

AIによる洪水予測と水位データの異常検知

岡田 治¹・柿澤 一弘¹・田所 正²・中安 正晃³

¹一般財団法人河川情報センター 情報基盤整備部 副参事

²一般財団法人河川情報センター 情報基盤整備部長

³一般財団法人河川情報センター 河川情報研究所研究第三部長

深層順伝播型ニューラルネットワークによる洪水予測技術を紹介する。深層順伝播型ニューラルネットワークは、浅層順伝播型ニューラルネットワークの隠れ層を多層としたモデルであり、複雑な洪水波形を学習する能力の獲得と洪水予測精度の向上が期待できる。また、ノイズ除去自己符号化器と呼ばれるニューラルネットワークによる水位データの異常検知技術を紹介する。

本稿では、上記に加え、降雨という自然現象において、降雨発生時刻は逆転しないという制約条件をニューラルネットワークに加える手法に関するアイデア、洪水到達時間、つまり流速の変動の影響を取り除き、洪水ピーク生起時刻の予測と実績との誤差を減らす手法に関するアイデアを提案する。また、水位の時系列データをウェーブレット変換し、いわゆる事前に特徴量抽出を行い、ニューラルネットワークが時系列データの特徴について学習する能力を向上させるアイデアを提案する。河川情報センターでは、GPU、CUDA、Python、Tensor Flow、Kerasを用いたプログラムを開発中である。以下これらの取り組みについて報告する。

Key Words : AI, 機械学習, ニューラルネットワーク, ノイズ除去自己符号化器, 洪水予測, 異常検知, ウェーブレット変換

1. はじめに

(1) 洪水予測技術の技術水準と課題

日本学術会議⁴²⁾は、2011年1月、国土交通省河川局長から「河川流出モデル・基本高水の検証に関する学術的な評価について」と題する依頼を受け、2011年9月1日、「これまで多数の流域で適用実績を持っていて信頼性がある貯留関数法を、ある程度、分布型のモデル形式にして利用すること、雨水流出現象の物理機構を捉えてモデル化する方法、土地利用の変化の効果を定量的に評価しうる分布型・連続時間の流出モデルによるシミュレーション技術等を取り込んだ、より合理的な手法が求められる」と回答した。

椿ら⁴⁰⁾は、2013年度・河川技術シンポジウム（土木学会）における「危機管理の実務に供する洪水予測技術」オーガナイズドセッションでの議論に向け、洪水予測技術を（1）水位相関法、重回帰分析、ニューラルネットワーク（分布なしモデル）、（2）概念的貯留関数法、概念的タンクモデル（集中型モデル）、（3）物理的斜面流出モデル（分布型モデル）に分類した。また、椿ら⁴⁰⁾は、ピーク生起時刻の予測と実績の差、ピーク水位の予測と実績の差を評価項目とし、国内洪水予測システムの予測精度を調査し、ニューラルネットワークによるシステムを含め、実務レベルに達していないシステムもあり、洪水予測技術の精度向上、様々な流域・洪水イベン

トを統一的な手順で予測できる技術開発を進める必要性について言及した。天野ら³⁶⁾は、中小河川における洪水予測モデルの課題として、10分降水量の利用、集中豪雨を対象とした解析を提言している。

(2) 水位データの異常検知技術の技術水準と課題

国土交通省水管理・国土保全局では、全国の一級水系において水文観測を行っており、その観測データは、河川計画・管理上の重要な基礎資料として利用されている。

水位データについては、水文観測業務規程³²⁾において、観測成果の保存、照査、公表が義務付けられており、観測機器等による異常値補正後、事務所等が1観測所ごとに行う標準照査、地方整備局が複数観測所、複数観測項目について広域的な視点から行う高度照査の2段階の品質照査が実施されている。

洪水時に誤ったデータを公表することは、自治体、住民が避難判断を誤る要因となる。また、誤ったデータを含んだ過去データを公表することは、各種シミュレーション、AIの学習能力低下、学習データ浄化処理のためのコスト増の要因になると危惧される。

国内における水位データの異常検知へのAI、機械学習の適用は遅れており、上限値・下限値、標準偏差、上下流相関等の旧来の手法に技術水準がとどまっており、AIによる自動化が求められる。

(3) AIの歴史と中心人物、主要モデル、最新事例

(a) AIの歴史と中心人物 G. Hinton, イギリス人、アメリカカーネギーメロン大からカナダに移住した研究者と2006年に発表された基本論文¹³⁾

海外におけるAI, 特にニューラルネットワークの技術水準は, ここ数年急速に進歩した。これは, 2006年, Geoffrey E. Hinton (トロント大・グーグル研究所), Simon Osindero (トロント大), Yee-Whye Teh (シンガポール国立大)¹³⁾らが, 人類が将来に渡って記憶に留めると考えられる記念すべき論文を発表したことから全てが始まった。2006年の論文の題名は, 「A Fast Learning Algorithm for Deep Belief Network」である。

2006年以前, AIの中心は, ニューラルネットワークでもディープニューラルネットワークでもなく, ベイズの定理を基礎とする確率分布とその周辺モデルにあった。2006年以前には, 線形回帰, カーネル関数を用いたガウス過程回帰, ベイズロジスティック回帰, ベイジアンネットワーク, サポートベクトルマシン, グラフ表記を用いた決定木, ランダムフォレストモデル, サンプリング手法を用いたマルコフ連鎖モンテカルロ法, 系列データを扱うのに適した隠れマルコフモデルが盛んに研究されていた。

2017年現在, 世界中の大企業, コンピュータ・パターン認識・数理統計を専門とする研究者, これらを専門としない研究者, Pythonプログラマーが世界中で大騒ぎしている原点の論文がこれである。

論文の内容は, 従来困難とされていた深層ニューラルネットワークにおける学習, 具体的には深層ニューラルネットワークにおける逆誤差伝播 (バックプロパゲーション) を可能とする手法についてである。

(b) AIの主要モデル

2006年以降, G. Hintonらの研究コミュニティを中心に深層学習の各種モデル, 各種確率的勾配降下法, クリッピング, L1, L2制約等の手法の発表と関連するソフトウェア開発環境の発表が相次いだ。2017年現在, 畳込みニューラルネットワーク (CNN : Convolutional Neural Network)¹⁴⁾, LSTM : Long Short Term Memory¹⁵⁾, 自己符号化器 (Auto Encoder)¹⁴⁾, 2016年のNIPS国際会議で盛んに紹介された敵対的生成ニューラルネットワーク (GAN : Generative Adversarial Neural Network)¹¹⁾は, Geoffrey Hinton (トロント大・グーグル研究所), Yann Le Cun (ニューヨーク大・フェイスブックAI研究所), Yoshua Bengio (モントリオール大), Ian Goodfellow (グーグルブレイン・オープンAI) 等が開発に携わった技術である。

(c) AIの最新事例

海外におけるAI, 特にニューラルネットワークの適用範囲は, ここ数年急速に拡大した。画像認識, 動画認識, 動体認識, 衛星画像解析, 音声認識, 音声合成, 機械翻訳等, 従来から機械学習で取り組まれていた分野, これらに加え, 自動運転, 農業, 気象, 医療, 創薬, 教育, 音声対音声の自動翻訳, 対話型AIスピーカ, 自律兵器も現実と成りつつある。自然科学への適用範囲も広がり, 欧州原子核研究機構大型ハドロン衝突型加速器 (CERN LHC) のノイズな生データからヒッグス粒子の検出¹⁶⁾ (2013年3月発表), アインシュタイン一般相対性理論から導かれる重力レンズの探索にニューラルネットワークの技術が用いられている²¹⁾ (2017年8月発表)。AIを用いた手術用切開ロボット¹⁶⁾ (2017年10月発表) も報告されている。NASAは月面のクレータの詳細な探索にAIを適用すると発表した²¹⁾ (2017年8月)。J. Tompsonら³²⁾は, 数値流体力学 (CFD: computational fluid dynamics) 分野において畳込みニューラルネットワークを用いたシミュレーションを実施し, より自然な流体の振る舞いをシミュレートするパラメータの同定に成功したと報告した。Ruppら²⁸⁾は, シュレディンガー方程式の近似計算に機械学習を用いる手法を報告している。A. Morningstarら²²⁾は, イジングモデルにおける相転移の近似シミュレーションをAIで再現したと報告している。X. Song, H. Kanasugi, R. Shibasaki (東大)ら³⁰⁾は, LSTMを用い, 市レベルでの人の移動と交通状況の予測とシミュレーションを行う手法を報告している。今後, 人類が蓄積してきた科学技術をニューラルネットワークに学習させるための効果的で汎用的な手法が開発される可能性もある。

2-1. AIによる洪水予測の類似事例

AIによる洪水予測では, 以下の手法が可能である。この節では, 一般論 (技術水準) を簡単に解説する。

(1) 浅層順伝播型ニューラルネットワーク

浅層順伝播型ニューラルネットワーク¹¹⁾は, 入力層1層, 隠れ層1層, 出力層1層の単純な構造をしている (図-1, 図-2)。

日本国内では, 関ら³⁸⁾による阿武隈川, 竹村ら³⁹⁾による猪名川, 稲吉ら³⁷⁾による天白川, 一言ら⁴¹⁾による斐伊川を対象とした洪水予測への適用事例が報告されている。海外では, A. Brathら⁴⁾によるイタリアのアルノ川, M. Valipourら³³⁾によるイランのデズ川ダムへの流入量を対象とした適用事例が報告されている。

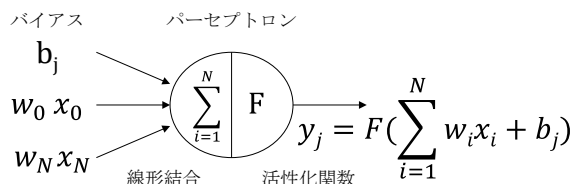


図-1 パーセプトロン

ピーク生起時刻の予測と実績の差，ピーク水位の予測と実績の差の精度が十分ではなく，概念的貯留関数法，概念的タンクモデル，物理的斜面流出モデルと組み合わせる方法，カルマンフィルタ等の同化技術を適用し予測精度の向上を図る等している。

