東海大学紀要工学部
Vol.60, No.1, 2020, pp. 33-38
DOI: 10.18995/24343633.60-1.33

機械学習による揚水発電所週間運用手法の開発

田中　貴裕[*1]　石丸　将愛[*2]

Development of Weekly Operation Method for a Pumped Storage Hydropower Plant by Machine Learning

by

Takahiro TANAKA[*1] and Masachika ISHIMARU[*2]

(Received on Mar. 31, 2020 and accepted on Jun. 19, 2020)

Abstract

"The Act on Special Measures Concerning Procurement of Electricity from Renewable Energy Sources by Electricity Utilities" was enacted in August 2011 in part to accelerate the interconnection of photovoltaic power generation. Some electric power companies have performed trial calculations showing that photovoltaic power generation output will exceed the total demand for power. In such cases, pumped storage hydropower plants have been used to pump up water to an upper reservoir during the daytime.

In recent years, artificial intelligence has been used in various fields, including machine learning. The support vector machine and linear regression described in this paper are kinds of machine learning, and are used to understand regularity and patterns from various data, and to clarify and predict the current situation.

In this study, we used weekly operation data of a pumped storage hydropower plant, which was calculated by fuzzy reasoning, as training data and learned it using a machine learning algorithm to create a prediction model. When new data was input to the prediction model, we verified how close it was to the operational data of fuzzy reasoning.

Keywords: Renewable energy, Pumped storage hydropower plant, Weekly operation, Machine Learning, Photovoltaic(PV) power generation

1. はじめに

日本の電力供給は，化石燃料による発電が主となっている[1]．しかしながら，CO_2 排出規制を考慮すると，化石燃料による発電に頼ることは避けなければならない．このような状況のなか，クリーンな再生可能エネルギー(Renewable energy)を利用した発電が注目されている．再生可能エネルギーとは，自然界で繰り返し起こる現象から永続的に利用できるエネルギー源であり，その代表例に太陽光発電や風力発電，地熱発電等がある．現在では，2011 年に制定された「電気事業者による再生可能エネルギー電気の調達に関する特別措置法」により，再生可能エネルギーの普及が進んでいる．

太陽光発電や風力発電からの出力は，天候により大きく変動するため，これらの変動電源が電力系統に大量連系された場合，電力の安定供給に影響が生じることが懸念されている．また，需要の低い時間帯に大量に発電すると，供給力過多となるため太陽光発電の出力抑制が必要である．その対策として，揚水発電所の昼間揚水が実施されている[2]．Fig. 1 に太陽光発電大量連系時の需給バランスの概要を示した．

揚水発電所の運用は，週間単位で行われることが多く，需要量と揚水発電所の上池の貯水量を考慮しながら，出力値を決定している．昼間揚水を行う場合は，需要量と揚水発電所の上池の貯水量に加えて，太陽光発電の出力を考慮し，出力値を決定することとなる．

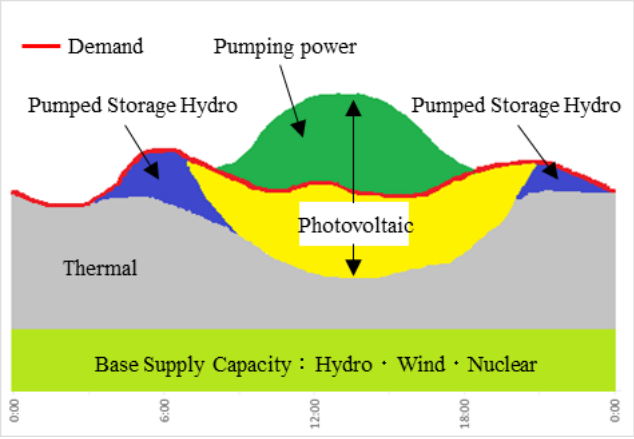


Fig. 1 Load curve & generation types.

