

# 機械学習を用いた 熱負荷予測手法の開発に関する研究

渡邊 郁

## 1. はじめに

建築物の省エネ性能を向上させるためには、空調稼働時に要する冷暖房負荷の予測及び空調制御の最適化が必要であり、高い精度でそれらを実現する方法の開発が求められている。

従来、建物熱負荷は温熱環境シミュレーションソフトに建物モデルの形状や外皮性能、各種気象条件などの詳細なデータを入力し、物理現象を考慮した数値計算を行うことにより求められてきた。しかし、その特性から熱負荷が導出されるまでに時間がかかることや将来予測が難しいという課題を持っている。そこで、より簡易的な熱負荷予測の手法として、学習に用いるデータを一括で入力し、入力された各データの間にある関連性を自ら学習して予測を行うことが可能なアルゴリズムを開発することによって過去の経験データより将来予測を可能とする機械学習が注目されている。

そこで、本研究では、機械学習による熱負荷予測手法を確立するための基礎的検討として、少ないパラメータ数の場合でもその中から関連性を見出しモデルの構築が行われる機械学習を用いて、簡易的に翌日の日積算熱負荷を目的変数とした熱負荷予測モデルを、気象条件を中心とした入手可能な各種パラメータを説明変数として用いて構築し、予測精度の評価を行うことを目的とする。

予測モデルは重回帰分析とディープラーニングの2通りで構築し、モデルによる予測精度の差についても検討する。

また、日積算熱負荷を求める際に、気象条件を中心とした変数の選択が適当かどうかを確認するため、予測対象日前日の気象条件を中心に予測した場合と予測対象日当日の気象条件を中心に予測した場合との2通り算出し、説明変数による予測精度の差についても検討する。

## 2. 機械学習による建物熱負荷の予測手法

### 2.1 重回帰分析の概要

重回帰分析とは、多変量解析の一種であり、回帰分析の中でも複数の説明変数を用いて目的変数を導くものである。以下の式のように各説明変数 $x_1 \sim x_n$ と目的変数 $y$ は偏回帰係数 $a_1 \sim a_n$ 及び切片 $b$ を用いて一次式の線形関係で表される。

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n + b$$

### 2.2 ディープラーニングの概要

ディープラーニング（以下、DL）の構造を図1に示す。DLとは、脳機能を模したモデルであるニューラルネットワーク（以下、NN）を多層に接続し、与えた課題に対して学習を行わせる機械学習の手法の一つである。

NNにおいては、入力層と出力層のニューロンをたがいに接続し、入力された特徴量に各値の重要度を調節するための重みを掛け、その合計値を活性化関数へ通すことにより次の層へ渡す出力値とする。

NNを用いた機械学習において複数の中間層を持つものをDLといい、モデルからの出力値とテストデータとの間の誤差を計算しニューロンの重みを更新することにより学習を進めるという仕組みを持つ。また、中間層を増やすことにより非線形の課題や複数のパラメータを持つ課題に対しても精度の良い予測が行えるという特徴を持つ。<sup>1)</sup>

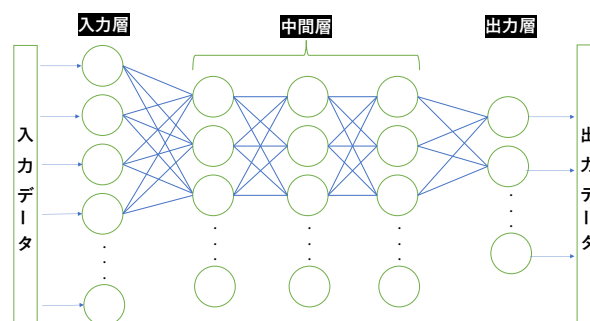


図1 DLの構造

## 3. 物理モデルを用いた建物熱負荷計算と学習データの構築

### 3.1 使用したソフト及び建物モデル

本研究では、学習データを作成するために動的熱負荷計算ソフトTHERB for HAM<sup>2)</sup>（以下、THERB）を用いた。建物モデルは、自立循環型住宅開発プロジェクトにおいて設定された標準プラン<sup>3)</sup>を用いた。表1に住宅モデルの概要を示す。