(2) 深層順伝播型ニューラルネットワーク

深層順伝播型ニューラルネットワーク¹¹⁾は，入力層1層，隠れ層多数，出力層1層等の層が深い構造を特徴としている（図-3）．1つのニューロンは，複数入力を重み付きで加算することで，複数入力に対する線形回帰等の線形性を学習できる．活性化関数により非線型性も学習できる．隠れ層が100層を超える事例も報告されているが，多層とすることでより複雑な現象を学習し，学習に成功すれば，分類・予測性能が向上すると期待される。

多層とすると，逆誤差伝播法（バックプロパゲーション法）による重み，バイアス値の最適化が困難となるが，2006年，G. Hintonらの研究グループがこの問題を解決する手法を発表したことから，「1.(3)AIの歴史と中心人物，最新事例」で紹介したとおり，2017年現在，世界中がAIで大騒ぎとなり，自動運転，自動翻訳等が実用可能となりつつある。

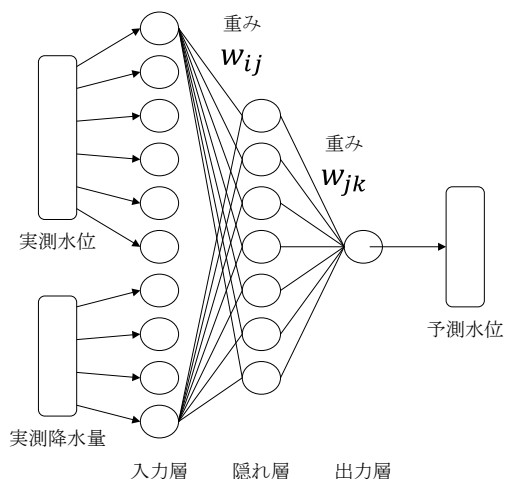


図-2 浅層順伝播型ニューラルネットワーク

(3) 動的（時間遅延）ニューラルネットワーク

動的（時間遅延）ニューラルネットワーク^{35), 20)}は，順伝播型コネクション（Feed Forward Connection）から構成されるものと，逆伝播型，再帰型コネクション（Back Forward, Recurrent Connection）を含むものに分類できる（図-4）．例えば，図-4で時刻Tに値10，時刻T+1に値0のデータを入力すると，時刻Tに値8，時刻T+1に値2を出力できる。

水位等の時系列データは，各データが独立に生起せず，ある程度従属関係（自己相関）を持つ特徴があり，統計処理も特殊性を反映したものとなる。

このため，動的（時間遅延）ニューラルネットワークは，洪水予測を含む時系列データの予測，異常検知，音声認識で有用とされている．動的（時間遅

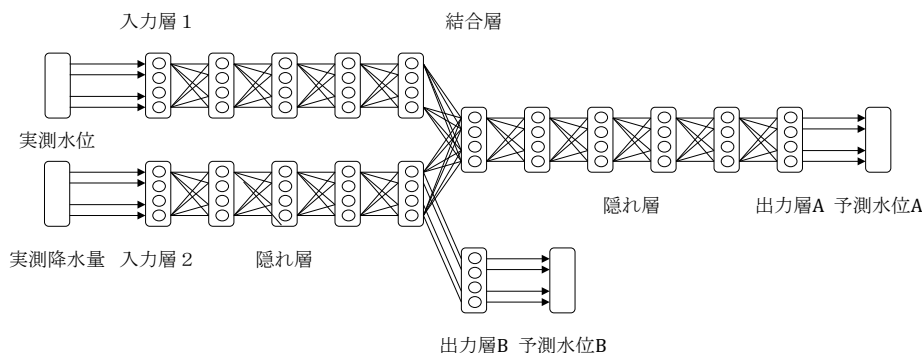


図-3 深層順伝播型ニューラルネットワークの例

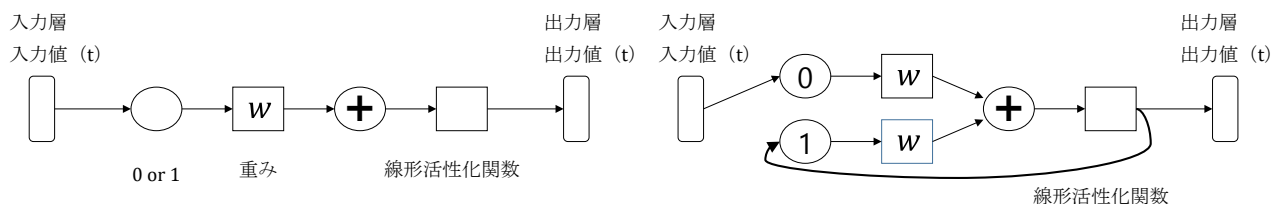


図-4 動的（時間遅延）ニューラルネットワークの例

延) ニューラルネットワークについては, Y. M. Chiang (台湾国立大) ら⁹⁾による, レーダデータから地上降水量見積, 地上降水量予測への適用事例が報告されている. A. R. Fazlina, A. M. Samad, R. Adnanら¹⁰⁾, R. Parthajit, S. Manabendraら²⁶⁾による, 洪水予測への適用事例が報告されている.

(4) WaveNet

A. Oord (グーグルディープマインド) ら^{23), 24)}は, Pixel Recurrent Neural Networkと呼ぶピクセル・チャンネル単位の自己回帰型モデルで並列化によって高速に勾配計算できるものを開発し報告している. また, A. Oord, S. Dielemanら²⁵⁾は, Pixel Recurrent Neural Networkという生成モデルをベースとし, 音声波形を生成できるWaveNetと呼ばれる深層ニューラルネットワークを開発したと報告した.

ニューロンの出力を再度同じニューロンに入力する等のモデルの再帰型では, 時系列性を取り込めるものの, 学習が困難になる. WaveNetの構造から明らかのように, WaveNetは, 右端の部分を除き順伝播型であり, 学習が困難になるという課題を回避しつつ時系列性を取り込める考えられる(図-5).

(5) Long Short Term Memory : LSTM

LSTMは, 再帰型ニューラルネットワークのひとつで, 1997年に, S. Hochreiter, J. Schmidhuberら¹⁵⁾により提案された(図-6). LSTMは, 時系列データ及び会話, 文章等の系列データを取り扱うのに適している. また, セル内部に過去, 前のデータの影響を引き継いだ値を保持することができるというメリットがある. これを周辺の更新ゲート, リセッ

トゲートにより更新, リセット可能とすることで, 再帰ニューラルネットの課題とされた学習の困難さが解消されている.

LSTMを用いた洪水予測のニューラルネットワークの例を図-6に示す.

学習時, A, B雨量観測所, C, D水位観測所の過去の時系列データから4次元, 4タイムステップ分の入力データを作成する. 予測時の時刻をTとし, 過去T-1, T-2, T-3の4次元の入力データを作成する(図-6 左下参照). 次に, 予測対象水位観測所の過去の時系列データから, 学習用の予測水位T+1, 予測水位T+2, 予測水位T+3, 予測水位T+4を作成する(図-6 右上参照). これらのデータを図6の1層, 7タイムステップのLSTMセルから構成されるニューラルネットワークに入力する. Keras等のFIT関数を起動する. Keras等のFIT関数では, 入力層から入力されて計算された出力値と, 予測水位との差分を計算し, 誤差が小さくなるようにモデルを逆方向に伝播し, LSTMセルの重み, バイアス値の最適値を計算する. これを繰り返し, LSTMセルの重み, バイアス値の最適値を求める.

予測時, 予測対象時刻(T)に対して, T, T-1, T-2, T-3の水位, 雨量データから4次元, 4タイムステップ分の入力データを作成する. この入力データを引数として指定し, Keras等のPREDICT関数を起動する. Keras等のPREDICT関数では, 学習済のLSTMセルの重み, バイアス値の最適値に従って, 出力値=予測水位を計算する.

