

# 深層学習モデルを活用した降水量の逆推定法に関する検討

長岡技術科学大学 非会員 畠山周愉  
長岡技術科学大学 正会員 陸 旻皎

## 1. 緒論

日本の積雪寒冷地域では、冬季から春季における気温上昇によって短期間で融雪が生じ、河川が増水することで洪水が発生する。そのため、融雪期の流量をより正確に推定することは、積雪寒冷地域が国土の6割以上を占める日本において、とても重要であるといえる。しかし、融雪流出を高精度で推定するためには、冬季の降水量(降雪量)を正確に把握する必要があり、これが現在の大きな課題となっている。降水量は、主に貯水型雨量計、転倒ます型雨量計、レーダー雨量計の3種類の雨量計によって測定されている。前二者の問題点は、雪や雹が容易に風の影響を受けるため測定が難しいこと、雨量計に雪が付着し蓋をすることで、捕捉の障害になることである。後者の問題点は、多様な形状をした個体降水粒子の判別が難しいことである。上記以外にも、雨量計は、豪雪・山岳地帯に設置してあることが多いために管理が難しく、降雪前に撤去する必要があるといった問題点が挙げられる。今後、正確な降水量の推定法が開発されることによって、より精度の高い利活用可能なデータが得られたとき、融雪期における流出予測、水位予測、地すべり予測など多方面に渡って多大なメリットがあると考えられる。

本研究では、「第3次AIブーム」のきっかけともなった「深層学習モデル」を活用し、降水量の逆推定法を検討する。現在までに、多くの科学者によって、多種多様な深層学習モデルが開発されてきたが、本研究においては、1997年に Sepp\_Hochreiter ら<sup>1)</sup>によって提唱された Long\_short\_term\_memory (以下、LSTM)モデルを活用する。LSTMモデルを用いるメリットは、時系列データにおける長期依存性を学習できることにある。Frederik\_Kratzert ら<sup>2)</sup>は、LSTMモデルが、冬季の降水量と気温の高い春季の流出生成との間のタイムラグを学習すること、つまり、初めに認識していないはずの水文パターンを正しく再現することを示した。さらに、彼らは、LSTM内部セルの

状態は、積雪、土壌含水量、地下水の貯留へと水文的解釈できる可能性を示唆している。

本研究における降水量の推定には、入力データとして、国土交通省及び気象庁の流量データ、気象観測データを活用した。入力に用いる特徴量は、[流量、日時]を基準とする。さらに、特徴量同士、また、各特徴量と目的変数(降水量)との相関関係を把握し、多重共線性を避けた上で相関があると考えられるものをさらに入力へ追加し検証した。

本研究の目的は、冬季の降水にも対応した高精度な降水量推定ではあるが、今回は、訓練用データが豊富にあり、且つデータセット作成が容易であった土器川流域(降雪のない地域)のデータを用いて上記モデルを精度検証し、データ駆動型としてのLSTMモデルの性能を評価した。

## 2. 流量・降水量データ及びその他気象観測データ

本精度検証の対象流域として、香川県の土器川常包橋地点の上流を選択した(Fig.1)。上記の対象流域面積は、90.7 km<sup>2</sup>である。降雨量は、香川県の5地点の気象庁データ及び国土交通省の9地点のデータ、流量は常包橋地点のデータを活用した。その他気象観測データは、香川県仲多度郡多度津町における観測値を用いた。

