

電力使用量予測のための深層学習手法

小松 大起* 黒岩 丈介** 小高 知宏** 諏訪 いずみ*** 白井 治彦****

A Deep Learning Method for Predicting Electricity Usage

Daiki KOMATSU*, Jousuke KUROIWA**, Tomohiro ODAKA**
Izumi SUWA*** and Haruhiko SHIRAI****

(Received September 30, 2021)

In this paper, we investigate how to apply deep learnings for multimodal inputs. The target problem is the prediction of power consumption, which enable us to control the operating time of each power plant in the short term and adapt the necessary amount of fossil fuels and other resources in the long term. In this paper, therefore, we perform the prediction of power consumption by using LSTM, which is a model that can handle time-series data. We employ the combination of temperature, precipitation and/or weather as multimodal inputs, which should be meteorological factors for the power consumption. The prediction is depend on the model structure and the combination of data.

Key words :Deep Learning, Machine Learning, Electricity Usage, LSTM

1. はじめに

2011年3月11日に発生した東日本大震災以降、低コストかつ効率的な発電方法であり全体の2割から4割を占めていた原子力発電は安全性に問題があるとされ2021年5月現在では停止中の2基も合わせて合計9基の原子力発電所が稼働するのみとなっている。^[1] それにより、原子力発電よりコストの高い火力発電での発電量が増え発電コストの増加による電気料金の値上げが生じ、我々消費者の負担増加に繋がっている。^[2]

電力使用量の予測を行うことができた場合、長期的

には季節ごとの電力使用量に基づいた最適な時期での化石燃料の調達、短期的には電力使用量が増える時間帯において各発電所の稼働状態を適切に決めることが可能であると考えられる。大まかな電力使用量の推移は経験から判別が可能であると思われるが、深層学習を用いて時系列データ予測を行うことによってより詳細な予測が可能となり、今後の計画を立てることができるようにと考えられる。^[3] 適切な予測を行い適切な量の電力供給や燃料調達を行うことができれば電力会社は過剰な発電や、供給不足による突発的に高コストな発電方法、効率の悪い発電方法を行う必要がなくなる。

近年、人工知能に関する多くの分野で情報処理技術として知的処理技術の一つである深層学習が用いられている。深層学習とは、ニューロンの層が多段に組み上げられたニューラルネットワークのことを指す。^[4] 深層学習が用いられる分野としては、人物の行動認識や表情認識に挙げられるような画像処理に関わるものや、話し言葉や書き言葉などの我々が普段使うような自然言語を対象として、それらの言葉が持つ意味を解析する自然言語処理や、株価や電気使用量など

* 大学院工学研究科 知識社会基礎工学専攻 知能システム科学コース

** Fundamental Engineering for Knowledge-Based Society Intelligent Systems Science Course,
Graduate School of Engineering

** 大学院工学研究科 知能システム工学講座

* Department of Human and Artificial Intelligent Systems,
Graduate School of Engineering

*** 仁愛女子短期大学

*** Jin-ai Women's College

**** 福井大学

**** University of Fukui

の予測にも用いられる。

そこで本研究の目的は、電力使用量のみを入力とした電力使用量予測では気象要因による変化を捉えきれないと考えられるため、気象要因を入力とした際に最適な入力の組み合わせを選択し、最適なモデル構造を見つけてより詳細な深層学習を用いた予測を実現することである。

2. 深層学習モデルについて

ニューラルネットワークとは、人間の脳の仕組みから着想を得たものであり、神経回路網をコンピュータ上で表現しようと作られた数理的モデルである。人間の脳はニューロンと呼ばれる神経細胞の結びつきで情報伝達や記憶の定着が行われており、その人間の脳神経系の持つ強力な学習能力を数学的に応用することにより、画像認識や音声認識などへの利用が期待され研究が行われている。通常ニューラルネットワークは、入力層、中間層、出力層から構成されており、これらの層と層の間にニューロン同士のつながりを示すシナプス結合荷重 W があり、これを重みと呼んでいる。ニューラルネットワークの中間層は多層にすることができ、中間層が多数存在する多層構造のニューラルネットワークを用いた手法を深層学習（ディープラーニング）と呼ぶ。