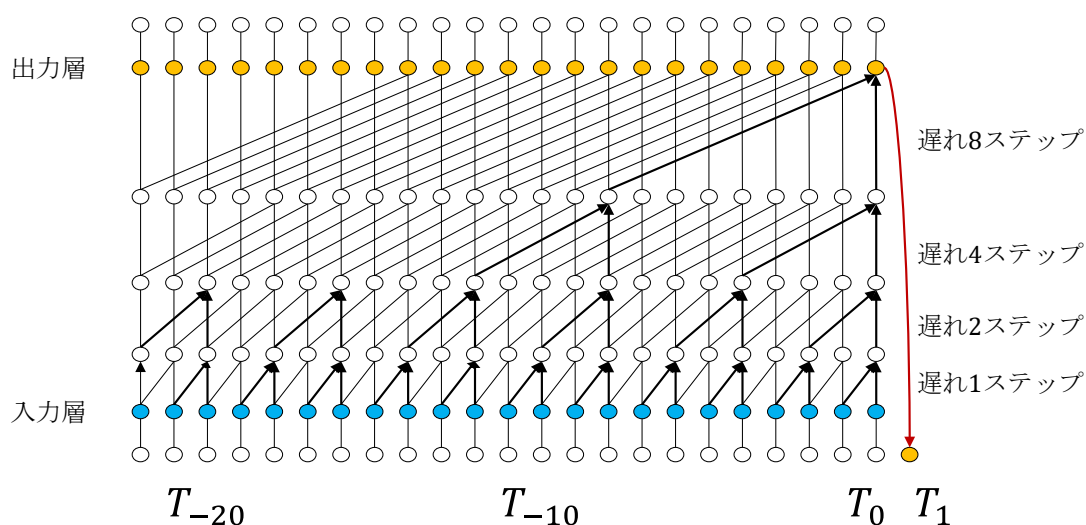


図-5 WaveNetの例

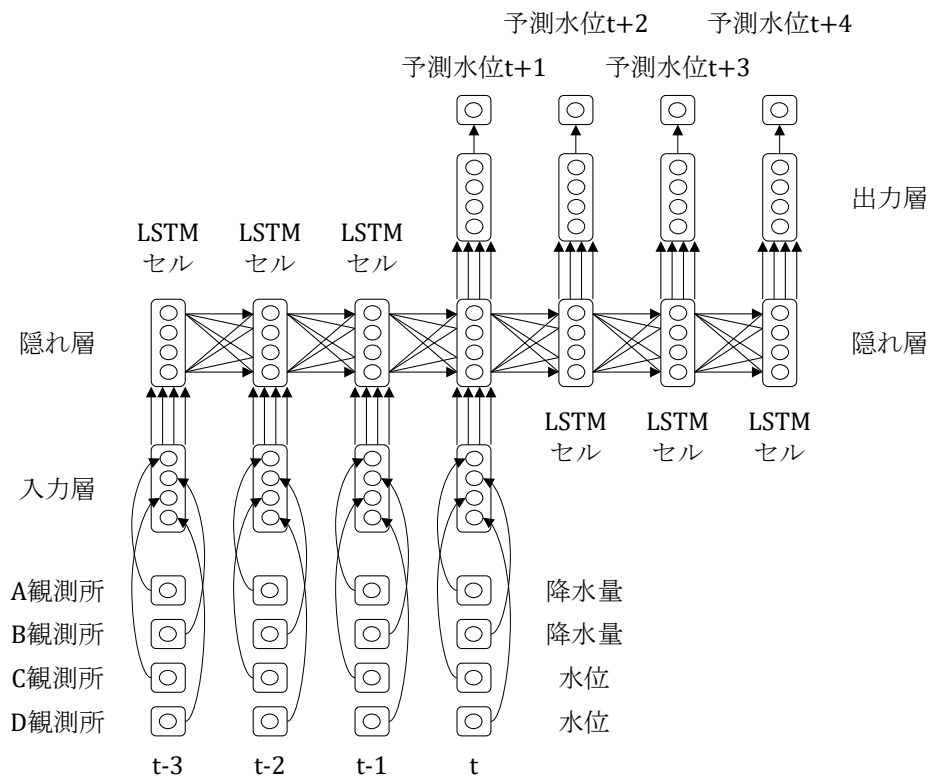


図-6 LSTM : Long Short Term Memoryの例

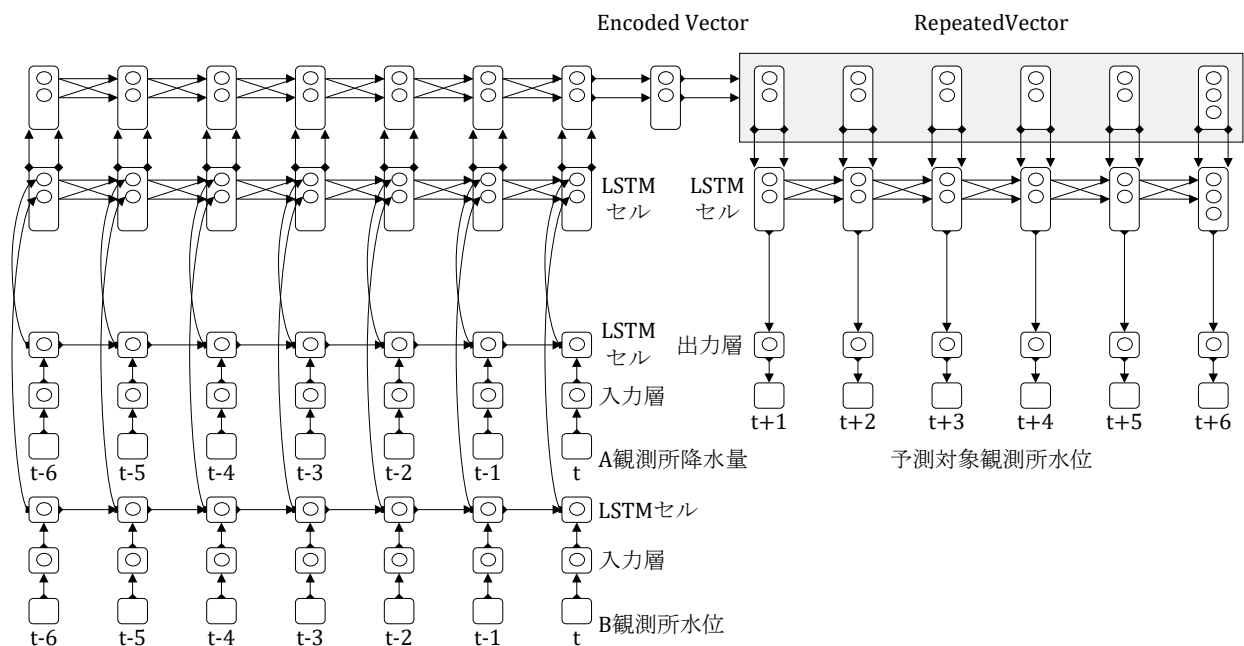


図-7 多層マルチモーダルLSTMの例

(6) 多層マルチモーダルLSTM

S. Aaron¹⁾は、2015年Kaggleコンテストの「How much did it rain?」において、米国ドップラーレーダーデータからの降水量の予測を多層マルチモーダル双方向LSTMを用いて予測し、優勝した。

S. Aaron¹⁾は、隠れ層のLSTMを多層（10層程度）

とした。また、マルチモーダルな複数種別データを別々のLSTMに入力した。個々のLSTMからの出力を上層のLSTMへの入力とした。

多層マルチモーダルLSTMを用いた洪水予測のニューラルネットワークの例を図-7に示す。

一般に、LSTMで時系列データの予測を行う場合、ワンステップスアヘッド問題という、実測値と1ス

テップずれた値を予測値と学習してしまう問題が発生する。つまり、過去の如何なる値よりも、統計的に処理すると、1ステップ前の値が予測値に近くなるためである。これを避けるため、図-7の左側の上層のLSTMの右上の出力（encoded vectorと呼ばれる）をKerasのRepeatVector関数を使用し、予測タイムステップ分、encoded vectorをコピーする。これを別のLSTMに入力し、対象観測所の水位を予測する。

(7) その他のニューラルネットワーク

M. Shafaei, O. Kisiら²⁹⁾は、ウェーブレット変換したデータを用い、水位の日データの予測を行ったと報告している。E. Racah, C. Beckham, T. Maharajら²⁷⁾は、敵対的生成ニューラルネットワーク（GAN : Generative Adversarial Neural Network）を用い、本物と見分けのつかない画像の自動生成を行ったと報告している。敵対的生成ニューラルネットワークは、Y. Le Cunが「この10年間で最も期待の持てるモデル」と評している。この敵対的生成ニューラルネットワークを用い、水位時系列データの予測をすることは、今後の課題である。また、画像の類似度を計算できるシヤムネットワーク⁶⁾（Siamese Neural Network）を用い、シミュレータから出力された予測水位と実測水位の類似度を計算し、シミュレータのパラメータの最適値を求めることにより、最終的に水位を予測するニューラルネットワークも実現可能と考えられる。

上記の他、改めて言及することも無いが、レーダデータは、地上雨量観測計では捉えられない降水強度の分布、時間変化等に関する情報が得られ、今後、洪水予測をするための必須のデータである。また、AIの優れた能力として、動画像識別能力、動体識別能力が挙げられる。畳込みニューラルネットワーク（CNN : Convolutional Neural Network）と時系列データを扱うのに優れたLSTMを組み合わせることで、まだ地上に降っていない降水量の予測も今後可能と考えられる。本報告では言及しないが、これはAIで取り組むことが望まれる大きな課題である。

2-2. AIによる水位データ異常検知の類似事例

AIによる水位データの異常検知では、以下の手法が可能である。この節では、一般論（技術水準）を簡単に解説する。

(1) 機械学習を用いたモデル

2006年から2012年頃に始まった深層学習ブーム以前、様々な分野において、機械学習と呼ばれる技

術が予測精度の最高順位にランクしていた。

機械学習³⁾を用いて異常検知するモデルには、LAG1+DBSCANを用いた手法、g-ESDを用いた手法、ARIMAモデルを用いた手法がある。それぞれ、水位データの異常検知が可能である。

LAG1は、アメリカ国立標準技術研究所（NIST）が時系列データの品質確保の手法の一つとして紹介している。LAG1では、当該時刻の水位と当該時刻+1の水位の差を横軸にプロットする。当該時刻+1の水位と当該時刻+2の水位の差を縦軸にプロットする。水位データには、一般に自己相関があるとされる。過去データの傾向の影響を当該時刻のデータが受けている。

正常データは、傾き1の直線上にプロットされ、異常は、傾き1の直線から離れた位置にプロットされる性質がある。DBSCANにより上記各点をクラスタに分割し、同じクラスタに属する点の数が閾値以下のものを異常と検知する（図-8）。

Generalized - Extreme Studentized Deviate (g-ESD) は、アメリカ国立標準技術研究所（NIST）が時系列データの品質確保の手法の一つとして紹介している。時系列データから標準偏差を計算し、観測値から平均値を引いて標準偏差で割った値の絶対値が最も大きい値を異常と検知する。異常と検知した値を取り除き、この計算を繰り返す。異常の最大値の件数を指定し、これを超えるまで異常を検知する（図-9）。

Auto Regressive Integral Moving Average (ARIMA) は、時系列データの分析手法の一つであり、時系列データを自己回帰パラメータ、差分の階差、移動平均パラメータを用いて分析（各成分に分解）することで、時系列な現象が時間とともにどのように変化するか、将来どのようになるかを確率分布付きで分析、予測できる。信頼範囲等、確率範囲の閾値の外側のデータを異常と検知できる（図-10）。