*1　工学部電気電子工学科学部生
*2　工学部電気電子工学科講師

近年我々にとって，人工知能はより身近な存在となっている．世界中の様々な企業が人工知能の研究を行っている．メールのスパムフィルタなどがその代表であり[3]，注目度は非常に高い．人工知能の一分野として発展した機械学習は，数値・文字・画像・音声など多種多様なデータの中から，規則性・パターンを発見し，現状の把握や将来の予測を行うことを目的として広く使用されている.機械学習の学習手法は大きく分けて,「教師あり学習」と「教師なし学習」の２つに分類される．

本研究では，ファジィ推論[4]によって作成された，週間運用データを訓練データとして機械学習アルゴリズムに学習させ，予測モデルを作成した．検証として，テストデータが入力された場合,予測モデルが出力する値が,ファジィ推論モデルにどの程度近似するか確認した．

2. 機械学習

2.1 教師あり学習

教師あり学習とは，入力データが与えられたとき，これに対する出力を正しく予測することが目的の学習方法である．訓練データと呼ばれる入出力のペアの事例を与え，これをもとに，新しい入力データが来た時にそれに対する正しい出力をするようなモデルを構築する．訓練データでは，入力を説明変数(特徴量), 出力を目的変数と呼ぶ．学習の出力が質的データであれば分類，量的データであれば回帰と呼ばれる．

2.2 分類

分類は教師あり学習の 1 つである．過去の観測に基づき，新しいデータが，どの部類に分けられるかを予測することが目的となる．メールのスパムフィルタは，二値分類の典型的な例である．訓練データによっては，複数に分類することも可能である．このような複数に分類するものを多値分類という．本研究においては，揚水発電所の運用を検証するため，発電，揚水，待機の 3 つに分類する必要がある．二値分類の概念を Fig. 2 に示した．

この二値分離の概念では，入力が 2 つであり，X_1とX_2として関連付けられる．ここで，教師あり学習のアルゴリズムを使い，分類境界(Decision Boundary)を学習することで，新しいデータを 2 つに分類することができる．

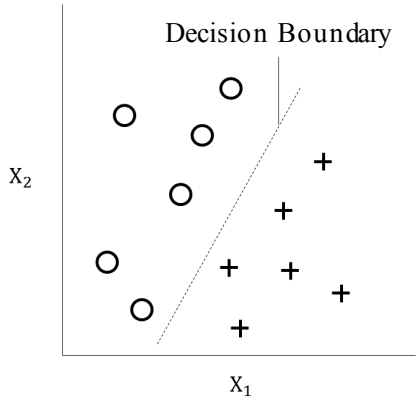


Fig. 2 Concept of binary classification.

分類のアルゴリズムの 1 つとしてサポートベクトルマシーン(Support Vector Machin：SVM)がある[5]．SVM は，マージンを最大化することにより分類を行う分類器である． マージンとは，分類境界とこの分類境界に最も近いデータとの間のマハラノビス距離として定義される．マハラノビス距離とは，ユークリッド距離を標準偏差で除して標準化した距離である．分類境界と最も近いデータは，サポートベクトルと呼ばれる．

Fig. 3 に SVM の分類境界とマージンについて示した．分類境界のマージンを大きくすると，汎化誤差が小さくなる．これに対して，マージンの小さいモデルは過学習に陥りがちである．

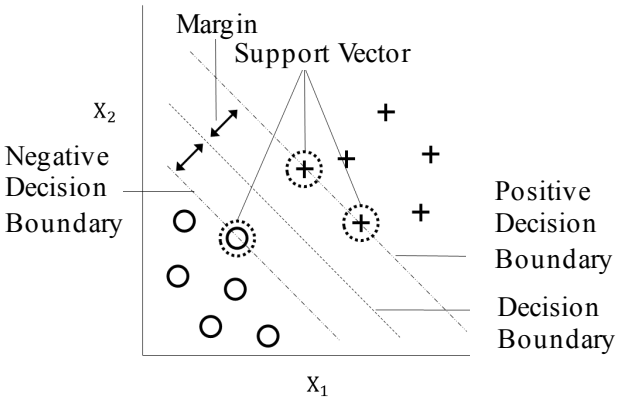


Fig. 3 Classification boundaries and margins of SVM.

2.3 回帰

回帰は教師あり学習のもう 1 つのサブフィールドである．回帰モデルは，連続値をとる目的変数を予測するために使用される．このため，変数間の関係を理解や，予報を作成するなどの形で応用されたり，科学の様々な問題に取り組んだりするのに適している．例としては，企業の今後数か月後の売り上げを予測することなどが挙げられる[6]．

単変量の線形回帰(Linear Regression)の目的は，単一の特徴量(説明変数 x)と連続値の応答(目的変数 y)との関係をモデルとして表現することである．説明変数が 1 つだけの線形モデルの方程式は以下のように定義される．

$$y = w_0 + w_1 x \qquad (1)$$

ここで，w_0 は y 軸の切片を表し，w_1 は説明変数の係数を表す．ここでの目標は，説明変数と目的変数の関係を表す 1 次式の重みを学習することである．この重みを利用すれば，訓練データに含まれていなかった説明変数に対して応答が予測できる．