本節では、用いるニューラルネットワークモデルを中心として深層学習モデルについて概要の説明を行う。

2.1 RNN (Recurrent Neural Network)

ある時刻での中間層の出力を次の時刻の層の入力として利用できるような再帰的構造を持ったニューラルネットワークモデルのことを RNN (Recurrent Neural Network) と呼ぶ。RNN は 1986 年の David E. Rumelhart による研究に基づき開発されたモデルである。^[6] 前の時刻での中間層の出力を扱うことにより、ある時点での中間層の出力がそれ以降の出力にも影響を及ぼしていくことになる。そうすることにより過去の情報を扱うことになり、時系列データを扱うことができるようになっていく。時間を追って得られたデータを時系列データと呼び、画像などのデータが一つのベクトル \mathbf{x}_n で表されるのに対し、 $(\mathbf{x}(1), \dots, \mathbf{x}(t), \dots, \mathbf{x}(T))$ という T 個のデータが 1 つの入力データ群として扱われる。時系列データには、本実験でも扱っている気温の推移や降水量の変化などの気象データ、株価や売上の推移などが挙げられ、これらのデータは一般的なデータの集まりではなく、データの並び

自体に意味を持つ。RNN において、時刻 t での入力を $\mathbf{x}(t)$ 、入力層と中間層の間の重みを U 、また中間層と出力層の間の重みを W 、 f, g を活性化関数とすると中間層の出力値 $\mathbf{s}(t)$ とネットワークの出力 $\mathbf{y}(t)$ は、以下のように表すことができる。

$$\mathbf{s}(t) = f(U\mathbf{x}(t) + W\mathbf{s}(t-1)) \quad (1)$$

$$\mathbf{y}(t) = g(W\mathbf{s}(t)) \quad (2)$$

2.2 LSTM (Long Short-Term Memory)

LSTM (Long Short-Term Memory) は RNN の拡張として 1997 年に Sepp Hochreiter 及び Jürgen Schmidhuber によって発表された時系列データを扱うモデルである。^[5] Long Term Memory (長期記憶) と Short Term Memory (短期記憶) という神経科学における用語から取られており、RNN はその構造上、時刻が離れているデータ間の依存関係を学習することが難しく、その弱点を克服するために開発されたモデルが LSTM である。LSTM は近い過去を扱うことのできる短期記憶と遠い過去を扱うことのできる長期記憶が可能である。LSTM は RNN の中間層を LSTM block と呼ばれるメモリと入力ゲート、忘却ゲート、出力ゲートの 3 つのゲートを持つブロックに置き換えることで実現されている。メモリは入力の依存性を記憶し、入力ゲートはメモリへ与える入力の調節を行い、忘却ゲートはメモリ中の値をどれだけ忘れるかを調節、出力ゲートはメモリ中の値のどれだけを活性化関数に与えるかの調節を行う。また、中間層の内部状態を以下のように表す。

$$\mathbf{f}(t) = \sigma_g(U_f\mathbf{x}(t) + W_f\mathbf{h}(t-1)) \quad (3)$$

$$\mathbf{i}(t) = \sigma_g(U_i\mathbf{x}(t) + W_i\mathbf{h}(t-1)) \quad (4)$$

$$\mathbf{o}(t) = \sigma_g(U_o\mathbf{x}(t) + W_o\mathbf{h}(t-1)) \quad (5)$$

$$\mathbf{c}(t) = \mathbf{f}(t) \odot \mathbf{c}(t-1) + \mathbf{i}(t) \odot \sigma_c(U_c\mathbf{x}(t) + W_c\mathbf{h}(t-1)) \quad (6)$$

$$\mathbf{y}(t) = \mathbf{o}(t) \odot \sigma_h(\mathbf{c}(t)) \quad (7)$$

RNN と同じく入力層と中間層の間の重みを U 、中間層と出力層の間の重みを W 、 σ を活性化関数とすると、 $\mathbf{f}(t)$ は忘却ゲート、 $\mathbf{i}(t)$ は入力ゲート、 $\mathbf{o}(t)$ は出力ゲートでの活性化ベクトル、 $\mathbf{c}(t)$ はメモリの活性化ベクトル、中間層の出力値 $\mathbf{h}(t)$ とネットワークの出力値 $\mathbf{y}(t)$ は、以上のように示される。 \odot はアダマール積を示し、ベクトルの要素ごとの積を行う演算子である。また、本研究では LSTM を用いる。