N. Laptev, S. Amizadeh, I. Flintら¹⁹⁾は、EGADSというYahooで時系列データの異常をリアルタイムで検知するためのシステムを報告している。外れ値は、ARIMA, Exponential Smoothing, Kalman Filter等で検知する。変化点は、カーネス密度分布、カルバック・ライブラーダイバージェンスで検知する。

(2) 順伝播型ニューラルネットワーク

順伝播型ニューラルネットワークでは、出力層の活性化関数にソフトマックス関数を用いると、確率つきで入力データを分類する分類器とすることが出来る。学習は、教師あり学習、ラベル付き学習、スーパーバイズド学習と呼ばれる手法で行う。

具体的には、学習用データを過去18タイムステップ分の水位10分データとする。これを1データとする。1データ毎に正常、異常（スパイクノイズ）、異常（長期無変動）等のラベルを人間が付ける。こ

のとき、ラベルは、ワンホットベクトルとよばれる [1,0,0], [0,1,0], [0,0,1] 等のベクトル形式とする。3種類に分類する場合、3次元のベクトル、10種類に分類する場合、10次元のベクトルとする。

このラベル付き学習用データを過去水位データから10万件等、作成する。この場合、入力層は18次元（18ニューロン）、出力層は3次元（3ニューロン）となる。隠れ層を10次元（10ニューロン）の層を2層等とする。

KerasのFIT関数を起動すると、ニューラルネットワークのプログラムが、学習用ラベル付きデータ10万件から、出力層での計算値とラベルの値の誤差が最も小さくなる重み、バイアス値を自動で計算する。

異常検知時、このニューラルネットワークの入力層に過去18ステップ分のデータを入力すると、このデータが正常、異常（スパイクノイズ）、異常（長期無変化）である確率が0.7, 0.2, 0.1等と出力される。

河川情報センターでこの手法でプログラムを開発したところ、無学習、つまり、出力値の確率が0.33, 0.33, 0.33と均等の確率を出力した。このため、時系列データの階差（1データ前との差）を計算し、18タイムステップ分のデータについて、この階差を20区間に分割したヒストグラム表記とし、各区間毎の出現確率を入力層に入力したところ、良好に正常、異常（スパイクノイズ）、異常（長期無変化）の確率を出力するように改善された（図-11）。

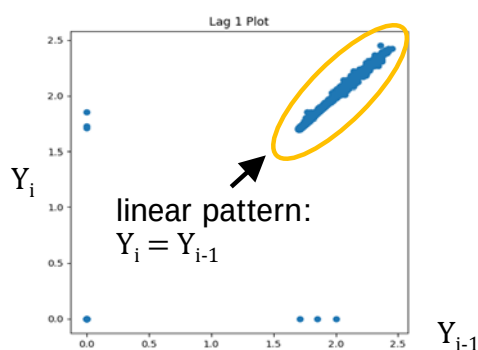


図-8 LAG1 + DBSCANを用いた異常検知

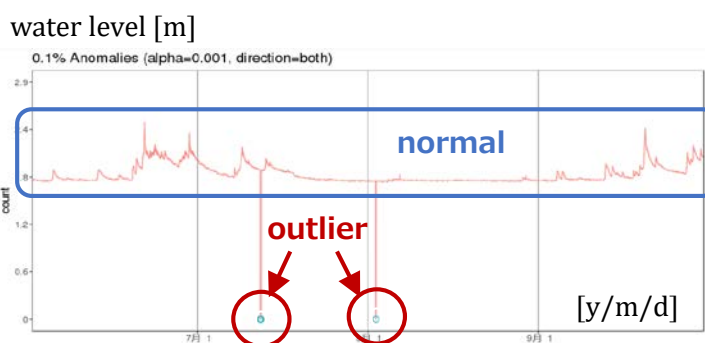


図-9 g-ESDを用いた異常検知

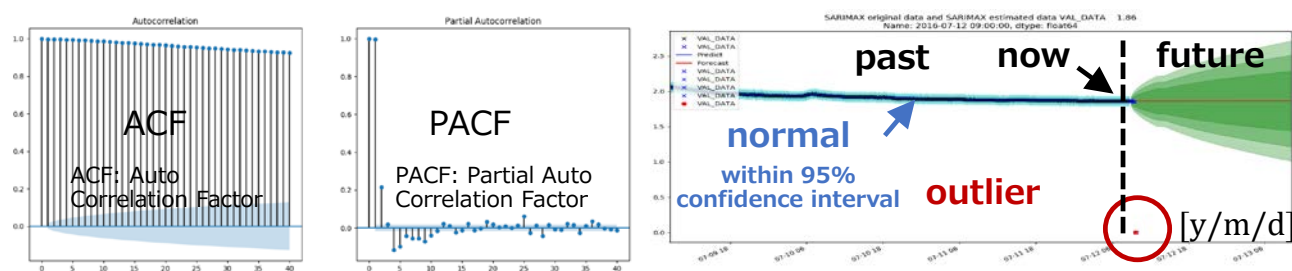


図-10 ARIMAを用いた異常検知

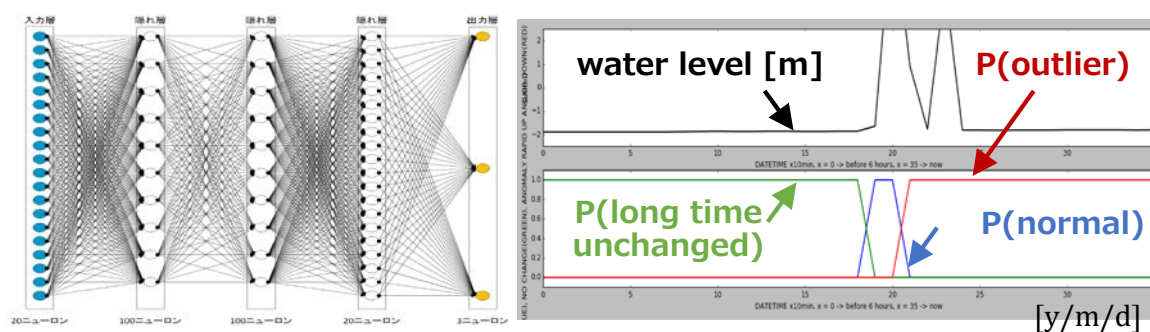


図-11 順伝播型ニューラルネットを用いた異常検知

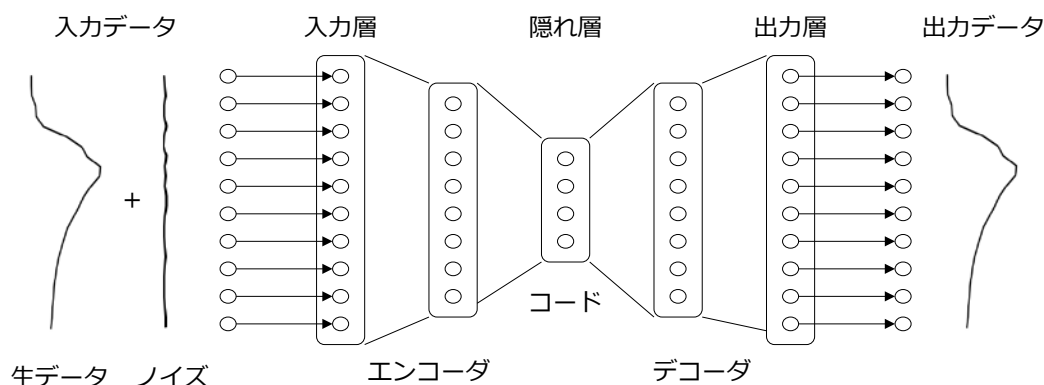


図-12 ノイズ除去自己符号化器の例

(3) ノイズ除去自己符号化器

G. Hintonら¹⁴⁾は、2006年、高い次元数を持つデータを低い次元数のエンコードされたデータに変換することが可能な多層ニューラルネットワーク（自己符号化器）を開発した。自己符号化器は、例えば入力層が10、中間の隠れ層が8, 4, 8、出力層が10ニューロン等、中間の隠れ層のニューロンの数が削減されるという特徴的な構造をしている。

次元削減は、クラスタリング、分類等で利用される。簡単で広く普及している手法は、主成分分析（PCA：Principal Components Analysis）で、最も大きな偏差がある方向を最初の軸とし、元のデータを座標変換し、これを最後の軸が決まるまで繰り返す。

自己符号化器は、このPCAをニューラルネットワークで実現したもの言われている。

ビンセントら³⁴⁾は、2010年、学習時、正規分布等に従う誤差を加えたデータを入力層に入力し、誤差のないデータを出力層に出力するように学習させるノイズ除去自己符号化器を開発した。図-12にノイズ除去自己符号化器を水位データの異常検知に適用する例を示す。予測時、入力データに未学習な異常が含まれる場合、入力データと出力データの誤差が大きくなる。この性質を利用し、水位データから異常を検知できる。

(4) 構造化ノイズ除去自己符号化器⁹⁾

T. Tagawa, Y. Tadokoro（トヨタ中央研究所）、T. Yairi（東大）ら³¹⁾は、2014年、ノイズ除去自己符号化器を拡張し、自動車のエンジン等から発生する40種類を超えるセンシングデータを利用し、異常の原因を分析できると報告した。

(5) LSTMノイズ除去自己符号化器

R. Kienzler（IBM）¹⁷⁾は、2014年、ノイズ除去自己符号化器を拡張し、パーセプトロンで構成していたレイヤを、LSTMに置き換えたニューラルネットワークを報告した。

3-1. FRICS洪水予測AIモデル（プロトタイプ）の概要

この節では、我々が提案するアイデアと手法を詳しく解説する。また、モデル流域は、雨量観測所3箇所、水位観測所2箇所、ダム1箇所とする。10分降水量、10分水位からダムの流入量を予測するという課題を設定した。