表 1 住宅モデルの概要

自立循環型住宅開発プロジェクト標準プラン	
対象敷地	福岡県福岡市
敷地面積	210.00m <sup>2</sup> (63.5坪)
建築面積	69.45m <sup>2</sup> (21.0坪)
1階床面積	62.93m <sup>2</sup> (19.0坪)
2階床面積	57.14m <sup>2</sup> (17.3坪)
合計床面積	120.07m <sup>2</sup> (36.3坪)
窓面積	27.92m <sup>2</sup>
UA値	0.51(W/m <sup>2</sup> ・K)

### 3.2 THERBによる学習データの構築

THERBによる熱負荷計算時の計算条件を表2に示す。本研究では、福岡における2015年～2020年の気象データを用い、対象建物に対して8月及び1月の熱負荷計算を行った。

各種条件として空調設定温度は8月と1月でそれぞれ2通り、空調稼働時間は7通り設定した。

また、計算対象はリビングのみとし空調もその一室のみ稼働するようにした。各種計算条件に対してTHERBによるシ

ミュレーションを行い、出力された熱負荷データを基に日積算熱負荷を算出した。

表 2 THERBIによるシミュレーション条件

項目	内容	
基本条件	気象	拡張アメダス気象データ(2015~2020年)
	計算期間	夏季: 7/31~8/31 冬季: 12/31~1/31
	計算間隔	600秒
	計算対象	リビング
	空調方式	ルームエアコン方式
各種条件	設定温度	夏季: 26°C, 28°C 冬季: 20°C, 22°C
	空調時間	24時間空調
		6:00~18:00
		18:00~6:00
		9:00~21:00
		21:00~9:00
		6:00~12:00, 18:00~24:00 9:00~15:00, 21:00~27:00

## 4. ビッグデータを用いた建物熱負荷の予測

### 4.1 熱負荷予測の計算条件および計算フロー

図2に計算条件および計算フローを示す。2015~2020年の学習データを基に、重回帰分析とDLそれぞれにおいて説明変数A及び説明変数Bを用いた日積算熱負荷の予測モデルを構築した。その際、日積算熱負荷と各説明変数間の関連性を調べるため相関係数も確認した。その後、2020年のテストデータを用いて重回帰分析とDLの予測結果の精度を比較した。

なお、重回帰分析においては、設定温度・空調時間ごとの式(a)、空調時間ごとの式(b)、全パターンを統合した式(c)の3つの予測式を導出し、各式による予測精度の差を比較した。

本研究においては、気象データを中心として予測を行った場合の精度を比較することを目的とするため、変数選択方法として強制投入法を用いた。翌日の説明変数として最高気温、最低気温、降雨の有無と共に浦野・他(2017)の研究より高い精度で予測可能であることが明らかにされている翌日の自然室温も採用した。

また、天気などの元々数値ではない質的変数を説明変数として用いる場合には0または1で表されるダミー変数に変換する必要があるため、本研究では降雨があった場合を1、それ以外を0とするダミー変数に変換した。

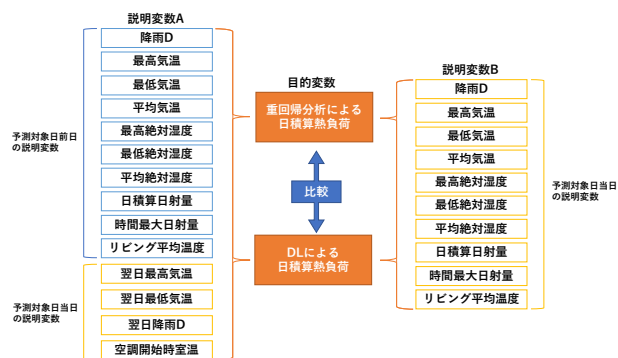


図 2 計算条件および計算フロー

## 4.2 モデルの概要

### 4.2.1 重回帰分析モデルの概要

2015~2019年の学習データを基に算出した予測式に関する重回帰分析結果のうち、一部結果を抜粋し夏季の説明変数Aを用いた場合における空調時間が(i)6:00~18:00、(ii)18:00~6:00、(iii)6:00~12:00, 18:00~24:00の3パターンの結果を以下の表3に示す。なお、式(a)の空調設定温度は28°Cの場合とする。