3. 実験方法

本節では、入力に用いるデータについての詳細及び実際に用いるネットワークについての説明を示す。また本実験で用いるデータは、予測を行いたい電力使用量に加えて気温、降水量、天気などのデータを用いる。用いるネットワークはLSTMを用いる。

3.1 用いるデータについて

3.1.1 電力使用量

本実験では東京電力パワーグリッド社が提供している電力使用量のデータ2016年4月から2020年12月のデータを用いる。電力使用量の推移を図1から図5に示す。

図1, 図2では、2019年と2020年の年間での電力使用量の推移を示している。それぞれ縦軸が電力使用量の 10^4 kW, 横軸が日にちを示している。春と秋に使用量が減り夏と冬に使用量が増加するという、ともに同じような傾向が見て取れる。

図3から図5は2020年4月の第2週から第4週の電力使用量の推移を示している。こちらも同様に縦軸が電力使用量の万kW, 横軸が時間を示している。図3より、赤系統の色は休日である土日を示し、青系統は平日を示している。通常は、平日の方が休日よりも活動が多いため電力使用量が増えていることを読み取ることができる。しかし、図4と図5を見ると外れ値的な動きをしている日があることが読み取れる。図4だと13日と18日、図5だと20日である。1日に100mm降ると大雨と呼ばれるがそれぞれ13日は132mmで18日は90mmと大雨に準ずる降水量であり、そのために電力使用量が増えたのだと考えられる。20日の場合は19日が平均気温14.7℃, 21日が14.2℃なのに対して9.7℃と約5℃低下し、気温の低下により電力使用量の増加に繋がったのだと考えられる。このように電力使用量の推移は気温や降水量等の気象条件に左右されるため、電力使用量の予測には気象要因を考慮する必要があると考えられる。

3.1.2 気温

気温のデータは気象庁が提供している、2016年4月から2020年12月のデータを用いる。図6に各年の気温の推移を示す。縦軸が温度(℃)を示しており、横軸が日にちを示している。基本的には、他の年の気温の推移も同じく夏場を頂点とする負の2次関数の形をとっている傾向を読み取ることができる。

3.1.3 降水量

降水量のデータは気象庁が提供している、2016年4月から2020年12月のデータを用いる。例として2020年の降水量の推移を図7に示す。縦軸が降水量(mm)を示しており、横軸が日にちを示している。最大降水量だけを見ると8月から10月ごろ、4月ごろ、梅雨である6月ごろに多く降っていることを読み取