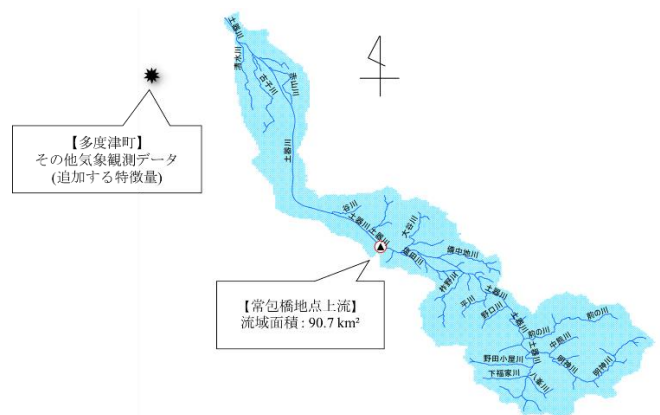


Fig.1 土器川流域\_本精度検証の対象流域

### 3. データセットの加工・前処理

土器川流域の降雨量, 流量, その他気象観測データは, 2005/12/31 23:00 ~ 1990/1/1 0:00 (1 時間値)の値を活用した. 本研究では, 深層学習モデルに流量及び気象観測データ (時系列では未来)を入力することによって, 出力として過去の降雨量を推定するため, 上記のように, 時系列を反対に並べている. 各特徴量データの欠損値は, 欠損値の前後の値で線形補間した. また, 各変数は, データスケールが異なるため, 一般的には下記 (A), (B)のような処理を加える必要があり, 今回は (B)を用いた. 気象観測データのように複数種類の変数を扱う場合は, スケール変換処理を施すことによって, データスケールの違いによる計算結果への影響を避けながら最適化させる.

(A): 標準化

平均値を 0, 分散を 1 としてスケール変換する.

$$Y = \frac{X - \mu}{\delta}$$

Y: 変換後のデータ X: 変換前のデータ

$\mu$ : X の平均  $\sigma$ : X の分散

(B): 頑健な標準化

四分位数を元にスケール変換する.

$$Y = \frac{X - Q_2}{Q_3 - Q_1}$$

Y: 変換後のデータ X: 変換前のデータ

$Q_1$ : X の第一四分位点  $Q_2$ : X の第二四分位点

$Q_3$ : X の第三四分位点

### 4. LSTM ネットワークの詳細

LSTM ネットワークの詳細を説明する前に, 標準的な Recurrent\_Neural\_Network (以下, RNN)を知る必要がある (Fig.2). 標準的な RNN におけるシーケンスの処理は, シーケンスの要素を反復的に処理する方法で行われる. また, 繰り返しの構成要素には,  $\tanh$  層のみという非常に単純化された構造である. 反復過程で得られた入力情報は, 状態として維持されていく仕組みである. しかし, 実際には, 勾配消失問題により, 長期間 (10 以上の入力シーケンス)の依存関係を学習することが不可能であるとわかっている. 上記の勾配消失問題という欠点を克服するために開発されたモデルが LSTM と Gated\_Recurrent\_Unit (以下, GRU)である. 本実験においては, 前者を活用

した精度検証を行ったため, LSTM の構造を Fig.3 に示した. LSTM 内部セルには, 情報の削除及び追加する機能がある. この情報の削除・追加機能は, シグモイド層によって構成された特殊なゲートにより制御されている. LSTM は, セル状態を維持するため上記のゲートを 3 つ所持している. Fig.3 に示した LSTM 内部セルの動作を以下に説明する. 数式内の  $W$  は重み行列,  $b$  はバイアスペクトルを意味する.

まずは, 忘却ゲートによってセルに不必要な情報を判定する.

$$F_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \quad (1)$$

次は, セルに必要な情報の保存を行う. 入力ゲートによって更新すべき情報を判定し,  $\tanh$  層によってセルに追加されるベクトル値  $\tilde{C}_t$  を作成する.

$$I_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \quad (3)$$

そして, 不必要な情報に  $F_t$  を掛けて忘却,  $I_t$  に  $\tilde{C}_t$  を掛けることで新たなセル状態を追加する.

$$C_t = F_t \cdot C_{t-1} + I_t \cdot \tilde{C}_t \quad (4)$$

最後は, 出力すべき情報を判定するため, セルへ  $\tanh$  及びゲートを適用させる.

$$O_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (5)$$

$$h_t = O_t \cdot \tanh(C_t) \quad (6)$$

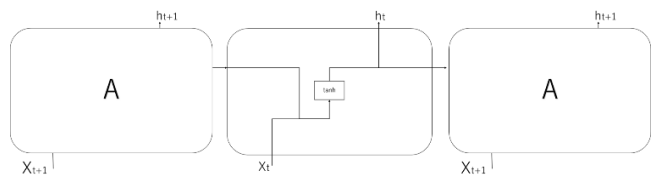


Fig.2 標準的な RNN の動作

$\tanh$ : ハイボリックタンジェント  $X_{t-1}$ : 前のセルの入力  $X_t$ : 入力  $X_{t+1}$ : 後のセルの入力  $h_{t-1}$ : 前のセルの出力  $h_t$ : 出力  $h_{t+1}$ : 後のセルの出力 A: 中央セルと同一構造