予測値と実測値との相関を表す重相関係数、説明変数が目的変数をどの程度説明できているかを表す重決定係数、及び自由度調整済重決定係数は、式(a)において概ね0.9を上回っており、建物熱負荷予測において気象条件を中心とした説明変数は有効であるということが確認できた。しかし、パターンを統合した式(b)、式(c)において各値は式(a)と比較して低下していることが確認できた。

また、予測式の係数を示す偏回帰係数(表中:偏)はパターンごとに異なっており、標準化偏回帰係数(表中:標)を基に算出される目的変数に対する各説明変数の寄与率(表中:寄)もパターンごとに異なる値を示した。

以上より、予測の際にはパターンごとの予測式を用いること、パターンを統合する際には条件の変化を考慮した説明変数を随時追加していく必要があると考えられる。

表 3 夏季のパターンごとの重回帰分析結果(説明変数A)

空調条件	空調時間(i)-式(a)		空調時間(ii)-式(b)		空調時間(iii)-式(a)		空調時間(ii)-式(b)		
重相関係数	0.9747		0.9434		0.9771		0.9608		
重決定係数	0.9500		0.8999		0.9546		0.9231		
自由度調整済決定係数	0.9450		0.8874		0.9501		0.9213		
説明変数	(偏)	(標)	(偏)	(標)	(偏)	(標)	(偏)	(標)	
切片	-96843.9	-0.980	-5477.7	-0.130	-50851.4	-0.562	-1894.1	-0.325	
降雨D	173.4	0.005	0.1%	356.1	0.010	0.2%	-196.1	-0.006	0.4%
翌日降雨D	202.3	0.006	0.1%	-304.3	-0.009	0.2%	1018.9	0.030	2.0%
最高気温	223.0	0.173	4.2%	23.5	0.018	0.3%	-104.8	-0.081	5.4%
最低気温	378.3	0.283	6.8%	722.0	0.540	10.3%	96.1	0.072	4.7%
平均気温	-1461.9	-1.092	26.2%	-739.8	-0.553	10.6%	35.5	0.027	1.7%
最高絶対湿度	149.6	0.077	1.9%	575.4	0.298	5.7%	79.7	0.041	2.7%
最低絶対湿度	301.9	0.136	3.3%	181.8	0.082	1.6%	14.6	0.007	0.4%
平均絶対湿度	-403.6	-0.195	4.7%	-919.7	-0.444	8.5%	-92.2	-0.045	2.9%
日射量合計	0.0	0.005	0.1%	-0.3	-0.043	0.8%	-0.1	-0.014	0.9%
日射量1時間最大値	0.5	0.010	0.2%	2.0	0.037	0.7%	0.5	0.010	0.7%
リビング平均温度	2029.0	0.662	15.9%	-3019.2	-0.986	18.9%	454.6	0.148	9.8%
翌日最高気温	488.0	0.378	9.1%	439.3	0.340	6.5%	226.6	0.176	11.5%
翌日最低気温	543.6	0.404	9.7%	475.3	0.354	6.8%	526.9	0.392	25.8%
翌日空調開始時室温	1346.4	0.737	17.7%	2746.1	1.503	28.8%	864.6	0.473	31.1%
							1644.0	0.900	25.1%