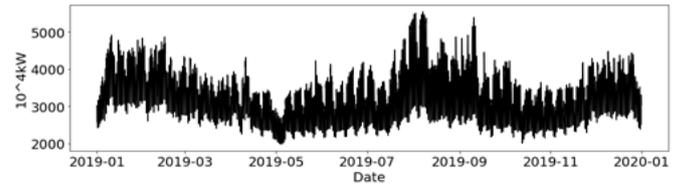


図1: 2019年の電力使用の推移

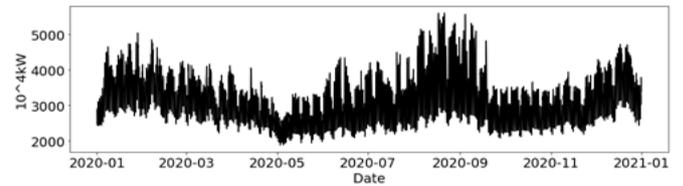


図2: 2020年の電力使用の推移

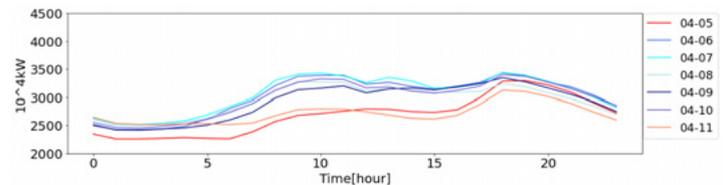


図3: 2020年4月2週目の電力使用の推移

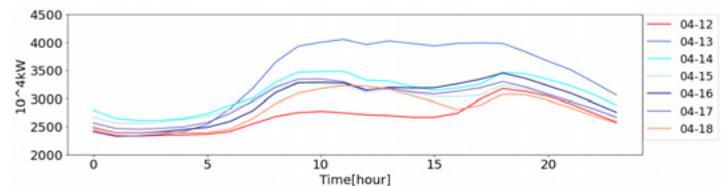


図4: 2020年4月3週目の電力使用の推移

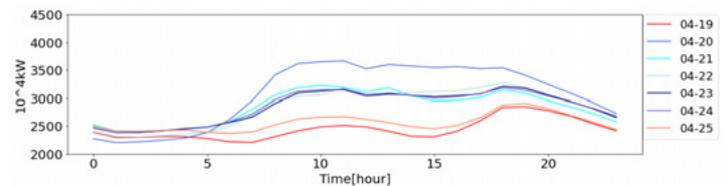


図5: 2020年4月4週目の電力使用の推移

表 1: 天気概況による数値付け

天気概況	対応する数値
「快晴」	0.0
「晴」	0.0
「曇」	0.5
「薄雲」	0.5
「大風」	0.3
「霧」	0.6
「霧雨」	0.7
「雨」	1.0
「大雨」	1.0
「暴風雨」	1.0
「みぞれ」	1.0
「雪」	1.0
「大雪」	1.0
「暴風雪」	1.0
「地ふぶき」	1.0
「ふぶき」	1.0
「ひょう」	1.0
「あられ」	1.0
「雷」	1.0

ることができる。また、他の年の降水量の中には 200 mm を超えるような外れ値的な降水量を記録することもある。

3.1.4 天気

天気のデータは気象庁が提供している、2016 年 4 月から 2020 年 12 月のデータを用いる。本研究では、天気を入力として扱う際に雨を 1.0、曇りを 0.5、晴れを 0 として各天候に数値付を行なった。各天気概況における数値付けを表 1 に示す。例として 2020 年の天気の推移を図 8 に示す。また、縦軸は天気を与えた数値、横軸は日にちを示す。雨を示す数値が 1.0 の点を見ると大まかに降水量が多かった期間において雨の数が増えている傾向を読み取ることができる。

3.2 データ間の相関関係

データ間の特性が離れすぎていると入力として適切ではないと考えられるため、これらのデータ間の相関関係について調べる。相関関係は相関係数を用いて計算を行なった。本計算は python の拡張モジュールである数値計算を行うためのライブラリ numpy を用

表 2: 電力使用量と各項目の期間ごとの相関係数

	気温	降水量	天気
時間	0.085	-0.024	-0.014
日にち	-0.069	-0.069	-0.13
週	-0.075	-0.20	-0.050

いた。また、相関係数 r は以下の数式で示される。

$$r = \frac{S_{XY}}{S_X \times S_Y} \quad (8)$$

S_X は X の標準偏差、 S_Y は Y の標準偏差、 S_{XY} は X と Y の共分散を示す。また、電力使用量は時間ごとのデータが与えられているため、それらを足し合わせたものを日にちごとのデータと週ごとのデータとして用いている。それ以外のデータについては公開されている時間ごとのデータから平均を算出することで日にちごとと週ごとのデータとして扱っている。

それぞれの期間ごとの相関係数を見ると、時間ごとでは気温との間に全体で唯一の正の相関を持っていることがわかる。また、日にちごとでは天気との間に強い負の相関を持ち、週ごとでは降水量との間に強い負の相関を持っていることがわかった。

