(Fig-1 オレンジ点)を使用した。この結果より、3時間後(赤線)までの予測値は、波形全体を通じて観測値のハイドログラフを再現できている。しかし、4時間後以降の予測値は、立ち上がりでは、過小に評価する傾向があり、水位の低減部では、過大に評価する傾向があることが分かる。

次に、予測精度の評価方法を説明する。避難するとき重要な時間であるハイドログラフにおける水位の立ち上がりからピーク水位までを用いた。その評価式は、2種類の方法を用いており、それぞれ次に式を

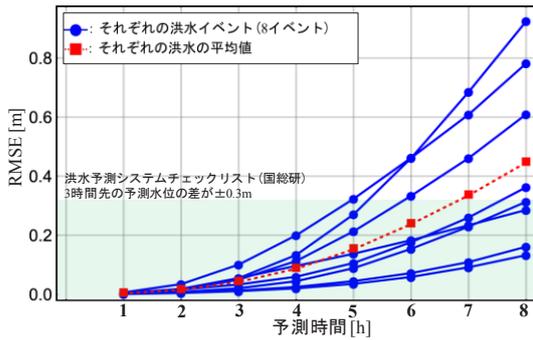


Fig-8 Difference between observed and predicted water level for each predicted time by RSME

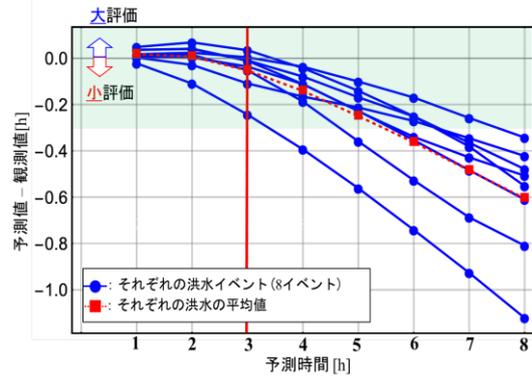


Fig-9 Difference between observed and predicted water level for each predicted time

示す。

$$RMSE = \sqrt{\sum_{n=1}^{N_{peak}} (h_{ob}^i - h_{pre}^i)^2 / N_{peak}} \quad (7)$$

$$RE = \sum_{n=1}^{N_{peak}} (h_{ob}^i - h_{pre}^i) / N_{peak} \quad (8)$$

(7)式は、平均二乗誤差である。しかし、この評価方法では、予測値が観測値と比較し、大きめに予測しているか、小さく予測しているか判断がつかない。これは、水位上昇期における予測値の傾向が安全側(観測値と比較して、大きめに予測)であるのか分からないということである。そこで、(7)式で評価すると同時に(8)式による評価を行った。Fig-8、Fig-9がそれぞれ(7)式、(8)式で精度評価を行った結果である。図中の緑色で塗りつぶされた部分は、それぞれの洪水予測手法で誤差が30cm以下のところを示している。これは、国土総合技術研究所の洪水予測の精度指標の目標例[8]を参考に決定した。Fig-8の青実線は、それぞれの洪水イベント(全8イベント)を示しており、赤の破線は全洪水イベントの平均値を示している。RMSEが30cm以内である部分は、8洪水イベントの平均値であると6時間程度であるが、洪水イベントによっては、4時間程度である。また、どのイベントにおいても予測時間が4時間あたりから傾きが大きくなっている。これはFig-1で使用している観測所と基準点との水位の時間差が大体同じであることから、水位の情報が伝わる時間以上では大きく精度が落ちるということを表している。次に予測時間により観測値に対する予測値の傾向を調べるために(8)式を用いたFig-8より、予測時間が3時間を超えたあたりから、予測時間が長くなるほど、観測値と比較して予測値は低めに予測することが分かる。これまでの計算では、Fig-1の中程度の支川の下流端である白色の観測所を用い、洪水予測精度はおおよそ3~4時間程度まで良く、これは観測所の位置による水位情報の伝達時間であることを述べた。そこで、今まで用いていた観測所より更に上流を用いることで、精度の良い予測時間を延ばすことができるのか計算を行う。Fig-10は、下流側と上流側の観測所を用いた場合の水位予測精度の比較である。赤線がFig-1の白点の観測所(下流側)を用いた結果であり、青線がオレンジ点の観測所(上流側)を用いたものである。この結果より、水位情報の伝わる時間の長い上流側の観測所を用いたほうが、3時間から5時間先の精度が向上していることが分かる。しかし、予測時間が7、8時間と長時間になると、観測所の位置による精度の違いはない。

6-2 水位予測の安定性に関する検証

これまで本手法による水位予測の精度検証方法として使用できるデータの全てを用いた。しかし、ある流域では、近年に水位計が設置されたため、大きな出水データの蓄積がされてないことも考えられる。そこで、パラメータを推定する際に用いるデータ数を変化させ、予測精度への影響分析を行った。その方法として、ある洪水イベントを水位予測するとき、パラメータ推定に用いる洪水イベントを1~6洪水と数を

変化させた。Table-1は、パラメータ推定に用いる洪水イベント数に対して作成できるイベント数である。

推定に用いるイベント数に応じて、作成できる数は異なるが、それぞれの洪水イベント数において、予測時間ごとに平均し、予測精度検証を行った。対象洪水は、最も水位が上昇した洪水イベント4を対象に行った。その結果がFig-11であり、それぞれの点は、それぞれの洪水イベント数であり、それを予測時間ごとに平均したものである。また、実線でつないでいるものは、それぞれの洪水イベント数における平均値である。この結果より、対象イベントでは、パラメータを推定するとき使用する洪水イベント数が少なく、予測時間が長いほど予測値に幅を持つことが分かる。また、その幅も観測値より大小両方の評価していることが分かる。しかし、その平均値としては、2イベント以上用いてパラメータ推定した場合は大きな変化がない。また、使用した洪水イベント数が6の時には、ほとんど予測値とその平均値が同じであることから、大きな出水が6イベントあれば、安定することが分かった。

