

# ニューラルネットワークを用いた アメダス気象データの類型化の試み

飯野 秋成\*, 廣井 俊介\*\*

(令和 3 年 11 月 30 日受理)

## Trial of Classification of Visualized Images of AMeDAS data with Deep Neural Network

Akinaru IINO\* and Shunsuke HIROI\*\*

We investigated how to visualize and categorize hourly temperature and precipitation data nationwide using the past annual meteorological data observed by AMeDAS. First, by imaging the annual meteorological data of AMeDAS meteorological data, the features that can be visually read were organized. Next, based on the idea of a deep neural network, a concrete method for dimensionally compressing the imaged meteorological data using auto encoder was shown. Furthermore, we showed the process of cluster analysis of the results of dimensional compression, visualized the similarity of data between measurement points nationwide, and evaluated its validity.

Key words: AMeDAS, Auto Encoder, Air Temperature, Precipitation, TensorFlow

### 1. はじめに

近年、気候の変化が指摘されている。落雷、突風、短時間強雨、長雨などの影響が顕在化しつつあり、災害発生時だけでなく個々の市町村ごとの特徴を生かした平時の気象防災活動の重要さが叫ばれている<sup>[1]</sup>。気象防災活動は、個々の市町村レベルの対策だけでは限界もあり、自治体間の防災ノウハウの共有に対する期待は大きい。降雨や強風、あるいは高温や低温の状況が年間を通じて類似したトレンドとなる市町村は、必ずしも近隣の自治体に限らず存在する可能性もあるため、気象災害を起こしうる要素の類似した自治体を統計分析的に示せるならば、気象災害に備えた意見交換を交わしうる平時の議論をより活性

---

\* 工学科 建築・都市環境学系 教授

Professor, Field of Architecture and Urban Environment, Department of Engineering

\*\* 工学科 建築・都市環境学系 学部生

Bachelor, Field of Architecture and Urban Environment, Department of Engineering



最も少ない記録を更新した。

このように、平年の気温・降水量と比較して特徴的な期間がある。その特徴に注目して、可視化および類型化の方法の妥当性を検討できると考えた。

## 2.2 アメダス気象データの可視化

気温の特別値を画像の1ピクセルの濃度に置き換えて表現する場合、1地点の年間の気温の特別値は、図1のように、カラム方向24時間分×ライン方向365日分(うるう年は366日分) = 8760ピクセル(うるう年は8784ピクセル)の画像で表現できる。日変化および年変化、さらには南北方向や標高の影響が濃淡により明確に可視化される。同様の処理をアメダスの年間の特別降水量のデータにも適用し、可視化の意義を考察した。

なお、本研究では、各都道府県から2~4か所程度の計測地点を選定し、全国138地点のデータを用いて試行的に分析を行うこととした。

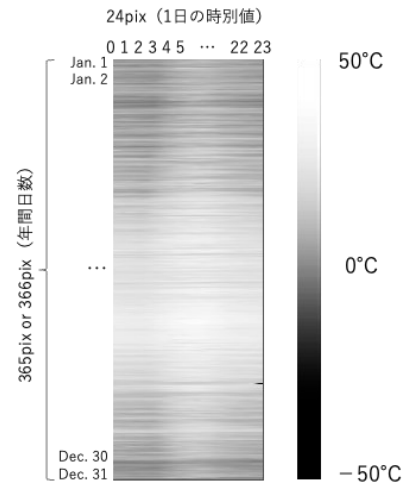


図1 アメダス気象データの画像化の方法 (桑名(三重県)の気温の特別値を用いた例)

## 2.3 AEによる次元圧縮

AEは、TensorFlowを用い、Pythonでコーディングすることにより実現させた。計算に用いたハードウェアおよびソフトウェアを表1に示す。

AEの構築では、入力層及び出力層の8784ノードに対して、中間層の数、および各中間層のノード数の検討が必要となる。次元圧縮の観点からは中間層を3層以上とする

ことの有用性が指摘されていることから<sup>[3]</sup>、図2の全結合型のディープニューラルネットワークを構築した。中間層1および中間層3のノード数は、入・出力層のノード数の20分の1程度としておき、中間層2はそのさらに20分の1程度のノード数を想定した。そして、中間層2のノード数を24~60程度に変化させながら、学習回数とともに出力層の画像が入力層の画像に近づくか否かを検討することとした。