3.3 用いるモデル構造について

本実験では、電力使用量の予測に LSTM を用いる。プログラミング言語には python を用いて、主に python のライブラリである tensorflow-1.15, pandas-1.2.4, numpy-1.18.5, scikit-learn-0.24.2 を用いて実装を行なった。

tensorflow は Google が開発したディープラーニングのフレームワークとして知られるオープンソースの機械学習ライブラリである。また、tensorflow はオープンソースのニューラルネットワークライブラリとして知られている keras と統合されたため tensorflow からインポートすることにより keras の機能を用いることが可能となっている。tensorflow は機械学習において直感的な構造の構築が可能となっている。

pandas 及び numpy はデータを扱うためのライブラリとして知られており様々な計算処理を行うことができる。相互に変換が可能であるためデータの用途によって使い分けることが可能である。

scikit-learn は主に機械学習のアルゴリズムを実装する際に用いられるライブラリであり、本実験では RMSE (Root Mean Square Error) を用いたモデルの

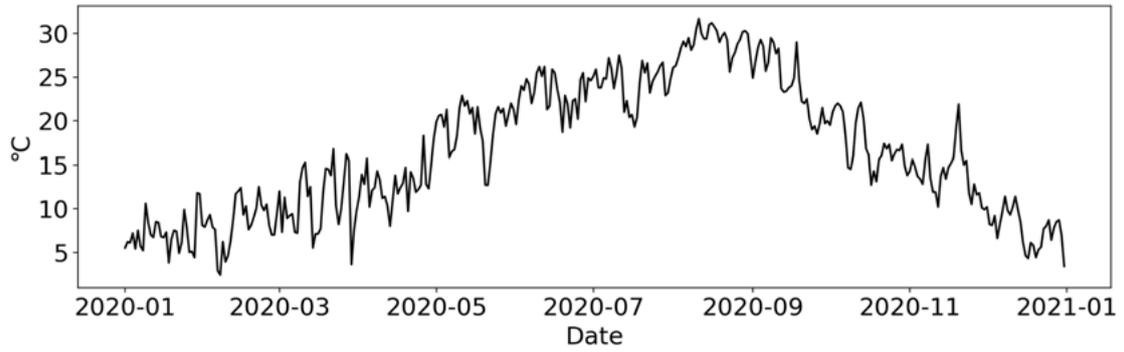


図 6 : 気温の推移

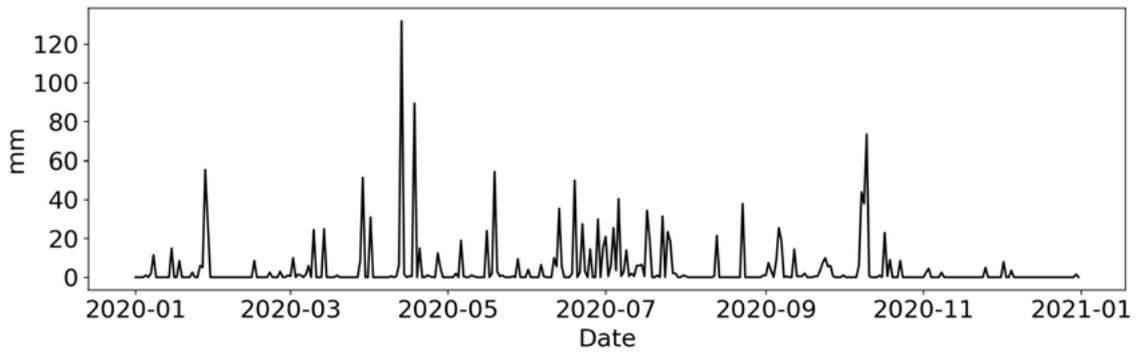


図 7 : 降水量の推移

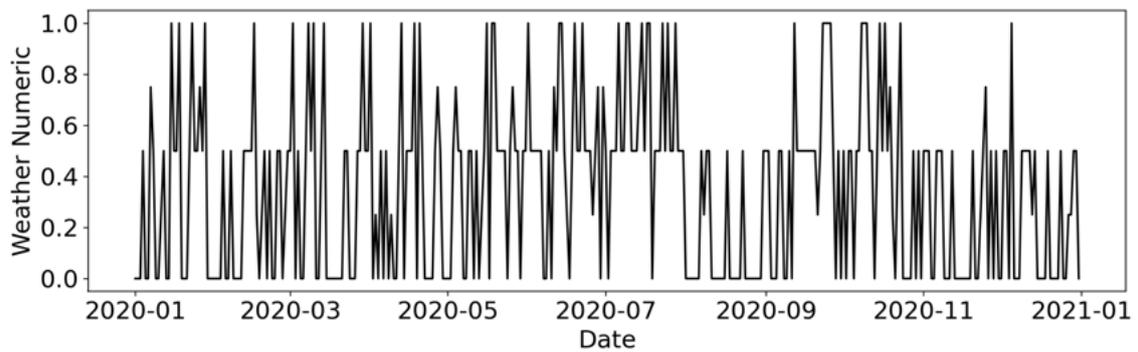


図 8 : 天気推移

評価に使用した。また、numpy 及び本実験では用いていないが scipy とのやりとりのみに対応しているため pandas で扱ったデータを numpy 形式に変換する必要がある。

本実験で用いる入力には時間ごと、日にちごと、週ごとそれぞれにおいて電力使用量に加えて気温、降水量、天気、気温と降水量、気温と天気、降水量と天気、気温と降水量と天気、各々電力使用量とともに入力として与えた場合、及び電力使用量のみを入力とした場合で実験を行った。また Train data と Test data は入力を 8 対 2 として分割を行い、予測結果は時間ごとなら 1 時間後、日にちごとなら 1 日後、週ごとなら 1 週後の結果の予測を行なった。最適化アルゴリズムには Adam を用い、学習率の初期値は 0.001 とし、バッチサイズは時間ごと 64、日にちごと 3、週ごと 1 として、epoch 数は validation loss が 10 回以上改善しなくなるまでとした。また層数は 1 層、2 層、3 層と変化させ、中間層の素子数は 100 個、200 個、300 個と変化させて電力使用量の予測を行なった。予測結果の評価には RMSE を用いた。