(1) 本研究ではWaveNetの構造をベースとする

本研究の意義、位置付け、目標を以下に示す。

- (a) WaveNetが洪水予測に適用可能であるか否かを明らかにすること、
- (b) WaveNetの適用により従来課題とされていたピーク生起時刻の予測と実績との差、ピーク水位の予測と実績の差が改善されるか否かを明らかにすること、
- (c) 従来の概念モデル、物理モデルによる繰り返し計算を必要とせず、簡易で汎用な手法をWaveNetに組み合わせることで、予測性能が改善するか否かを明らかにすることである。

(2) 降雨発生時刻及び水位予測時刻が前後しない（因果関係を破壊しない）という制約条件をニューラルネットワークに取り込むアイデア

- (a) WaveNetの構造をベースとし、以下の拡張を行う。
- (b) ニューラルネットワークの入力層に、実測の10分降水量48タイムステップ（8時間分）を

入力する。出力層に、ニューラルネットワークで計算した10分流入量48タイムステップ（8時間分）を出力する。これに加え、出力層に、予測用の10分流入量16タイムステップ（2時間40分）を出力する。

- (c) 1雨量観測所あたり4次元データ（48タイムステップ分の累加10分降水量、上記から仮定した減衰分を引いた量等）とする。
- (d) 4次元データは、次のニューロンに入力されるが、この間は、 4×4 の全結合とする。
- (e) WaveNetの次にタイムステップ毎の隠れ層を2層構成し、雨量観測所3箇所のデータを 3×3 の全結合により最適な線形結合、非線形結合となるように学習させる。

(3) 洪水到達時間、流速の変動を取り除くアイデア

- (a) 流速を洪水時にリアルタイムで取得するのは困難と言われている。洪水到達時間、遅延時間、ピーク水位の生起時刻を予測するためには、流速の情報が重要となる。流速の代理の値として水位観測所の水位を用いる。簡単のため、水位観測所毎の48タイムステップ分の中央値を計算し、さらに2水位観測所の中央値を平均する。
- (b) 上記流速の代理の値の最大値、最小値を全学習用データから計算し、流速の代理の値を正規化する。流速の代理の値を複数区間に分割する。ここでは、20区間（区間1：0.00～0.05、区間20：0.95～1.00等）とした。
- (c) ニューラルネットワークを20個構築する。構造は同じだが、重み、バイアス値は異なる。
- (d) 10分降水量、10分流入量の学習用のデータを上記区間毎に分割し、区間毎のニューラルネットワークの重み、バイアス値を出力層での流入量の計算値と実測値の誤差が最小となるように最適化する。
- (e) 予測時、48タイムステップ分の10分水位データから観測所毎の中央値、中央値の平均値を計算し、予測に使用するニューラルネットワークを20区間分から選択する。
- (f) 粒子フィルタ、アンサンブルカルマンフィルタ、モンテカルロ法の考え方を取り入れる。降水量の観測時刻に、ガウス分布に従って、ランダムな誤差を加え、流入量の予測値を計算し、平均値を最終的な流入量の予測値とする。

(4) その他のアイデア

10分降水量、累加降水量、水位変化のハイドログラフを観察する。累加降水量と水位のグラフは、洪水到達時間（遅延時間）にずれがあるものの、10分

降水量のグラフと比べ、良く合っているように見える。しかし、一時的に雨が止んだ等の場合、累加降水量は変化しないに関わらず、水位のグラフは、水位が一時的に減少する振る舞いを示すケースが見られる。

また、降水量がゼロであっても、流入量はゼロとまらないケースが見られる。これら、2点は流入量の予測の誤差の原因と仮定し、これらの影響を取り除くため、以下のアイデアを取り入れる。

- (a) ニューラルネットに入力するデータとして、以下4種類のデータを用いる。
 - ・ 10分降水量48タイムステップ分の累加降水量
 - ・ 上記から指数関数で減衰する量を引いた量
 - ・ 上記から48タイムステップ分のダム流入量の例えば5%分を差し引いた量、
 - ・ 上記から48タイムステップ分のダム流入量の例えば10%分を差し引いた量、
- (b) ニューラルネットの出力層の手前で、例えば、0.01mmの一定の降水量を入力する。

3-2. FRICS異常検知AIモデル（プロトタイプ）の概要

この節では、我々が提案するアイデアと手法を詳しく解説する。

(1) 本研究ではノイズ除去自己符号化器をベースとする

本研究の意義、位置付け、目標を以下に示す。

- (a) ノイズ除去自己符号化器が水位データの異常検知に適用可能であるか否かを明らかにすること、
- (b) 水位データをスパイクノイズ検知用に事前に特徴量変換することで、検知性能が改善するか否かを明らかにすること、
- (c) 水位データを長期無変動検知用に事前に特徴量変換することで、検知性能が改善するか否かを明らかにすること、
- (d) 水位データをウェーブレット変換により特徴量変換することで、スパイク、長期無変動を汎用的（1つのニューラルネットで）に異常と検知できるか否かを明らかにすることである。

(2) スパイクノイズ用に特徴量変換するアイデア

水位データの値は、平常時には一定で、出水時は、

10分データにおいてもデータの値が不連続になる程度に変動する．これをそのまま，特徴量変換せずにノイズ除去自己符号化器で学習させると，出水時の洪水波形は，異常と検知される可能性がある．本研究では，以下の式で水位データを特徴量変換した．

$$X(t) = \text{Log}\left(10 * \frac{x(t) - \min(x(t0,,,tN)) + 0.50}{\max(x(t0,,,tN)) - \min(x(t0,,,tN))}\right)$$

(3) 長期無変動用に特徴量変換するアイデア

水位データの値は，平常時には一定で，10分データにおいてもデータの値がほとんど変動しない場合がある．平常時の水位と長期無変動時の水位の変化量は洪水時の変化量に比べ，いずれも極わずかである．これをそのまま，特徴量変換せずにノイズ除去自己符号化器で学習させると，平常時の水位を異常と検知するか，長期無変動の異常を見逃す可能性がある．

本研究では，以下の式で水位データを特徴量変換した．最初に1つ前のタイムステップの水位データ間との階差を計算する．48タイムステップ分の階差の値について，標準偏差を計算し，逆数を変換後の値とする．これにより，無変動の一定値のみならず，機器故障時に見られる，斜め直線上のドリフトも検知可能としている．

$$X(t) = \frac{1}{\text{stdev}(x(t) - x(t-1)),,,(x(t-47) - x(t-48))}$$

(4) ウェーブレットに特徴量変換するアイデア

10分水位データをscipy.cwt関数を使用し，連続ウェーブレット変換した．パラメータは，以下を使用した．

マザーウェーブレット：リッカー
バンド幅：48

(5) 検知性能向上を図るため水位データを水位方向に区間分割し区間毎にノイズ除去自己符号化器を学習させるアイデア

- (a) 水位データの値の最大値，最小値を全学習用データから計算し，水位データの値を正規化する．水位データの値を複数区間に分割する．ここでは，5区間（区間1：0.00～0.20、区間5：0.80～1.00等）とした．
- (b) ニューラルネットを5個構築する．構造は同じだが，重み，バイアス値は異なる．
- (c) 10分水位の学習用データを上記区間毎に分割し，区間毎のニューラルネットワークの重み，バイアス値を最適化する．

表-1 開発環境

部品	型番
マザーボード	GIGABYTE X99P-SLI (x1)
CPU	Xeon E5-2609 v4 @1.70GHz 8core
メモリ	DDR4-2400 8GB (x4) 合計 32GB
SSD	240GB RAID 0 (ストライピング) (x2) 合計 480GB
GPU	nVidia GTX GeForce 980 Ti (x2)
OS	CentOS 7.5.1611
CUDA	CUDA 7.5.17
Python	Python 3.5.2
Tensor Flow GPU	Tensor Flow for GPU 1.2.0
Keras	Keras 2.0.4
Eclipse, PyDev	4.6.3, 5.6.0
Library	numpy, scikit-learn etc.

- (d) 異常検知時，12，24，36，48ステップ分の10分水位データから中央値を計算し，異常検知に用いるニューラルネットワークを5区間から選択する．

4. プログラムの開発と検証

洪水予測プログラムと水位データの異常検知プログラムを表1に示す環境をターゲットとし開発した．

(1) CPU使用率

ニューラルネットワーク学習時，CPU使用率は8コア全てにおいて50%前後のリソースを消費していた．また，学習前，学習後のファイル入出力処理時，CPUは1コアのみが100%となっていて，その他のコアは0%に近かった．

上記から，CPU周波数は，1.7GHzではなく，3.0GHz以上を推奨する．

(2) GPU使用率

GPU使用率は，「nvidia-smi」コマンドを実行し確認できる．洪水予測，水位データの異常検知の学習時，GPU使用率は，50%未満であった．このため，単独のモデルを学習させる場合，複数モデルの学習でも順番に学習させる場合，nVidia Geforce GTX 980 Ti 程度の一昔前のGPUでも処理可能であった．複数モデルを並列して学習させる場合，複数人で1台のサーバを共有する場合，nVidia VP-100、nVidia-DGX-1等の高性能で高価な環境を整備する必要があると考察する．なお，今回の実験では，

CPU, メモリ, GPU, GPUメモリ, ストレージ容量を多く消費する画像データを扱っていないことに留意いただきたい。また, nVidia Geforce GTX 980 Ti は1年ほど前に既に製造中止されており, 2017年現在, 後継のnVidia GTX 1080 Tiの購入 (10万円前後/1GPU) を推奨する。

(3) メモリ使用率

ソースコードの可読性, 保守メンテナンス性を高めるため, かつ, AIの学習速度を向上させるためには, 全学習用データを1度にFIT関数 (Keras の学習開始用の関数) に投入 (CPUのメモリ, SSDからGPUのメモリにデータを転送) することが望ましい。この方法は, CPUメモリ使用率を増大させる。予測精度を高めるため, 過去データを100万件程度とす

ると, 学習時にメモリ使用率が100%に達してしまう障害が発生した。このことから, 学習用マシンのメモリ量は, 洪水予測, 水位データの異常検知目的においても32GBは過小で, 128GB程度のメモリ搭載を推奨する。