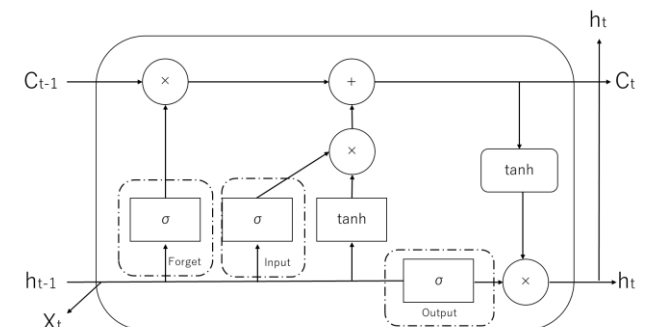


Fig.3 LSTM 内部セルの動作

$\sigma$ : シグモイド関数  $\tanh$ : ハイボリックタンジェント  $X_t$ : 入力  $C_{t-1}$ : 前のセル状態 (長期記憶)  $C_t$ : セル状態 (長期記憶)  $h_{t-1}$ : 前のセルの出力 (短期記憶)  $h_t$ : 出力 Input ( $I_t$ ): 入力ゲート Output ( $O_t$ ): 出力ゲート Forget ( $F_t$ ): 忘却ゲート +: 加算  $\times$ : 計算

## 5. 特徴量選択基準

まずは、本研究において活用すべき特徴量を選択する必要がある。データセットに含まれる目的変数と各特徴量の相関係数を Fig.4 に示した。順に降雨量 (mm/h), 日時 (sin, cos), 流量 (mm/h), 温度 (°C), 風速 (m/s), 現地気圧 (hpa), 相対湿度, 大気圧 (hpa), 蒸発散量 (mm/h) である。上記の日時データとは, 日時を sin, cos で表現 (365 日を一周, 例:  $\text{Cos}(2\pi(\text{JD} - \text{JD}_0) / 365))$  し, 周期性を抽出した特徴量である。各変数の相関関係を確認し, [流量, 日時]に加えて目的変数と相関がありそうな上位 2 つの特徴量を選択した。今回用いる特徴量は, 流量, 日時, 湿度, 現地気圧に決定し, 順に組み合わせながら降雨量を推定した。

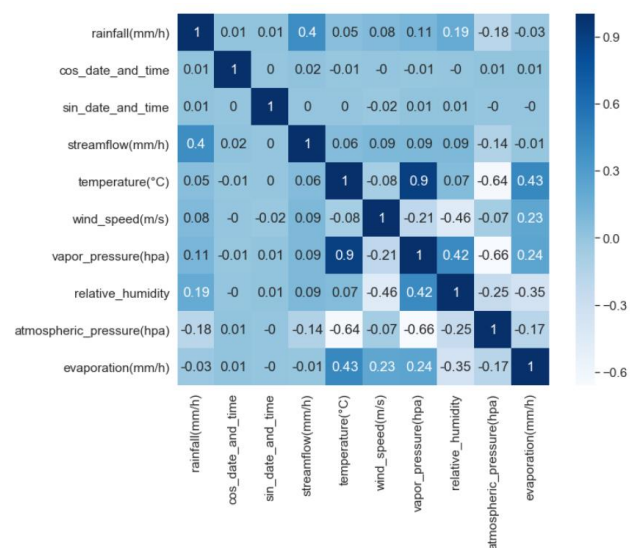


Fig.4 目的変数と各特徴量の相関係数

## 6. LSTM モデルによる降水量推定の評価法

LSTM モデルを評価する方法としては, 初期の特徴量 [流量, 日時]と, そこにさらに湿度及び現地気圧を加えた場合の降雨量推定結果を以下の指標で評価した。精度評価指標は, Root Mean Squared Error (以下, RMSE), Mean Absolute Error (以下, MAE),  $R^2$  (決定係数)を用いた。訓練時のパラメータは, Table1 に示した。固定パラメータは, Sequence\_Length, 出力