3.3.1 RMSE (Root Mean Square Error)

本実験で用いる指標である RMSE は、平均平方根二乗誤差と呼ばれ、入力と予測結果の誤差が大きいほど大きい値になる特徴を持つ評価指標である。そのため、値が大きく変動する点での予測の評価を主に行いたい本実験では適した指標であると考えられる。また、RMSE は以下の式で表される。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

上記の n はデータ数を、 y と \hat{y} は入力と予測値を示す。

4. 実験結果

時間ごと、日にちごと、週ごとの予測結果の推移を図 9 から図 11 に示す。それぞれの結果において、気温、降水量、天気、気温と降水量、気温と天気、降水量と天気、気温と降水量と天気、各々電力使用量とともに入力として与えた場合、及び電力使用量のみにおいて層数を 1 層、2 層、3 層と変化させ、素子数をそれぞれ 100 個、200 個、300 個としたときの結果を示している。

一番精度の良い予測結果はそれぞれ、時間ごとでは電力使用量のみを入力として、層数 2 層、素子数 300 個の時に RMSE が 0.0130、日にちごとでは気温と降水量を入力として、層数 3 層、素子数 100 個の時に RMSE が 0.0744、週ごとでは気温と降水量と天気を入力として、層数 2 層、素子数 200 個の時に RMSE が 0.0972 という結果を示した。時間ごとの予測結果の下位 4 つでは必ず気温の情報が入力として与えられていた。日にちごとの予測結果では、時間ごとの予測結果と比較して層数が少ない時に比べて、層数を増やした際の予測結果のばらつきが見られた。週ごとの予測結果では、下位 3 つが入力として気温を含んでいることがわかった。

5. 考察

予測精度は週ごと、日にちごと、時間ごとの順で高くなっていることがわかる。時間ごとではデータ数が多く、単純に日にちごとは時間ごとの 24 分の 1 そして週ごとはさらに 7 分の 1 のデータ数になっていることから、データ数が減ってしまうため予測精度の低下が見られたのだと考えられる。