空調条件	空調時間(iii)-式(a)		空調時間(ii)-式(b)		統合式(c)	
重相関係数	0.9767		0.9245		0.8996	
重決定係数	0.9540		0.8547		0.8093	
自由度調整済決定係数	0.9494		0.8513		0.8087	
説明変数	(偏)	(標)	(寄)	(偏)	(標)	(寄)
切片	-112533.1	-1.064		-8380.6	-0.162	
降雨D	91.7	0.003	0.1%	600.4	0.018	0.3%
翌日降雨D	201.2	0.006	0.2%	-333.4	-0.010	0.2%
最高気温	5.7	0.004	0.1%	173.4	0.135	2.2%
最低気温	233.4	0.175	4.6%	931.9	0.697	11.6%
平均気温	-1034.1	-0.773	20.5%	-737.3	-0.551	9.2%
最高絶対湿度	138.3	0.072	1.9%	658.4	0.341	5.7%
最低絶対湿度	174.5	0.079	2.1%	277.1	0.125	2.1%
平均絶対湿度	-294.3	-0.142	3.8%	-1313.0	-0.634	10.6%
日射量合計	0.0	0.003	0.1%	0.0	0.002	0.0%
日射量1時間最大値	0.3	0.006	0.2%	1.1	0.019	0.3%
リビング平均温度	2351.4	0.768	20.3%	-3295.5	-1.076	17.9%
翌日最高気温	599.8	0.465	12.3%	490.6	0.380	6.3%
翌日最低気温	674.9	0.502	13.3%	684.6	0.509	8.5%
翌日空調開始時室温	1423.0	0.779	20.6%	2755.8	1.508	25.1%
				1965.5	0.108	2.9%

### 4.2.2 ディープラーニングモデルの概要

学習に用いたモデルの構成を図3に示す。中間層が直列に接続されたSequentialモデルであり、中間層はすべて全結合ニューラルネットワークであるDenceとして配置している。また、勾配消失を抑制するためにBatchNormalizationを中

間層の出力に対して適応している。入力層のパラメータ数は可変として説明変数に様々な値を用いることができるようにしている。<sup>1)</sup>

なお、DLの学習においては一般的に数千〜数万のデータが必要とされるが、用意したデータ数が1つのモデルに対して2170であり、不足している可能性があることを踏まえうえて様々なパターンを検討した結果エポック数を1000、学習率を0.001、バッチサイズは16とした。

夏季に説明変数Aを用いて構築したモデルのエポック数と損失関数の関係を図4に示す。なお、損失関数は平均二乗誤差(MSE)を用いている。学習が進むにつれ学習データに対する予測値の誤差とテストデータに対する誤差はともに低下していく傾向がみられることからわかるように、モデルが学習データに対して過度に適合してしまいテストデータに対する精度が落ちてしまう過学習の傾向はみられない。

しかし、局所的に評価データに対する誤差が大きくなっていることから、学習不足が起きていると考えられる。学習不足が起きる原因としては、データ数が少ないことによりモデルが各変数間に存在する関係性を十分に学習しきれていないことや、有効な説明変数の個数が少ないことにより熱負荷予測時に考慮すべき現象の複雑さに対してモデルが単純になり学習時の取りこぼしが多くなることがあげられる。

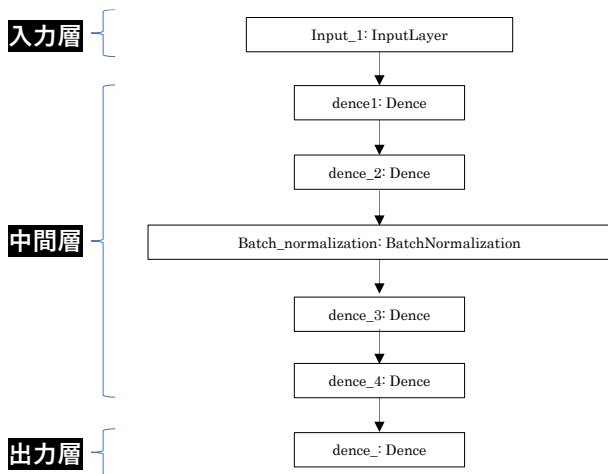


図 3 学習に用いたモデルの構成

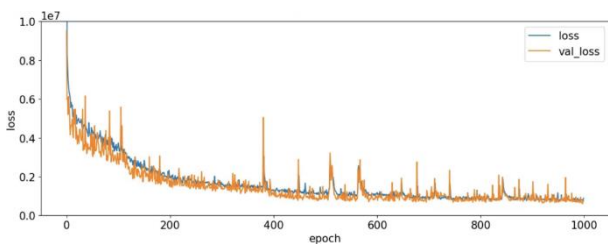


図 4 夏季のDLモデルによるエポック数と損失関数の関係(説明変数A)