先に定義した 1 次式を前提とすると，線形回帰については,「サンプル点を通過する直線のうち最も適合するものを見つけ出すこと」として理解することができる．Fig. 4 に線形回帰の概要について示した．

この最も適合する直線は，回帰直線(Regression line)とも呼ばれる．回帰直線からサンプル点への直線は残差(Residual)と呼ばれ，予測の誤差を表す．

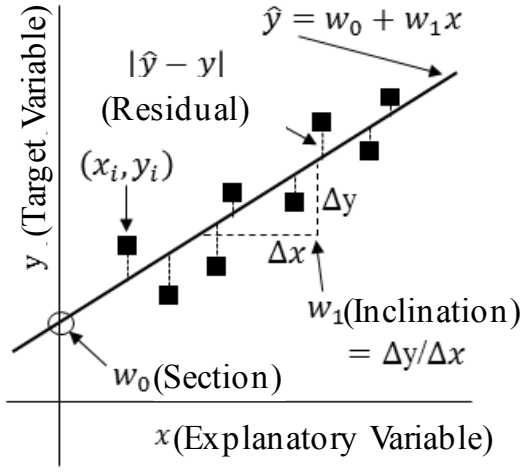


Fig. 4 Concept of linear regression.

説明変数が，1 つだけの線形回帰モデルは，単回帰(simple linear regression)という特殊なケースもある．また，線形回帰モデルを複数の説明変数に合わせて一般化することも可能であり，これは重回帰と呼ばれる．

$$\hat{y} = w_0 x_0 + w_1 x_1 + \cdots + w_m x_m = \sum_{i=0}^{m} w_i x_i = \boldsymbol{w}^T \boldsymbol{x} \qquad (2)$$

ここでw_0は，$x_0 = 1$として y 軸の切片を表している．

サンプルに近似直線となるような係数を求める手法として，最小二乗法がある．最小二乗法は，サンプル点と直線との差(残差)の 2 乗の和を最小化とする係数を求めることができる．

$$\mathrm{J}(\boldsymbol{w}) = \frac{1}{2}\sum_{i=0}^{n} \left(y^{i)} - \hat{y}^{(i)}\right)^2 \qquad (3)$$

ここまでは，説明変数と目的変数の関係が線形であることを前提としていた．この前提が外れる場合は，多項式の項が追加された，「多項式回帰モデル」を使用する方法がある．

$$\hat{y} = w_0 + w_1 x + w_2 x^2 \cdots + w_d x^d \qquad (4)$$

ここで，d は多項式の次数を表す．多項式回帰を使って非線形関係をモデリングすることは可能だが，線形回帰の係数 w については，線形であるため，重回帰モデルとみなされる．

2.4　学習モデル構築

Fig. 5 に，予測モデルに機械学習を使用する場合の一般的なワークフローを示した．

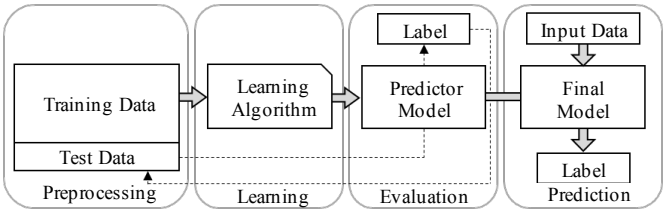


Fig. 5 Machine learning workflow.

2.5　訓練データの前処理

生のデータが，機械学習アルゴリズムの性能を最適化するのに適した形式で提供されることは，滅多にない．機械学習を適用する際，常にデータの前処理が最も重要なステップの 1 つとなる．多くの機械学習アルゴリズムでは，最適な性能を得るために，データ値の大きさの水準を意味する特徴量の尺度が同じであることが要求される．

抽出された特徴量によっては，相関関係が高いために一定の重複がみられることもある．このような場合には，特徴量を低次元の部分空間に圧縮するのに次元削減の手法が役に立つ．特徴空間の次元を減らせば，必要となる記憶域が少なくなり，機械学習アルゴリズムの処理を大幅に高速化することが可能である．