表1 本研究で用いたハードウェアとソフトウェア

■ハードウェア	
PC	Core i7(第11世代)+16GBMem OS: Windows 10 pro
■ソフトウェア	
可視化	Gfortran, Adobe Photoshop
ニューラルネットワーク	Anaconda 3
	Tensorflow, sk-learn, opencv-python
クラスター分類および地図化	Matplotlib, Adobe Photoshop

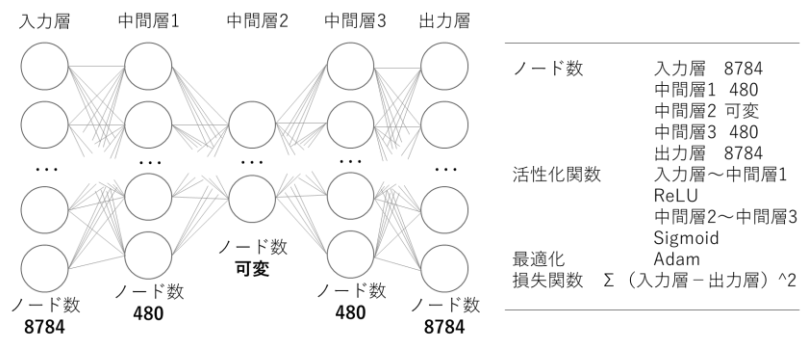


図2 本研究で構成したニューラルネットワーク

## 2.4 類型化結果の妥当性の検証と地図化

上記AEにより、中間層2に形成される28個の数字の組みあわせ、すなわち28次元のベクトルが138地点分形成される。これらは、非階層型クラスター分類の手法により類似性の高いクラスを形成できるものと考えた<sup>[4]</sup>。クラスター分類には、TensorFlowの計算結果を継承し、Jupyter Notebook上のMatplotlibを活用した。その結果は、各アメダス計測地点の北緯および東経に基づいて、日本地図上にプロットすることも可能となることから、年間の時別値のトレンドが近い地域の分布状況について、地図上のプロット結果から考察することとした。

## 3. 結果及び考察

### 3.1 アメダス気象データの可視化

図3(a)は、2020年度の全国のアメダス気象データから、各都道府県2地点以上の計測地点を選定し、計138地点の気温を画像化したものである。画像化プログラムのコンパイラにはGFortranを用いた。全国のデータから最低気温と最高気温を表現できるよう、 $-50^{\circ}\text{C}$  (黒)  $\sim +50^{\circ}\text{C}$  (白) の8bitグレースケールとしており、濃淡1段階あたり約 $0.4^{\circ}\text{C}$ 程度の精度となっている<sup>注2)</sup>。各地点の画像では、カラム方向に毎日の日中の温度変化、ライン方向には季節変化が明瞭に示されるほか、北端から南端にむけて白色となる様子や標高の高い地点の気温が低いことも明瞭に表現できる。図中に散在する黒点はデータの欠測を示しており、デジタル値0を置くことにより、のちのニューラルネットワーク処理における学習プロセスに影響を与えないよう配慮した。

図3(b)は降水量を画像化したものである。全国の分布の傾向としては、概ね西日本の観測地点の方が東日本の観測地点より降水量が大きい傾向が可視化されている。6~7月の梅雨の影響が、関東、東海以南に明瞭に可視化されている。また、9月の台風が関東甲信越、