4-1. 洪水予測AIモデル

この実験結果は, 当日発表予定である。

4-2. 異常検知AIモデル

この節では, AIによる水位データ異常検知の実験結果を図-13から図-19に示す。

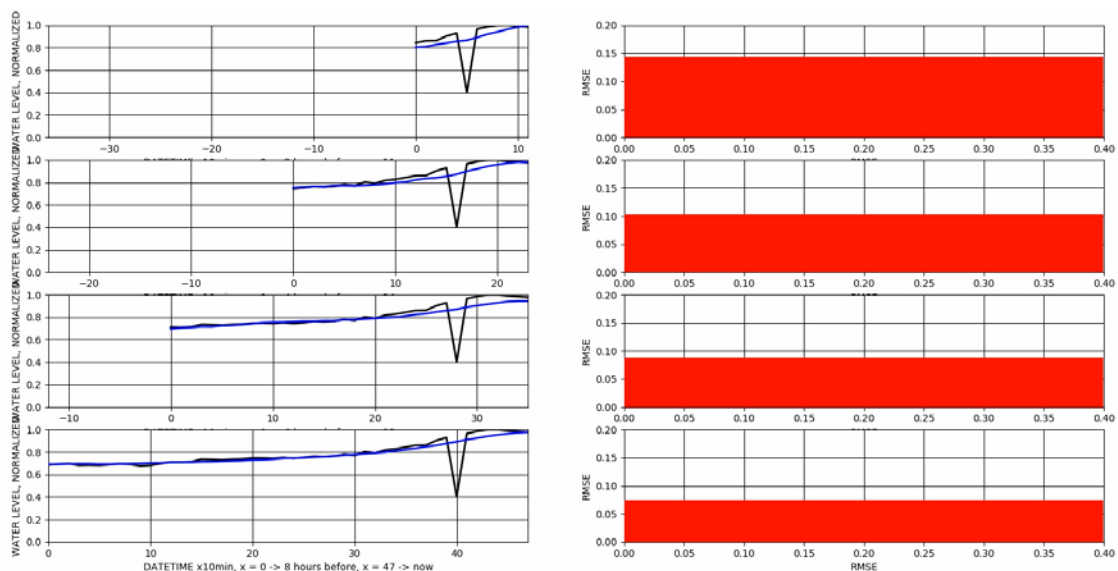


図-13 スパイクノイズ検知用に特徴量変換し学習済のDAEに, ピーク付近にスパイクノイズを合成した水位を入力した場合

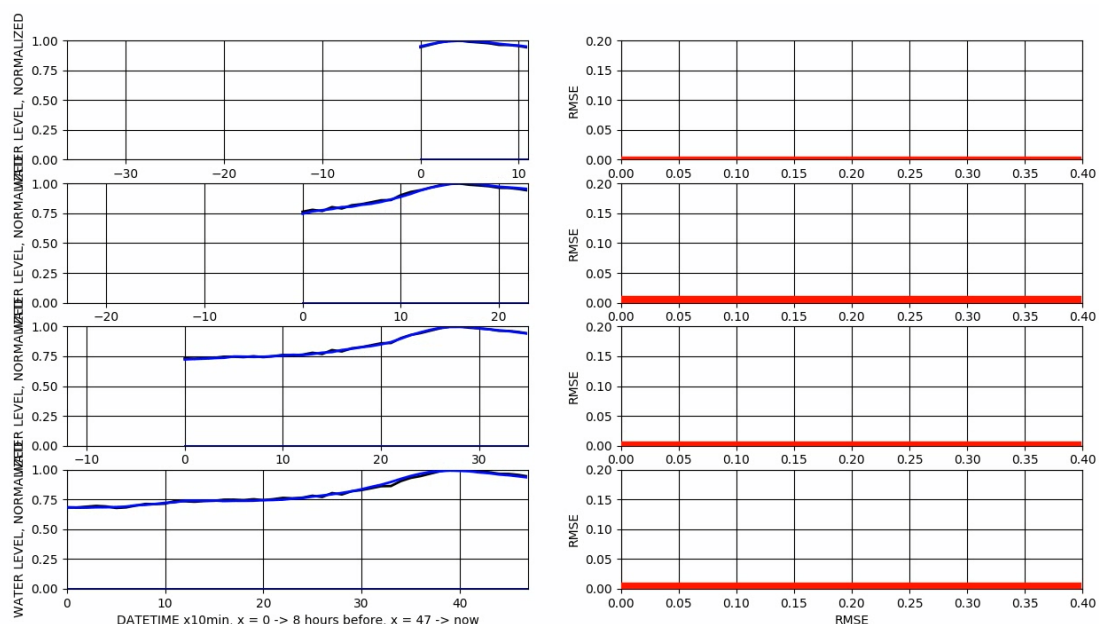


図-14 スパイクノイズを合成しない水位を入力した場合

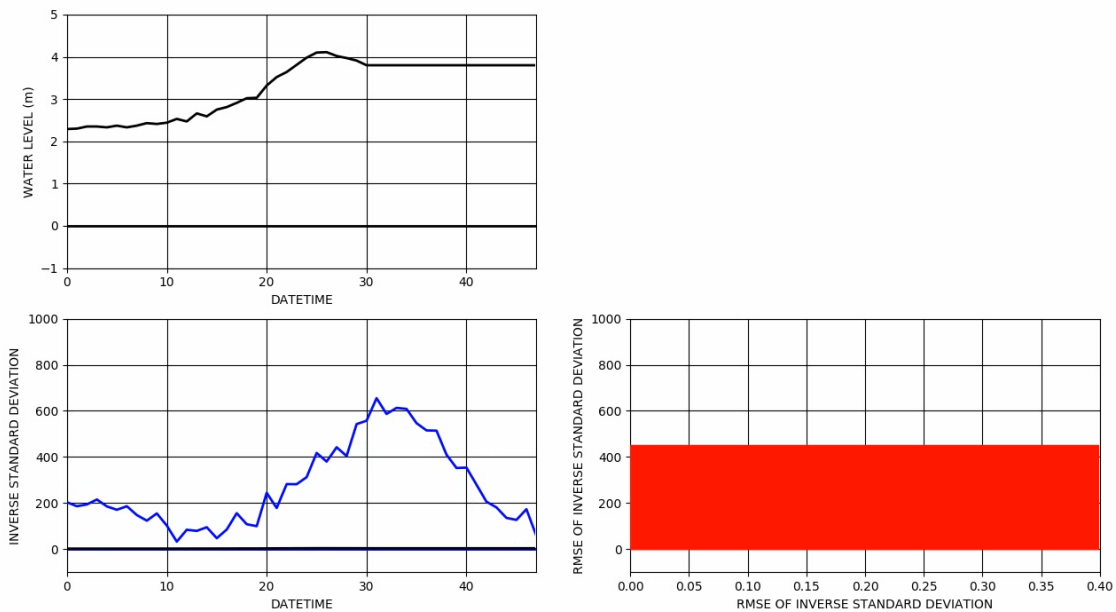


図-15 長期無変動検知用に特徴量を階差、標準偏差の逆数で変換し学習済のDAEに、水位ピーク付近に長期無変動を合成した水位データを入力した場合

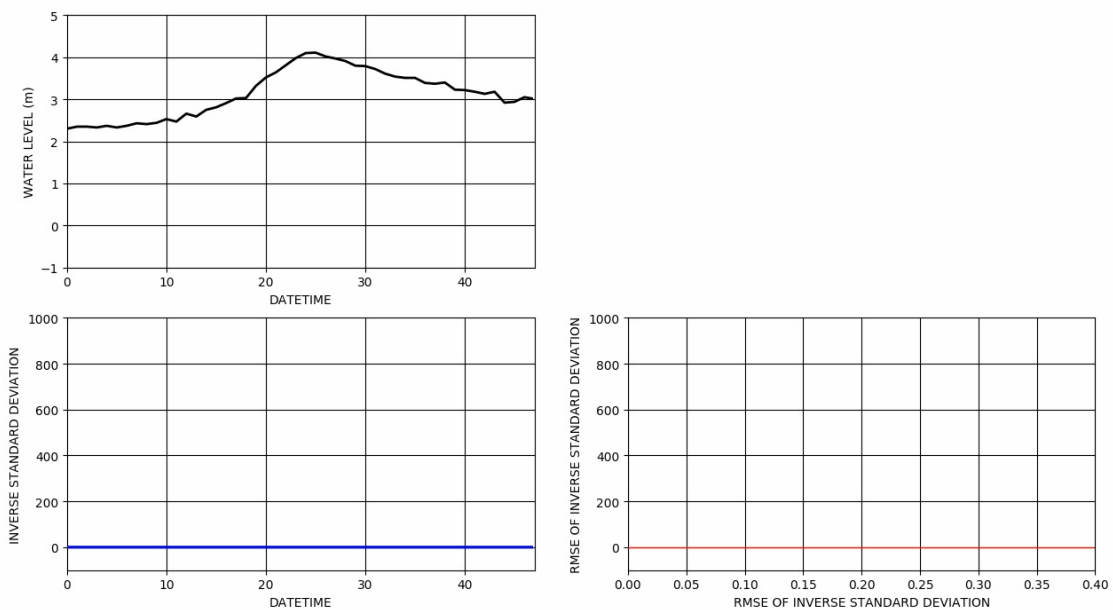


図-16 長期無変動を合成しない水位データを入力した場合

図（左上，左下）の横軸は、8時間（10分水位データ48タイムステップ）。左端の「0」が8時間前、右端の「47」が現在。図左上の縦軸は、水位、図左下の縦軸は、1タイムステップ前との階差の48タイムステップ分で処理した標準偏差の逆数、図右下の縦軸は、表示されている48ステップ分でRMSE（Root Mean Square Error）、つまり、誤差の平均値を計算したものである。