2.6　モデルの学習

モデルの汎化性能を推定するために，訓練データをさらに訓練用とテスト用に分割し，様々な交差検証の手法を適用する方法がある．交差検証の手法として，k 分割交差検証(k-Fold Cross-Validation)などがある．Fig. 6 に k 分割交差検証の概念を示した．k 分割交差検証では，非復元抽出を用いて，訓練データをランダムに k 個に分割する．そのうちの k-1 個をモデルの学習に使用し，1 個をテストに使用する．この手順を k 回繰り返すことで，k 個のモデルと性能評価を取得する．また，サブセットごとに評価された性能E_iを使って，モデルの推定平均性能 E が計算され，以下の式で表される．E_iは，正解率や誤分類率などを表す．

$$E = \frac{1}{k}\sum_{i=1}^{k} E_i \qquad (5)$$

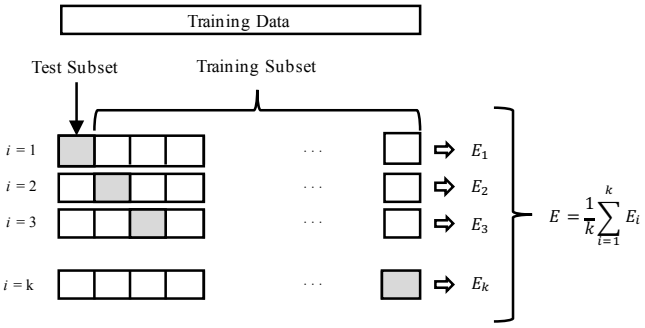


Fig. 6 Concept of k-fold cross-validation.

2.7　モデルの評価

分類モデルの性能指標として，正解率(accuracy)がある．正解率は，モデルの全体的な性能を数値化するのに便利な指標である．この他にも，適合率(precision)，再現率(recall)，F1 スコア(F1-score)などの性能指標により，モデルの妥当性を数値化することができる．

線形モデルの性能を数値化する手法の 1 つとして，平均二乗誤差(Mean Squared Error：MSE)がある．平均二乗

誤差は,線形回帰モデルを学習させるために最小化した,誤差平方和の平均値であり,様々な回帰モデルの比較や,グリッドサーチと交差検定を通じたパラメータのチューニングに役立つ.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y^{i} - \hat{y}^{(i)}\right)^2 \quad (6)$$

また，標準化された平均二乗誤差として，モデルの性能をより効果的に解釈することができる決定係数(R^2)を求めることもできる.

$$R^2 = 1 - \frac{SSE}{SST} \quad (7)$$

ここで，SSE は誤差平方和であり，SST(Sum of Squared Total)は以下の式によって表される.

$$SST = \sum_{i=1}^{n}\left(y^{i} - \mu_y\right)^2 \quad (8)$$

R^2の範囲は 0～1 をとる．$R^2 = 1$の場合，$MSE = 0$となり回帰モデルは完全に適合することを意味する.

3. 揚水発電所週間運用への機械学習の適用

3.1 訓練データ

Fig. 7 に訓練データの概要を示した．一般送配電事業者の運用データは入手できないため，ファジィ推論を適用することで作成した.

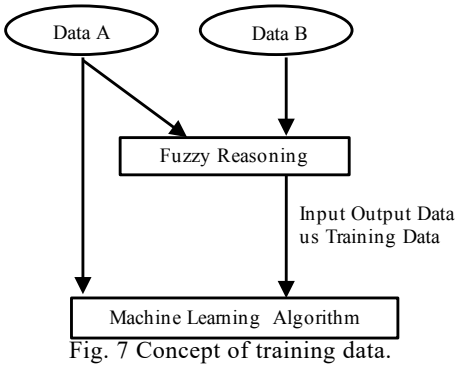


Fig. 7 Concept of training data.

3.2 検証方法

検証方法を Fig. 8 に示した．本研究では，機械学習の予測モデルが，ファジィ推論の週間運用モデルにどれくらい近似するか検証した.

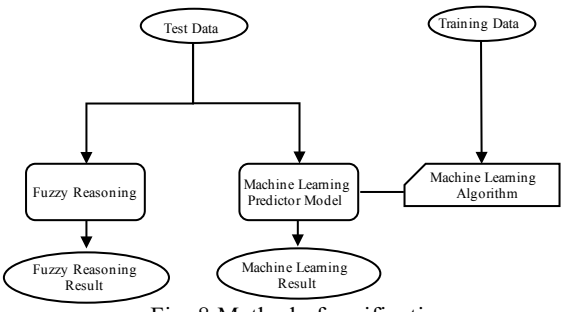


Fig. 8 Method of verification.