層の活性化関数 Parametric\_ReLU (以下, PReLU), 最適化関数 Adam, 損失関数 Mean Squared Error (以下, MSE), バッチサイズ 64 である。上記の固定パラメータは, 事前に試験した上で精度が高いものを選択している。Sequence\_Length は, 設定した値の入力数の依存関係を学習した上で, 目的変数を推定していく。例えば, 設定値の時間範囲  $[t, t-1, \dots, t-n]$  の入力値  $X$  の依存関係から,  $[t+N, N$ : 任意の時間分]の目標値  $Y$  を推定する。本実験では,  $n=10, N=1$  に固定し, 入力 10 時間分の特徴量の依存関係から, 次の 1 時間の降雨量を推定した。データセットの分割期間は, 2005~1993 年 (訓練用), 1992~1991 年 (検証用), 1990 年 (テスト用)である。変動させるパラメータは, LSTM\_層数 [A, B : 1, 2], 内部ユニット数 [A, B, C : 32, 64, 128]である (Fig.5, Layers - A\_1)。学習回数は, 最高を 100 回として, 最適な回数で自動停止させた。

Table1 訓練時のパラメータ

Parameter	Setting
Preprocessing	Robust_Scaler
LSTM_Layers	A_1, B_2
Units	A_32, B_64, C_128
Dense_Layers	1
Sequence_Length	10
Epochs	n = 1 To 100
Batch_Size	64
Activation_fn	tanh
Activation_fn (Ot)	PReLU
Recurrent_Activation_fn	hard_sigmoid
Optimizer_fn	Adam (lr=0.001)
Loss_fn	MSE
Training_Data	2005 - 1993
Validation_Data	1992 - 1991
Testing_Data	1990

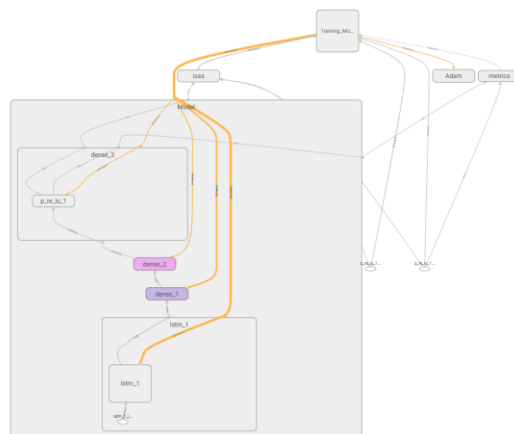


Fig.5 LSTM モデルの計算プロセス (Layers - A\_1)

## 7. 推定結果及び考察

LSTM モデルによって得られた推定結果を Table2 に示した. パターン P\_4 は, 決定係数  $R^2 = 0.70$  を示した. 決定係数  $R^2$  の良し悪しは, 明確に統計学的基準があるわけではないが,  $R^2 \geq 0.5$  であるため, モデルの説明力がある程度証明できたと考えることができる. また, P\_4 は, その他のパターンに比べ RMSE, MAE 共に値が小さく, 本検証では優れたモデルであることが示された. P\_4 の推定結果のグラフを Fig.6 に示した. 上記 LSTM モデルは, 大規模から小規模の降雨まで, おおよそ推定できていることがわかる. 本検証全体の結果から, 2つのことがわかった. 1つ目は, ユニット数, 層数を共に増やすと精度が下がりやすい傾向にあること, 2つ目は LSTM 層を 2層積み重ねると 1層より再現力が向上することである. 上記 1つ目の原因は, ユニット数が計算への影響力が強く, さらに層数が増えたため一気に過学習が起き, 精度低下に繋がったと考えられる. 上記 2つ目の要因を P\_4 の結果から考察する. P\_4 のユニット数は 32 であることから, おそらくネットワークキャパシティが不足していたため, その不足分を追加された層が補った可能性が高いと考えられる. P\_6 の 2層, ユニット数 128 のパターンは, 大きく精度が低下しており, 上記ユニット数の計算への影響と過学習に繋げて説明できる. 特徴量の基準 [流量, 日時]に加えて, 湿度及び現地気圧を追加したパターンは, 全体的に精度向上がみられなかったため, 共に特徴量の影響力が小さかったことが考えられる.