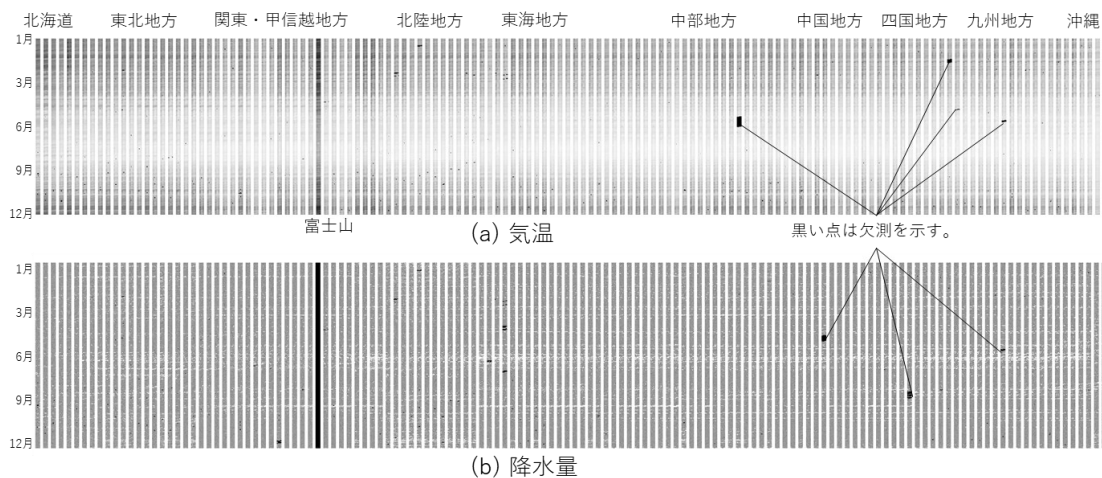


図3 全国138地点のアメダス気象データの画像化結果

東海、中部地方に白く細いラインとして読み取れていることなどが指摘できる。さらに、冬季の降雪については北陸地方にも甲信越地方にも明瞭にみられず、降雪量が少ないことが可視化されている。

### 3.2 AEによる次元圧縮

次元圧縮に用いたニューラルネットワークの詳細は、図2に示すとおり、入・出力層のノード数をそれぞれ8784、中間層1と中間層3のノード数を480に固定し、中間層2のノード数の変化と損失関数との関係を検討した。学習回数は、1地点目から138地点目までの学習を完了した時点までを1回とカウントし、10000回学習させた。損失関数は、学習回数ごとに入力層と出力層の差画像を作成し差画像の全ピクセルの2乗平均とした。

図3(a)の気温の画像を用いて、中間層2のノード数を12~60で変化させながら、学習回数と損失関数との関係を求めた結果を図4(a)に示す。ノード数12では損失関数は一定値以上減衰せず、24,28では学習回数2000以上でゆるやかな減少傾向となった。ノード数32以上で損失曲線はほぼ重なり、学習回数4000以降で繰り返し急激な上昇と下降が発生した。学習回数によってはノード数が大きいほど損失関数の値が小さくなるというわけではない。また、出力層は $-50\sim 50^{\circ}\text{C}$ を0~1の黑白濃淡画像となっていることから、気温の再現精度 $0.4^{\circ}\text{C}$  (画像濃淡0.004に相当)の2乗値 $1.6\times 10^{-5}$ 以下が損失関数の値として期待される。この判断基準のもとでは、ノード数は32以上、かつ学習回数は6000回以上が必要となる。

図3(b)の降水量の画像について、同様の検討を行った結果を図4(b)に示す。学習回数2000回以降、ノード数12~32程度で損失関数の値はほぼ単調減少している。アメダスの降水量時別値の出力層上の精度 $0.5\text{mm}$  (画像濃淡0.0036に相当)の2乗値は $1.28\times 10^{-5}$ となり、出力層の差をこの値以下とするためには、ノード数は12~32、かつ学習回数は4500回程度以上が必要となる。