（1）スパイクノイズ

図-13、図-14から分かる通り、スパイク検知用に特徴量を変換したことで、ピーク後にスパイクの異常が発生した場合でも、RMSEの閾値を設定するこ

とで検知可能であることがわかった。

（2）長期無変動

図-15、図-16から分かる通り、長期無変動用に特徴量を変換したことで、ピーク後に長期無変動の異常が発生した場合でも、RMSEの閾値を設定することで検知可能であることが分かった。

上記DAEの学習では、長期無変動の期間を8時間n（10分水位データ48ステップ）以上とし、特徴量の変換時に使用したが、この長さを微調整すれば、2時間、4時間、6時間、8時間、それ以上、それ以下の場合にも検知できると期待できる結果であった。

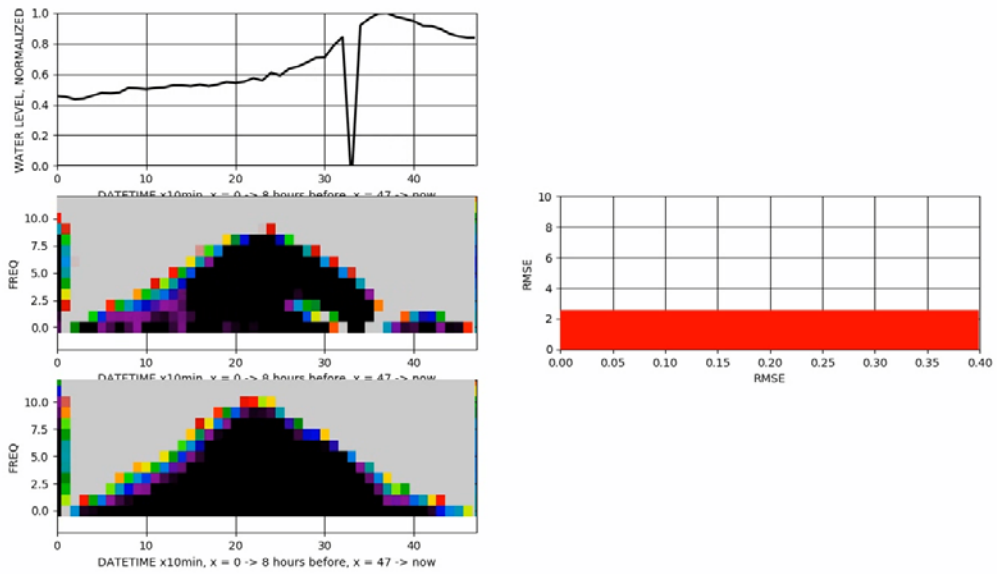


図-17 図ウェーブレット変換した水位データで学習済のDAE（畳込みNN自己符号化器）に、水位ピーク付近にスパイクノイズを合成した水位データを入力した場合

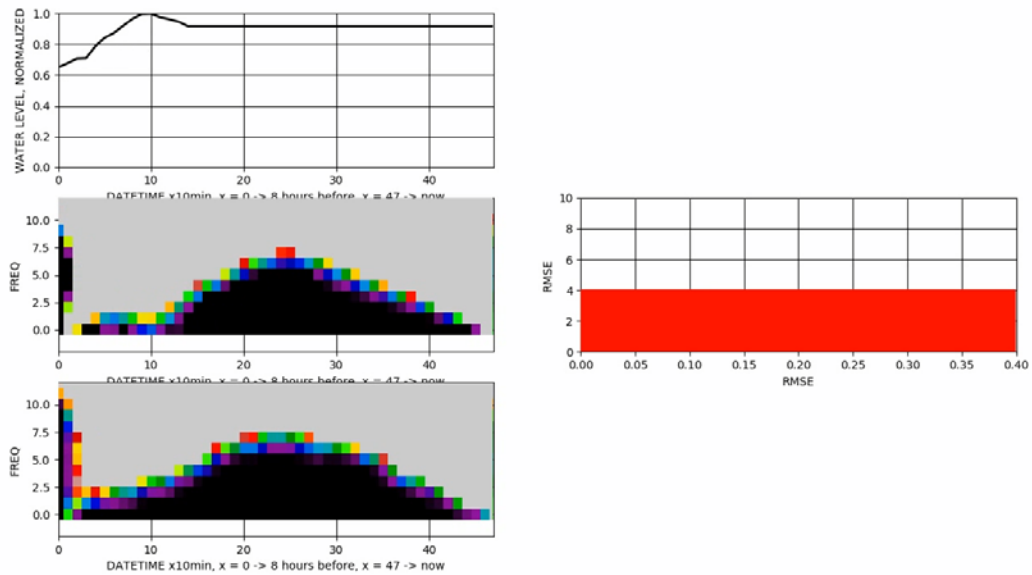


図-18 水位ピーク付近に長期無変動を合成した水位データを入力した場合

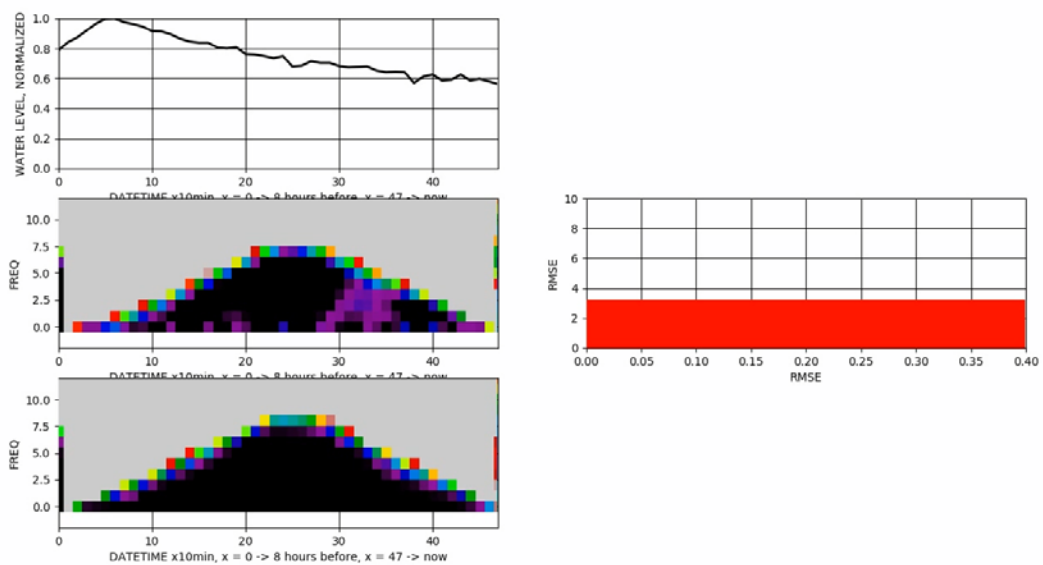


図-19 水位ピーク付近に何も合成しない水位データを入力した場合

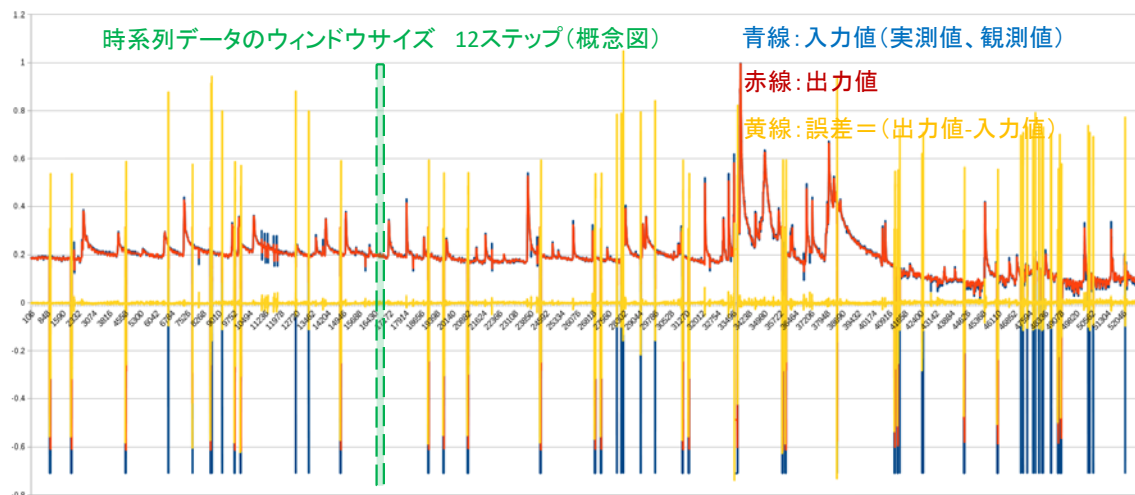


図-20 ノイズ除去自己符号化器で欠測データを抽出した結果

(3) ウェーブレットによるスパイク、長期無変動

図（左上，左中，左下）の横軸は，8時間（10分水位データ48タイムステップ）．左端の「0」が8時間前、右端の「47」が現在。図左上の縦軸は，水位，図左下の縦軸は，ウェーブレットの周波数である．図右下の縦軸は，表示されている48ステップ分でRMSE（Root Mean Square Error），つまり，誤差の平均値を計算したものである。

ウェーブレット変換とノイズ除去自己符号化器の（畳み込みNNを使用の）構成では，入力データにスパイクノイズが混入した場合でも，出力されるウェーブレットは，スパイクノイズをフィルタリングし，未検出している．このことは，この手法で異常検知できる可能性を示している．ウェーブレット変換を用いる場合でも，事前に水位データの特徴量抽出処理を行うことが望ましいのかもしれず，今後の課題とする。

(4) スパイクノイズ検知用に事前に学習したDAEで欠測データを抽出した結果

図-20に1年分の水位データ（欠測が多数）について，異常検知した結果を示す．本研究では，タイムスライドウィンドウと呼ばれる，ウィンドウを移動させながら異常検知する方法で行っている．実験では，12，24，36，48タイムステップでDAEを作成したが，スパイクノイズの場合，12ステップ程度が望ましいとの結果であった．システム化に向け，複数タイムステップで検知を多重化する等も必要と思われる。