Table2 LSTM モデルの評価

Features	Pattern	[Layers, Units]	$R^2$	RMSE	MAE
流量, 日時	P_1	[1, 32]	0.63	0.91	0.21
	P_2	[1, 64]	0.53	1.03	0.31
	P_3	[1, 128]	0.66	0.88	0.24
	P_4	[2, 32]	0.70	0.83	0.21
	P_5	[2, 64]	0.69	0.84	0.23
	P_6	[2, 128]	0.44	1.12	0.31
流量, 日時, 湿度	P_7	[1, 32]	0.56	1.00	0.24
	P_8	[1, 64]	0.58	0.97	0.27
	P_9	[1, 128]	0.51	1.05	0.25
	P_10	[2, 32]	0.55	1.01	0.23
	P_11	[2, 64]	0.64	0.90	0.23
	P_12	[2, 128]	0.59	0.96	0.23
流量, 日時, 現地気圧	P_13	[1, 32]	0.56	1.00	0.29
	P_14	[1, 64]	0.58	0.97	0.24
	P_15	[1, 128]	0.52	1.04	0.25
	P_16	[2, 32]	0.67	0.87	0.22
	P_17	[2, 64]	0.64	0.90	0.23
	P_18	[2, 128]	0.64	0.90	0.20

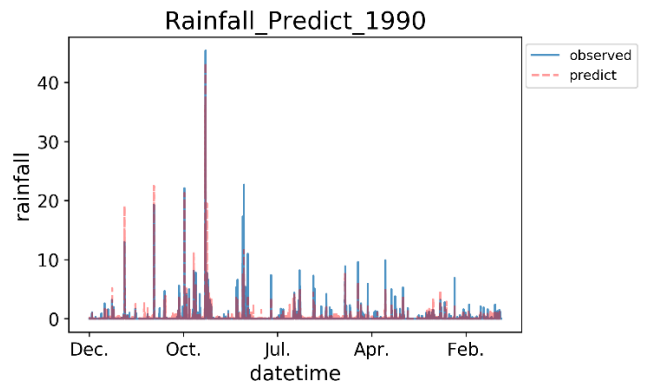


Fig.6 P\_4 の推定結果

## 8. 展望

本検証では, LSTM モデルとパターン P\_4 の組み合わせによって最も良い推定結果が得られた. 今後は, 本検証結果をもとに積雪寒冷地域を対象とした長期間の降水量推定を行う予定である. LSTM モデルが, 冬季の降水から春季の流出のタイムラグを考慮した逆推定を行えるのか検証していきたい.

## 9. オープンソースソフトウェア及び PC 性能

基本的な研究環境を Table3 に示した. 各項目は, 全て旧バージョンを用いている. Python3, tensorflow, keras は, 互換性問題に直面することが多く, 環境構築は初学者には大きな壁となっているのが現状である. 上記の3つは, アップデートによって追加・削除される機能があるため, 自分が試したい機能をいつのバージョンで扱うことができるか把握し, 且つ他項目のバージョンに対応させる必要がある.

Table3 基本的な研究環境

Main_Item	Version
ディストリビューション	Anaconda3-5.3.0-Windows-x86_64
プログラミング言語	Python 3.7.0
深層学習用フレームワーク	tensorflow 1.15.0 keras 2.2.4
機械学習用ライブラリ	scikit-learn 0.23.2
CPU	Core i7 7700HQ (Kaby Lake) / 2.8GHz / 4コア
GPU	GeForce RTX 2080

## 10. 参考文献

- Sepp Hochreiter, Juergen Schmidhuber : LONG SHORT-TERM MEMORY, Neural Computation, pp.1735-1780, 1997
- Frederik Kratzert, Daniel Klotz, Claire Brenner, Karsten Schulz, and Mathew Herrnegger : Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks, Hydrol. Earth Syst. Sci., pp.6005-6022, 2018