3.3 モデルの構成

予測モデルの構成を Fig. 9 に示した．予測モデルは，SVM を用いて，発電運転，揚水運転，待機状態の 3 クラスに分類し，その後クラスごとに線形回帰を行い，出力を予測するものを構成した.

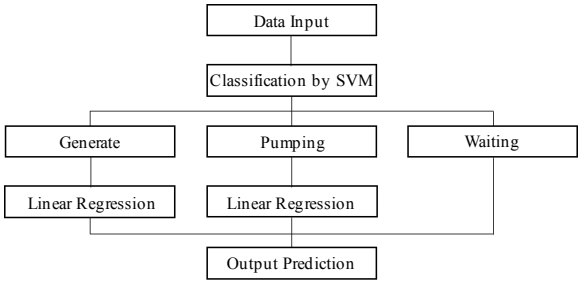


Fig. 9 Structure of prediction model.

3.4 開発環境の準備

機械学習を行うにあたり開発環境は，プログラミング言語は Python を使用した. 本研究では機械学習のプログラミングに使用される scikit-learn ライブラリを使用した.

4. 揚水発電所の週間運用シミュレーション

4.1 シミュレーション条件

本研究では，機械学習アルゴリズムに学習させる訓練データは，電力需要，太陽光発電出力，揚水発電所の上池貯水量，揚水発電出力である．需要は，2014 年東北電力株式会社の需要実績データを 1 年分使用した [7]．揚水発電所も東北電力㈱の揚水発電所である第二沼沢発電所（定格出力 460.0[MW]）を対象とした.

太陽光発電の定格出力[MW]は,東北電力㈱が試算した太陽光接続可能量である 5，050[MW]を想定した [8]．しかし，需要実績データのように太陽光発電の出力値の実測データはないため，2014 年東海大学 18 号館屋上太陽光発電の出力データを取得し，使用することとした．18 号館屋上太陽光発電の定格出力は 5.5[kW]であるため，太陽光発電出力の時系列データを，定格出力で除した発電比率として設定した．同時刻の東北電力管内での太陽光発電は，定格出力を接続可能量とみなし，この比率で発電するものとみなして使用した.

揚水発電出力，上池貯水量は実測データがないため，ファジィ推論モデルを適用し，需要の高低，太陽光出力の高低に基づき，揚水発電出力を作成した [4]．上池貯水量は，シミュレーション中の揚水発電出力によって変化する．発電では貯水量は減少し、揚水では増加する.

テストデータは取得できた，2017 年 7 月 10 日の午前 8 時から 7 月 17 日の午前 7 時までの需要及び太陽光出力データを使用した．Table 1 に訓練データとテストデータの種別を示した．また，Fig. 10 に需要と太陽光出力を示した．横軸は， 2017 年 7 月 10 日の午前 8 時から 7 月 17 日の午前 7 時までの，時刻を示している．平日の日中帯

の需要が高く，休日需要が全体的に低いという週間パターンである．

ファジィ推論モデル，機械学習モデルともに，初期水位をパラメータとして設定する必要がある．本研究では，テストデータの初期水位をそれぞれ，20 [%]，50 [%]，80 [%]とした際の，貯水量の変化を確認した．

Table 1 Data type of training data and test data.

Data Type	Demand	PV
Training Data	Tohoku Area, 2014	Tokai University Building No.18 Rooftop, 2014
Test Data	Tohoku Area, 2017	Tokai University Building No.18 Rooftop, 2014

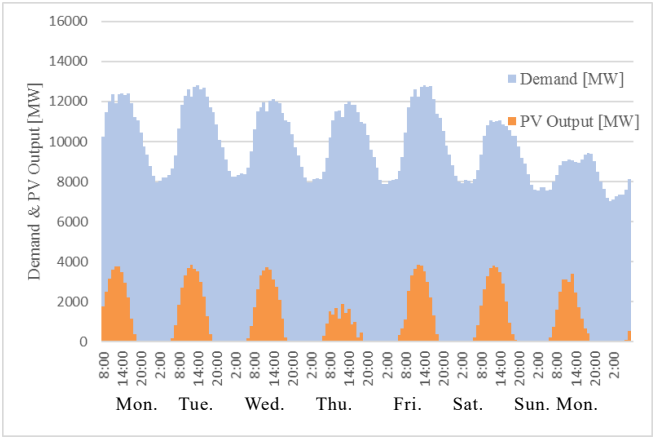


Fig. 10 Weekly demand & photovoltaic power generation on July 7, 2017.