6. おわりに

本研究は，「2001年宇宙の旅」「ターミネータ」等で描かれたSFの世界，世界中の科学技術，特にコンピュータ・サイエンスに取り組む研究者が夢のひとつとしていたAIの世界が2006年のG. Hintonの論文以降，実現に向けて急速に動き始めたタイミングで実施したものである．「AIが有用なことは日々のニュース等で知っているが，河川分野でも使えるのか？」という素朴な疑問に答えることを第一の目標とし，洪水予測と水位データの異常検知という課題にAIを適用することを試みた。

最後に本研究のメッセージを以下に示す。

- 「AIの適用範囲は洪水予測に留まらず，レーダ画像，衛星画像からはより有用な情報を取り出せる」
- 「nVidia GTX 1080 Ti を購入する（お金が無い）」
- 「nVidia VP-100, nVidia DGX-1を購入する（お金が有り余っている）」
- 「Pythonを勉強する、特にnumpyデータ形式」
- 「Hinton教授の講義，スタンフォード大の自然言語処理の講義等をYou Tubeで勉強する」
- 「Google, Facebook, トロント大，スタンフォード大，カーネギーメロン大，MIT等の最新の論文をarXiv.orgで無料でダウンロードし勉強し，無料で投稿する」
- 「人類が蓄積してきた科学技術をニューラルネットワークに取り込む手法を考える」
- 「貧乏国に落ちていく現状を止めるため，科学技術，AIとPythonプログラマを正しく位置付ける」
- 「ニューラルネットワークに入力するデータは，Feature Engineeringを行い，正規化，正則化，対数変換，ヒストグラム等の処理を行い，正規分布に近づける」

参考文献

- 1) S. Aaron: How much is it rain II, Kaggle, 2015.
- 2) C. C. Aggarwal: Outlier Analysis 2nd Edition: Springer, 2013.
- 3) C. M. Bishop: Pattern Recognition and Machine Learning: Information Science and Statistics, Springer, 2006.
- 4) A. Brath, A. Montanari, E. Toth, Neural networks and non-parametric methods for improving realtime flood forecasting through conceptual hydrological models, Hydrology and Earth System Sciences, 6(4) pp.627-640, 2002.
- 5) Y. M. Chiang, F. J. Chang, B. J. D. Jou, P. F. Lin: Dynamic ANN for precipitation estimation and forecasting from radar observations: Journal of Hydrology, 334, pp.250-261, 2007.
- 6) S. Chopra, R. Hadsell, Y. LeCun: Learning a Similarity Metric Discriminatively, with Application to Face Verification, Proc. of Computer Vision and Pattern Recognition Conference, IEEE Press, 2005
- 7) Y. LeCun, Y. Bengio, G. Hinton: Deep Learning, Nature vol.521, no.7553, pp. 436-444, 2015.
- 8) Y. LeCun, L. Bottou, Y. Bengio: Reading checks with multilayer graph transformer networks. In Acoustics, Speech, and Signal Processing, ICASSP-97., 1997 IEEE International Conference on, volume 1, pp. 151-154, 1997
- 9) S. X. Ding: Model-based fault diagnosis techniques, Springer, 2013.
- 10) A. R. Fazlina, A. M. Samad, R. Adnan: Model-based fault diagnosis techniques, Springer, 2013.
- 11) I. J. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIP Press, 2016.
- 12) A. Graves: Supervised Sequence Labelling with Recurrent Neural Networks, 2012.
- 13) G. Hinton, S. Osindero, Y. Teh: A Fast Learning Algorithm for Deep Belief Nets: Neural Computation 18, 1527-1554, MIT, 2006.
- 14) G. Hinton, R. R. Salakhutdinov: Reducing the dimensionality of data with neural networks: Science, Vol.313, no. 57866, pp.504-507, 2006.
- 15) S. Hochreiter, J. Schmidhuber: Long Short-Term Memory, Neural Computation, 9(8), pp.1735-1780, 1997.
- 16) IEEE Spectrum. URL <https://spectrum.ieee.org>
- 17) R. Kienzler, LSTM AutoEncoder
- 18) A. Krizhevsky, I. Sutskever, G. Hinton: Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, 2012.
- 19) N. Laptev, S. Amizadeh, I. Flint: Generic and Scalable Framework for Automated Time-series Anomaly Detection, KDD2015, 2015.
- 20) Mathworks: Documentation, How Dynamic Neural Networks Work, 2017. URL [https://jp.mathworks.com/help/nnet/ug/how To](https://jp.mathworks.com/help/nnet/ug/how>To)
- 21) MIT Technology Review. URL <https://www.technologyreview.com>
- 22) A. Morningstar, R. G. Melko: Deep Learning the Ising Model Near Criticality, arXiv: 1708.04622v1, 2017.
- 23) A. Oord, N. Kalchbrenner, K. Kavukcuoglu: Pixel recurrent neural networks. arXiv preprint arXiv: 1601.06759, 2016a.
- 24) A. Oord, N. Kalchbrenner, A. Graves: Conditional image generation with PixelCNN decoders. CoRR, abs/1606.05328.2016b.
- 25) A. Oord: WaveNet: A generative Model for Raw Audio, 2016. URL <http://arxiv.org/pdf/1609.03499.1097-1105>, 2012.
- 26) R. Parthajit, S. Manabendra: ICMOS 2010, AIP Conference Proceedings, Vol 1298, pp.219-225, 2010.
- 27) E. Racah, C. Beckham, T. Maharaj: Semi-Supervised detection of extreme weather events in large climate datasets, ICLR 2017, 2017.
- 28) Rupp, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning Physical Letters.
- 29) M. Shafaei, O. Kisi: Predicting river daily flow using wavelet-artificial neural networks based on regression analyses in comparison with artificial neural networks and support vector machine models, Natural Computing Applications Forum 2016, DOI 10.1007/s00521-016-2293-9, 2016.
- 30) X. Song, H. Kanasugi, R. Shibasaki: Deep Transport: Prediction and Simulation of Human Mobility and Transportation Mode at a Citywide Level, Proceedings of 25th IJCAI-16, 2016.
- 31) T. Tagawa, Y. Tadokoro, T. Yairi: Structured Denoising Autoencoder for Fault Detection and Analysis, ACML 2014, JMLR: Workshop and Conference Proceedings 39 pp.98-111, 2014.
- 32) J. Tompson, K. Schlachter, P. Sprechmann: Accelerating Eulerian Fluid Simulation With Convolutional Networks, 2017. URL <https://arxiv.org/abs/1607.03597>, 2017.
- 33) M. Valipour, M. E. Banihabib, S. M. R. Behbahani, Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir, Journal of Hydrology 476 pp.433-441, 2013.
- 34) P. Vincent, H. Larochelle, Y. Bengio: Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion, The Journal of Machine Learning Research, 11 pp.3371-3408, 2010.
- 35) B. Vries, J. Principe: A Theory for Neural Networks with Time Delays, 1991 NIPS proceedings, 1991.
- 36) 天野 卓三, 三輪 準二, 水草 浩一: 中小河川における各種洪水予測モデルの適用性に関する研究, 河川技術論文集, 第9巻, 論文, pp. 61-66, 2003.
- 37) 稲吉 明男, 長江 幸平, 宮田 達磨: ニューラルネットワークモデルによる二級河川での洪水予測の基礎的検討, 河川技術論文集, 第9巻, pp. 179-184, 2003. 6.
- 38) 関 基, 木村 晃: 河川特性を反映したニューラルネットワーク洪水予測の精度向上, 河川技術論文集, 第19巻, 報告, 2013.
- 39) 竹村 仁志, 川村 一人: 猪名川におけるニューラルネットワークモデルをもちいた水位上昇部の予測精度向上について, 新技術・新工法部門, No. 16 pp. 1-5
- 40) 椿 涼太, 小林 健一郎: 洪水予測技術の現状と課題について, 河川技術論文集, 第19巻, 総説, 2013.
- 41) 一言 正之, 桜庭 雅明: 深層ニューラルネットワークと分布型モデルを組み合わせたハイブリッド河川水位予測手法, 土木学会論文, B1 (水工学), Vol. 73, No. 1, pp. 22-33, 2017.
- 42) 日本学術会議: 河川流出モデル・基本高水検証に関する学術的な評価 (回答), 2011.
- 43) 国土交通省河川局河川環境課監修: 水文観測業務規程関係集, (一財) 河川情報センター, 2002.

An Overview of State-of-the-Art AI Technologies for Flood Forecasting and Anomaly Detection of Water Level Time Series Data

Osamu OKADA, Kazuhiro KAKIZAWA,
Masaaki NAKAYASU, Tadashi TADOKORO

In this paper, we introduce the state-of-the-art AI technologies for flood forecasting, using the deep feed forward neural network. This is the model in which the hidden layers of the shallow feed forward neural network are multilayered, for example, layers > 100 model was reported. As a result, the deep feed forward neural network can expect to acquire the ability to learn complicated flood waveforms and improve flood forecasting accuracy. And, the technology for detecting anomalies in water level time series data by the neural network called the deep denoising auto encoder (DAE) is introduced.

In addition to the above introduction, we propose the following three ideas to improve the ability of the neural network. The first idea is a method of adding constraint condition to the neural network that the rainfall does not occurred to violate the order of time of occurrence. Why ? Simple feed forward neural network does not handle time, easily violates the order of time of occurrence and this is problem to us. The second idea is to reduce the error between forecasted flood peak and observed flood peak by removing the factor due to the flood arrival time from the neural network, that is, the influence of fluctuation in flow velocity. The third idea is to improve the ability of the neural network to learn about the features of time series data by transforming the time series data of the water level to wavelet, so-called, feature extraction before input to the neural network. At FRICS, we are developing flood forecasting software and anomaly detection software, using GPU, CUDA, Python, Tensor Flow, and Keras. We also introduce future research plan of AI.