時間ごとの予測結果では、一番精度が良い入力は電力使用量のみを入力とした場合であった。時間ごとの予測の際には予測を行うにあたり十分な量のデータがあったと考えられるため、入力として複雑になりすぎない電力使用量のみを入力とした時に精度の上昇が見られたのだと考えられる。また、精度の下位 4 項目では全て気温を入力として含んでいることがわかった。時間ごとの気温は電力使用量と唯一正の相関をもっていることから正の相関を持つことが精度の低下につながる可能性があると考えられる。

日にちごとの予測結果では、気温と降水量を入力した際に精度が一番良いことがわかった。しかし、他の項目でも時間ごと、週ごとの結果と比べて差が小さいことがわかった。気温、降水量、天気それぞれを入力した際には、相関係数の値の順で精度が高いことがわかった。時間ごとの予測は日にちごと、週ごとと比較して安定しているためデータ数が適していると考えられる。

週ごとの予測結果では、気温と降水量と天気を入力した際に精度が一番高かった。精度の下位 3 つが気温を入力に含んでいることがわかった。下位の予測結果の全てが気温を含んでいるわけではないが、時間ごとのデータにおいて電力使用量と正の相関をもっている気温が入力として含まれることで精度が下がってしまっているため週ごとのデータにおいても相関係