以降の解析では、中間層2のノード数について、気温は60、降水量は32として、いずれも学習回数10000回の画像を用いることとした。

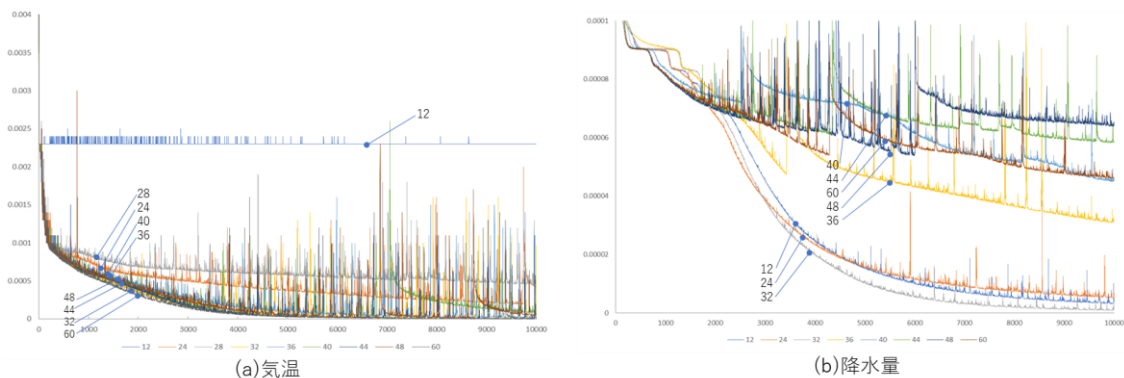


図4 中間層2のノード数ごとの学習回数と損失関数との関係

### 3.3 類型化結果の妥当性の検証と地図化

クラスター数を変化させながら k-means 法によるクラスター分類を行った。各クラスターに含まれる地点の特別観測値の標準偏差を年平均したものを図 5(a),(b)に比較した。気温については、クラスター数が 30 以上にすることで標準偏差 4°C 程度のクラスターはみられなくなり、いずれのクラスターの標準偏差も 0.0~1.0°C の範囲となった。また降水量については、クラスター数 8~10 で全てのクラスターの標準偏差が 0.12mm 以下となった<sup>注 3)</sup>。さらに図 6(a),(b)に、各クラスターの週ごとの平均気温及び積算降水量の年変動を求めた結果を示す。気温のクラスター分類結果には、主に冬季から中間期にかけての温度差が影響していることが読み取れる。また降水量のクラスターには、降雨の集中する時期の降水量のレベル、および 5 週や 37 週に見られるピーク時期の差が影響していることが読み取れる。

生成された各クラスターに含まれる観測地点を地図上にプロットした結果を図 7(a),(b)に示す。気温については、各クラスターとも地域ごとにほぼまとまっており、同クラスター内に、特に緯度方向に大きく離れた観測所の組み合わせはみられなかった。降水量については、北海道から九州まで、10 のクラスターが地域ごとにほぼまとまっている傾向が確認できる。しかしながら、クラスター No.0 の観測所の地域的なまとまりが東北地方と沖縄地方に存在した。その要因を考察するため、クラスター No.0 の東北地方 16 地点と沖縄地方

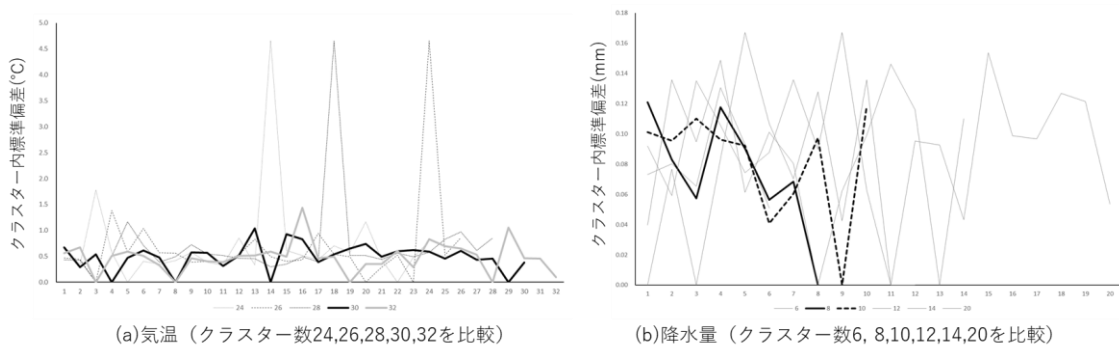


図 5 クラスター数を変化させたときの各クラスター内標準偏差 (年平均)