4.2　機械学習によるシミュレーション結果

上池の初期水位を 50 [%]とした際の揚水発電所出力と貯水量の変化を Fig. 11 に示した．揚水発電所出力（左側縦軸）の 0[MW]を境に，上側（出力:正）で発電運転，下側（出力:負）で揚水運転である．

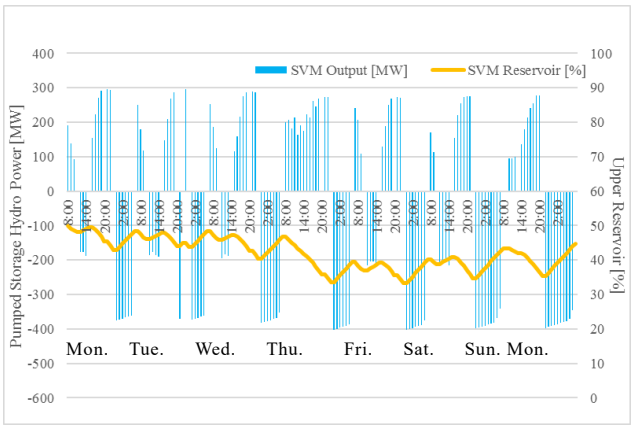


Fig. 11 Pumped storage hydro power operation on July 10, 2017 (upper reservoir 50 [%]).

シミュレーションでは，木曜が太陽光の出力が低いため，積極的に発電を行っている．木曜日以降は，水位が低くなったものの，翌月曜には初期水位の 50 [%]近くまで水位が上がっている．

4.3　ファジィ推論モデルとの比較(初期水位 50 [%])

機械学習モデルとファジィ推論モデルの出力と上池貯水量を併せて Fig. 12 に示した．平日の運転は，機械学習モデルが積極的に揚水を行う傾向にあったものの，太陽光出力が小さかった木曜に積極的に発電を行い，週末にはほぼ同じ水位となった．

日曜にファジィ推論モデルでは昼間に発電していないのに対し，機械学習モデルでは発電を行っている．この結果，翌月曜日における水位に差が生じたものの，水位を保つために揚水するファジィ推論モデルと同様の傾向が確認できた．休日に積極的な発電を行った理由として，休日ではあるものの，夏季のため年間を通して需要が高く，水位もそれほど低くないことから，SVM による分類を行った際，発電に分類されたと考えられる．

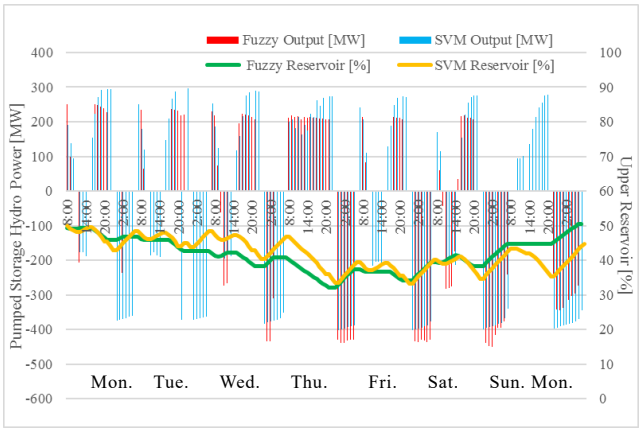


Fig. 12 Pumped storage hydro power operation on July 10, 2017 (upper reservoir 50 [%]).

4.4　ファジィ推論モデルとの比較(初期水位 20 [%])

初期水位を 20 [%]とした際の出力結果を Fig. 13 として示した．ファジィ推論モデルは木曜日まで水位があまり変化しなかったのに対し，機械学習モデルは木曜日まで揚水を行っていた．木曜日以降は水位や出力に差はあるものの，ファジィ推論モデルと機械学習モデルが，ほぼ同じタイミングで発電と揚水を行っていた．翌月曜日ではほぼ同じ水位になった．

4.5　ファジィ推論モデルとの比較(初期水位 80 [%])

初期水位を 80 [%]とした際の出力結果を Fig. 14 として示した．木曜日までは，ファジィ推論モデルと機械学習モデルともに，積極的に発電を行いながら水位は低下している．機械学習モデルの特徴として，待機状態が少ないことが確認された．前述したように SVM の分類はマージンによって分類境界を設定している．このマージンに