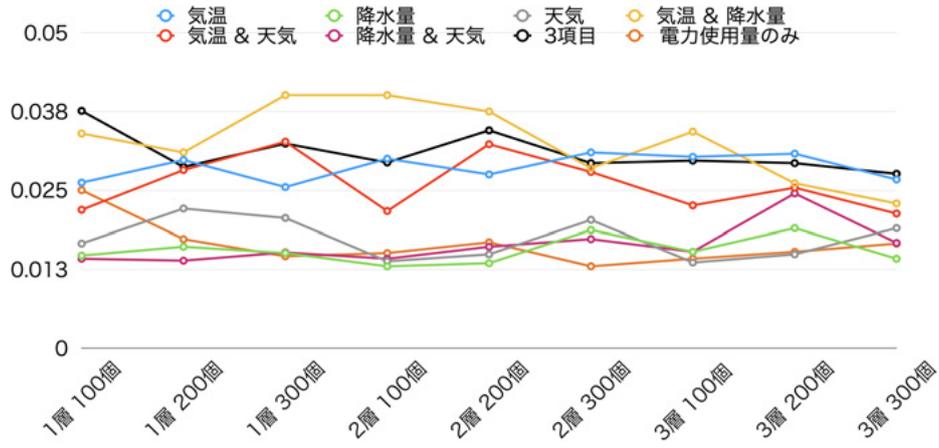


図 9：時間ごとの予測結果

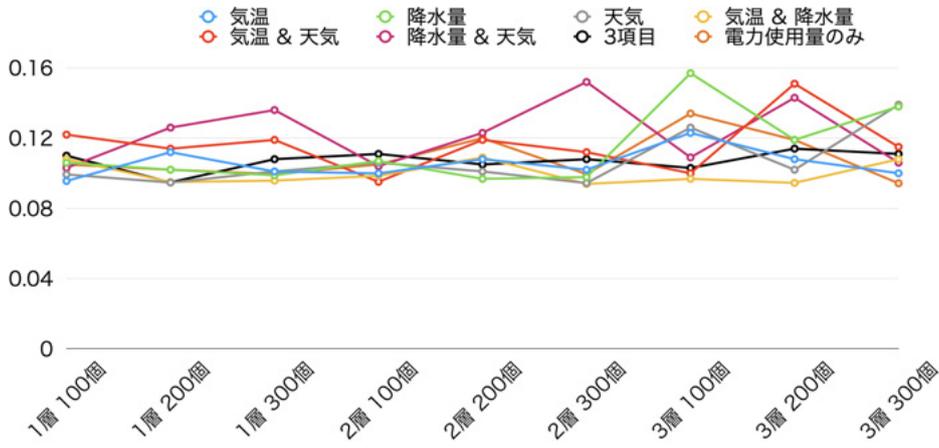


図 10：日にちごとの予測結果

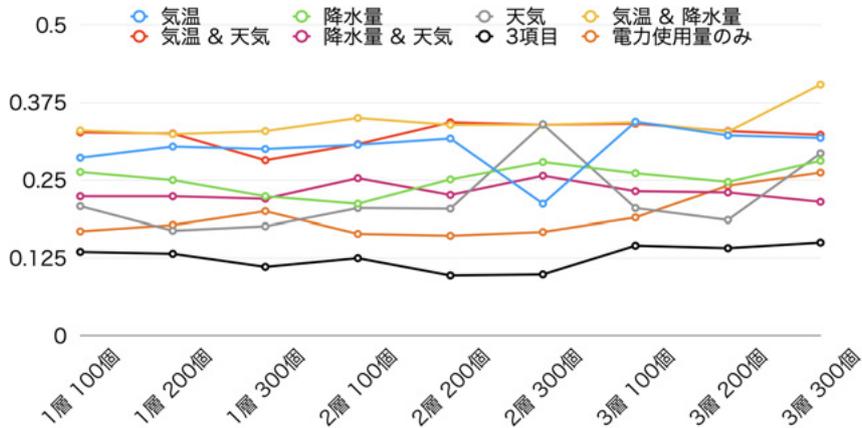


図 11：週ごとの予測結果

数以外の関係が存在していると考えられる。

また、層数は時間ごと、日にちごと、週ごとそれぞれにおいて2層及び3層の場合に良い精度を示していることから1層だと単純であるためこのデータ数を扱うためには、2層以上のモデルが適しているのだと考えられる。素子数については、数によって大きな差が見られないことから層数よりは重要視する必要のない項目であると考えられる。

以上のことから、精度の上昇には層数が少なすぎないモデルを用いること、データ数を増やすことが適していると考えられる。

6. まとめ

本研究では、気温、降水量、天気を組み合わせてマルチモーダルな入力として電力使用量の予測を行なった。その際には総数や素子数のモデル構造の変更を行い、RMSEの値による精度の違いを調べた。期間ごとの最良な予測結果はそれぞれ、時間ごとでは電力使用量のみを入力として、層数2層、素子数300個の時にRMSEが0.0130、日にちごとでは気温と降水量を入力として、層数3層、素子数100個の時にRMSEが0.0744、週ごとでは気温と降水量と天気を入力として、層数2層、素子数200個の時にRMSEが0.0972という結果を示した。相関係数による精度の違いも確認することができたが、精度向上に最適なのはデータ数が多いことだとわかった。

6.1 今後の展望

本実験では全ての入力を始めから等しく入力として扱うことで電力使用量の予測を行なったが今後作成したいモデルとして考えているモデルが電力使用量のみを主な入力として、他の項目を最終的な調節に用いることのできるモデルの作成を考慮する。時間ごとの予測では相関係数による精度の違いを見ることができたが、週ごとの予測では規則性は見つかったものの、相関係数でそれを示すものはなかったため他の関係性の評価を行うことを考える。また、RMSEを評価指標として用いたが、他の評価指標も用いることで結果を多角的に見られるようにしたいと考える。

参考文献

- [1] 日本の原子力発電所の状況, <
<https://www.enecho.meti.go.jp>> (2021/09).
- [2] 松尾 雄司, 永富 悠, 村上 朋子: 有価証券報告書を用いた火力・原子力発電コスト構造の分析, Journal

of Japan Society of Energy and Resources, Vol.33, No.5, 1-2(2012).

- [3] 鎌田 真, 市村 匠: リカレント構造適応型 Deep Belief Network による時系列データの学習, 計測自動制御学会論文集, Vol.54, No.8, 628-639(2018).
- [4] 浅川伸一: python で体験する深層学習, コロナ社, 302-303(2016).
- [5] Sepp Hochreiter, Jurgen Schmidhuber: Long short-term memory, Neural Computation, 9(8), 1735-1780(1997).
- [6] David E. Rumelhart, Geoffrey E. Hinton, Ronald J. Williams: Learning representations by back-propagating errors, Nature, 323, 533-536(1986).