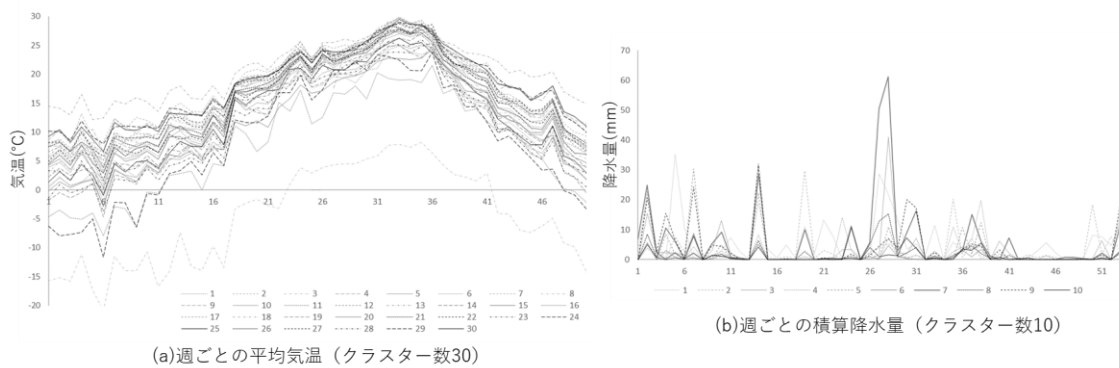
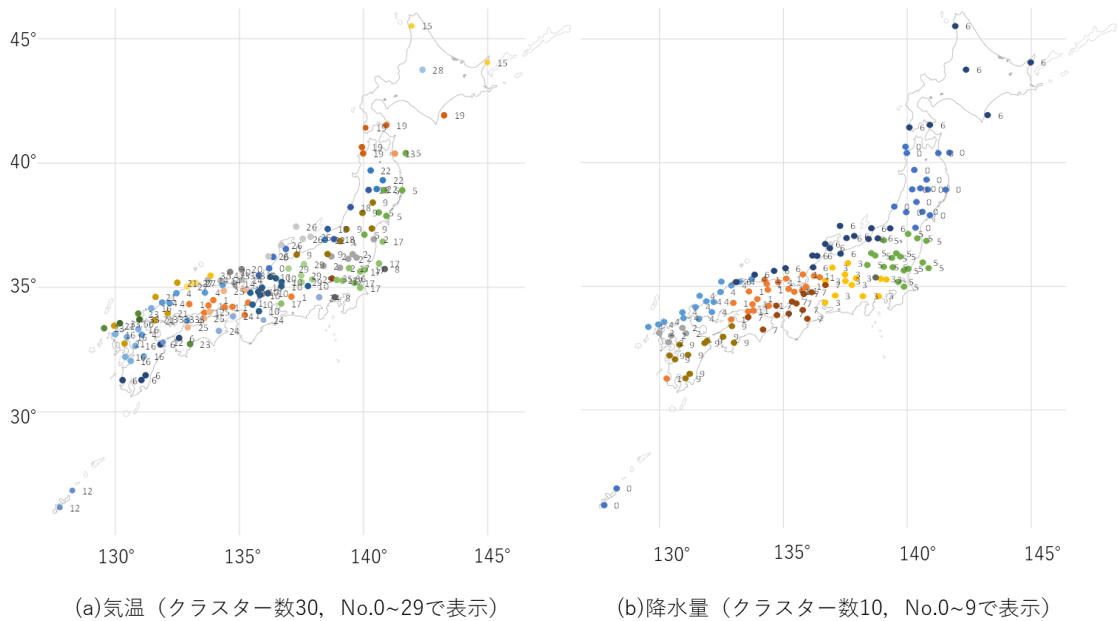


図 6 各クラスターの週平均気温および週積算降水量の年変動



※図中の数値はクラスター番号を示す。

図7 地図化結果

2地点の降水量の月積算値を年間で図8に示した。東北地方は、平年<sup>注4)</sup>に比較して4月と7月に極大、10月に極小となっている。また沖縄地方は、平年の極大が6月と9月であるのに対し、5月と8月に極大となり、11月に極小値がみられている。すなわち東北地方と沖縄地方では、月積算値自体は異なるものの、極大と極小の出現パターンが約1か月ずれた形でのトレンドの類似性がみられる<sup>注5)</sup>。