よって，揚水または発電に分類されやすくなったと考えられる．しかし，総じて水位はファジィ推論モデルとほぼ同じ値に落ち着いていることから，SVM のマージンの設定は適切であると判断できる．

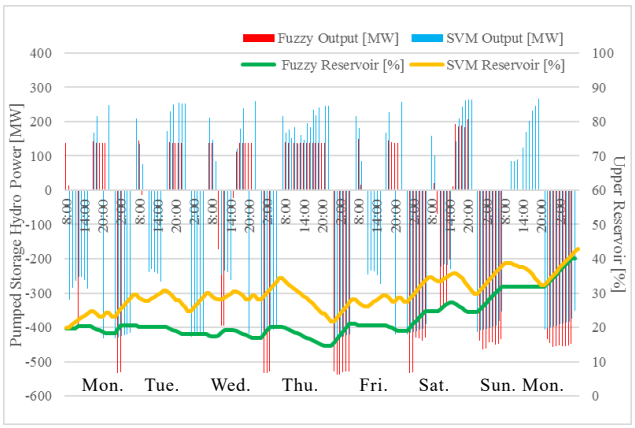


Fig. 13 Pumped storage hydro power operation on July 10, 2017 (upper reservoir 20 [%]).

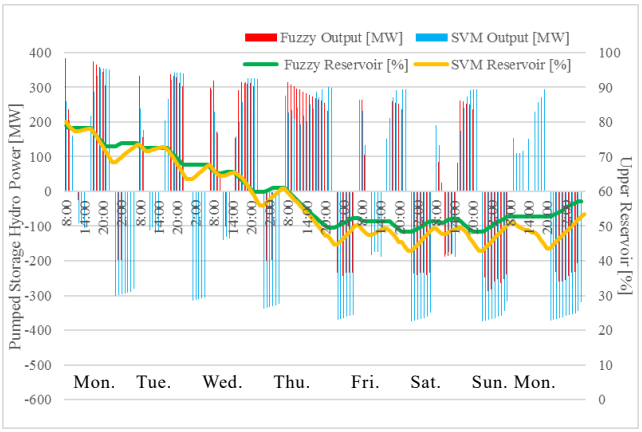


Fig. 14 Pumped storage hydro power operation on July 10, 2017 (upper reservoir 80 [%]).

5. まとめ

本研究では，揚水発電所の週間運用について機械学習を適用することを提案し，その性能を検証した．初期水位が 50 [%]，80 [%]の際は，概ねファジィ推論モデルと一致した．20 [%]の場合は水位，出力に差がでたものの，最終的な水位はほぼ同じ値となった．したがって，機械学習による本モデルの性能は適切であると判断できる．

本研究で提案した機械学習では，SVM のパラメータ調整や線形回帰の係数は訓練データを用いて決定しているため，訓練データの数を増やすほど，揚水発電所の運用決定の精度がさらに向上するものと考えられる．

本研究では一般送配電事業者が実際に運用している実績データを入手することができなかったため，ファジィ推論により作成した運用データを訓練データとして使用した．今後一般送配電事業者からの実績データを入手することができれば，より現実的な結果が得られると考えられる．

参考文献

1) エネルギー白書 2016，資源エネルギー庁.
2) 九州本土における再生可能エネルギーの出力制御について，経済産業省 第 17 回総合資源エネルギー調査会，2018 年 10 月 10 日，資料 2，参考資料 1.
3) Sebastian Raschka:「Python 機械学習プログラム 達人データサイエンティストによる理論と実践」インプレス (2016).
4) 石丸将愛，平子修二: ファジィ推論を適用した揚水発電所の週間運用策定手法の開発，東海大学工学部紀要，Vol.55, No.2, pp.15-22 (2015).
5) 奥村学: 言語処理のための機械学習入門，コロナ社 (2015).
6) 中井悦司: IT エンジニアのための機械学習理論入門，技術評論社 (2016).
7) 東北電力 HP:東北電力需要実績データ. http://setsuden.tohoku-epco.co.jp/download.html
8) 各社接続可能量(2015 年度算定値)の算定結果，総合資源エネルギー調査会 省エネルギー・新エネルギー分科会 新エネルギー小委員会 系統ワーキンググループ(第 7 回)配布資料 10-1.