7 おわりに

本研究では、水文データの乏しい地域を想定し、時系列解析の手法である時系列解析である多変量自己回帰モデルを適用させ、水位データのみによる洪水イベントの再現精度検証及び予測を行った。以下にそのまとめを示した後、今後の展望について述べる。

- 利根川上流域を対象にして、多変量自己回帰モデルの構築を行った。その結果、本研究の対象流域程度の大きさ(5100km²)ならば、流域内に水位観測点が数点あればピーク水位に対して、約±5%以内の再現性を持つモデルを構築することができる。また、基準点を含む数点以上の観測地点を用いることで精度の高い再現を行うことができる。

- 水位のみから水位予測を行った場合、高精度な予測を行うことができるのは、基準点と使用する観測所の水位の時間遅れ程度である。本流域の場合は、3~4時間程度である。また、基準点より、遠い地点の観測所を用いることで、長時間先に対しての予測精度が高くなる。

- パラメータ推定に用いる洪水イベント数を変化させ、洪水予測値の安定性を調べた結果、5、6イベント

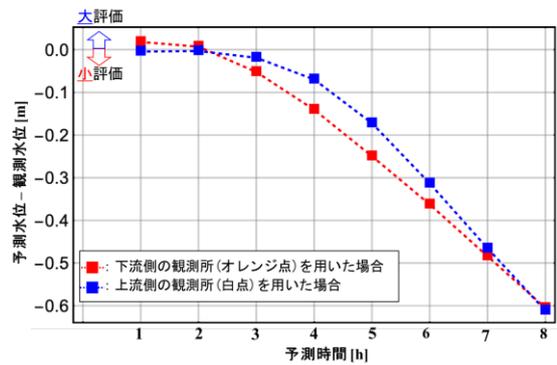


Fig-10 Comparison if prediction accuracy using stations located downstream and upstream

Table-1 The number of events for parameter estimation and number of data sets that can be created.

用いる洪水イベント数	作成できるデータセット数
1	$\gamma C_1=7$
2	$\gamma C_2=21$
3	$\gamma C_3=35$
4	$\gamma C_4=35$
5	$\gamma C_5=21$
6	$\gamma C_6=7$

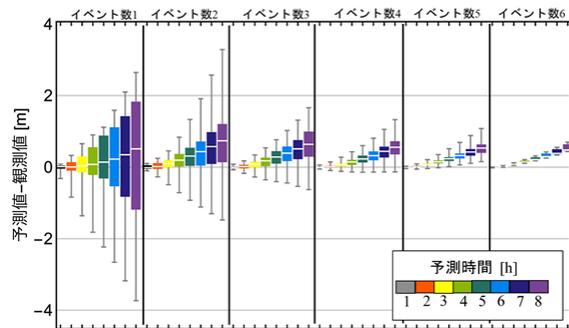


Fig-11 Stability of predicted values when the number of flood event used for parameter estimation is changed

あれば予測値が安定することが分かった。

本論文では、基準点の水位を上流の水位データから予測を行った。そのため、精度の高い水位予測を行える時間は、基準点と用いる観測所の水位の時間遅れ程度である。ここで降雨流出プロセスを考えると、ある流域に雨が降り、それが土壌に浸透・流出し、川に流れ、ある地点まで集ってくる。ある観測地点の水位情報は、この降雨流出プロセスの中では、最後の情報であり、一番の始まりは、降雨情報である。そこで、本手法に降雨情報を加えることにより、水位情報より前の情報を取り込むことができ、より精度の高い水位予測の長期化につながると考えられる。今後、さらに避難等に使える時間を増やすために、より先の時間においても精度の高い予測水位を行えるように本手法の拡張を行っていく予定である。

参考文献

- [1] 大規模氾濫に対する減災のための治水対策のあり方について ~社会意識の変革による「水防災意識社会」の再構築に向けて~, 社会資本整備審議会, 2015.
- [2] 例えば, 瀬戸心太, 芳村圭, 沖大幹: 高分解能衛星降水マップによる日本域の洪水検出能力, 水工学論文集, 第 52 巻, pp.355-360, 2008.
- [3] 小山直紀, 山田正: 降雨流出における降雨の空間分布の不確実性が河川流量・水位に与える影響に関する研究, 土木学会論文集 G(環境), Vol.75, No.5, pp.261-26, 2019.
- [4] VIETNAM CENTER OF HYDRO, <<http://cmh.com.vn/article/378-Rain-gauge-network.html>>, (入手 2019.2)
- [5] 例えば, 一言正之, 櫻庭雅明, 清雄一: 深層学習を用いた河川水位予測手法の開発, 土木学会論文集 B1(水工学), Vol.72, No4, I_187-I_192, 2016
- [6] 赤池弘次, 中川東一郎: ダイナミックシステムの統計的解析と制御, サイエンス社, 1972.
- [7] Hirotugu Akaike, A New Look at the Statistical Model Identification, IEEE Trans. Automat. Contrl., NO.6, pp.716-723, 1974.
- [8] 洪水予測システムチェックリスト(平成 21 年度版)【案】, 国土総合技術研究所, 2010.