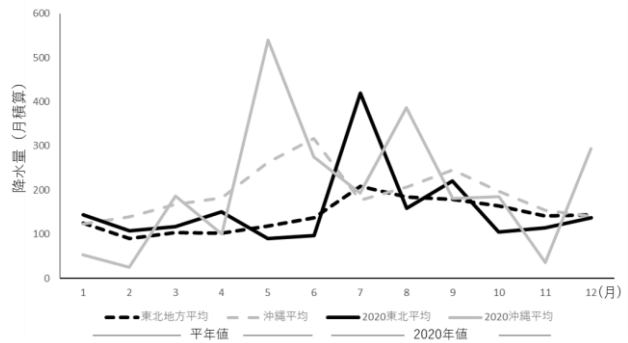


図8 降水量のクラスターNo.0における東北地方と沖縄地方の降水量の年変化の比較

その他、クラスターNo.1は近畿地方北部から瀬戸内にかけての地域的なまとまりのほか、九州南部の観測所1か所が含まれていた。また、北海道と北陸沿岸部にそれぞれクラスターNo.6の地域的なまとまりが確認された。

#### 4. 結論

本研究の知見は以下のとおりである。

- 1) 2020年のアメダス気象データにおける年間の特別気温および特別降水量を画像化し、目視で読み取れる情報を整理した。気温の画像には、昼夜や季節による比較的長周期の変化が明瞭に表現された。降水量の画像には、梅雨の多雨や夏～秋の台風の影響が強く

表れた地域がみられた。

- 2) アメダス気象データの 2020 年度の年間の気温と降水量の特別値を、中間層 3 層のオートエンコーダーで処理した。中間層 1 と 3 のノード数を 480 に固定した場合、中間層 2 の適切なノード数は、気温の場合 32 以上、降水量の場合 12~32 となった。
- 3) 中間層 2 のノードの数値を用いて k-means 法によるクラスター分類を行い、各観測地点のクラスターを地図上にプロットした。気温についてはそれぞれのクラスターは地域ごとにまとまった。降水量については、平年とは異なる降水の状況を反映して、東北地方と沖縄、北海道と北陸などのように離れた観測地点に同クラスターが形成される場合があった。

今後は、解析対象を全国の観測点約 1300 か所の 2020 年度以外の観測データに拡大することや、風速、日照時間等の観測データの分析方法の検討を行う。また、よりオートエンコーダーの多層化や畳み込みニューラルネットワークの利用の可能性など、より適切な次元圧縮の方法の検討を進める。

## 注

- 1) このパートの平年値は、文献[1]で参照されている 1981 年から 2010 年までの 30 年間の観測値の平均値を示す。
- 2) アメダスの気温の特別値の観測精度  $0.1^{\circ}\text{C}$  を表現するには量子化ビット数 8bit ではやや不足するが、本研究では画像処理プログラムや表示ソフト等における扱いやすさを優先した。
- 3) 各クラスターの気温と降水量の標準偏差を時刻別、日別、週別などで算出すれば、当然のことながら年平均で求めた値より大きくなる。クラスター内のばらつきの許容判断については適切な期間と代表値の選択方法を含めて総合的に判断する必要もある。
- 4) このパートの分析では、2021 年 5 月より気象庁 Web にて新たに公開された平年値、すなわち 1991 年~2020 年の 30 年間の観測値の平均値を用いた。
- 5) 降水量の月積算値に基づく考察のみを示したが、特別パターンの偶発的な類似の可能性も考えられる。その検証方法については今後の課題である。

## 文献

- [1]気象庁：地域における気象防災業務の在り方検討会報告書(2017.7.19)；気象庁ホームページ, 2021.11.15 アクセス
- [2]気象庁：2020 年 (令和 2 年) の日本の気象；気象庁大気海洋部報道発表(PDF), 2021.1.4
- [3]岡谷貴之：技術解説 ディープラーニング；映像情報メディア学会誌, Vol68, No.6, pp.466-471, 2014
- [4]Andreas C. Muller：Python ではじめる機械学習 (初版第 4 刷)；オライリー・ジャパン, pp.164-190, 2018