#### 4.2.3 各変数間の相関係数

熱負荷予測に用いた各変数間の夏季における相関係数を表4に示す。夏季・冬季ともに熱負荷と説明変数間には概ね相関関係がみられるが、0.7を超える強い相関を持つ説明変数は少ない。また、各説明変数間において気温関連、湿度関連及び日射関連の説明変数に強い相関関係がみられた。

予測モデルを構築する際に重要な要素の一つに、目的変数に対して有効な説明変数を組み合わせることがあげられる。

そのため、熱負荷と強い相関を持つ説明変数の選択や、説明変数間に相関を持つものを統合するなどの手法で、説明変数の個数や種類を調整することにより精度を上げることができると考えられる。

また、冬季においても各説明変数の間に同様の傾向がみられた。

表 4 夏季の各変数間の相関係数

説明変数	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1.熱負荷	1.00														
2.降雨D	-0.43	1.00													
3.翌日降雨D	-0.42	0.35	1.00												
4.最高気温	0.67	-0.64	-0.45	1.00											
5.最低気温	0.68	-0.45	-0.39	0.82	1.00										
6.平均気温	0.73	-0.57	-0.43	0.95	0.93	1.00									
7.最高絶対湿度	0.37	-0.10	-0.30	0.51	0.64	0.57	1.00								
8.最低絶対湿度	0.33	-0.09	-0.27	0.39	0.59	0.48	0.77	1.00							
9.平均絶対湿度	0.36	-0.12	-0.28	0.46	0.64	0.54	0.92	0.92	1.00						
10.日射量合計	0.48	-0.78	-0.44	0.67	0.47	0.61	0.14	0.11	0.14	1.00					
11.日射量1時間最大値	0.46	-0.80	-0.47	0.70	0.47	0.60	0.16	0.13	0.15	0.94	1.00				
12.リビング平均温度	0.04	-0.28	-0.22	0.46	0.46	0.49	0.25	0.18	0.21	0.30	0.30	1.00			
13.翌日最高気温	0.75	-0.49	-0.62	0.70	0.68	0.74	0.39	0.36	0.38	0.55	0.53	0.37	1.00		
14.翌日最低気温	0.77	-0.43	-0.43	0.74	0.76	0.81	0.47	0.45	0.48	0.49	0.46	0.40	0.82	1.00	
15.翌日空調開始時自然室温	0.19	-0.30	-0.28	0.46	0.45	0.49	0.24	0.20	0.22	0.33	0.31	0.81	0.45	0.46	1.00

相関係数 r  
 0.7 ≦ |r| < 1 強い相関がある  
 0.4 ≦ |r| ≦ 0.7 相関がある  
 0.2 ≦ |r| ≦ 0.4 弱い相関がある  
 0 ≦ |r| ≦ 0.2 ほとんど相関がない

#### 4.3 熱負荷の予測結果

##### 4.3.1 重回帰分析とDLによる予測結果比較

重回帰分析とDLによる予測結果を表5、表6に示す。予測精度の評価法としては平均絶対パーセント誤差(MAPE)を用いた。

初めに、重回帰分析における各予測式の予測精度を比較すると式(a)、式(b)、式(c)の順で予測精度が落ちる傾向がみられた。これは、パターンを統合していく際に条件の変化を考慮した説明変数を追加していないため予測モデルとしての説明力が低下していることが原因であると考えられる。

また、重回帰式とDLの予測精度では、全体的に重回帰式のほうが高い傾向がみられた。これは、DLのモデルに対しての学習データ数が不足していることから各変数間に存在する関係性を学習しきれておらず、予測精度を十分に高めることができているためと考えられる。

次に、空調条件ごとに比較すると、冬季では設定温度が低いほうが、そして稼働時間帯が夜間より日中であるほうが精度が悪くなっている。これは、住宅モデルの断熱性能が高く室温が上昇しやすくなっており、室温が設定温度より高くなった場合に空調が自動停止してしまう影響を予測に反映できていないことが原因であると考えられる。

また、夏季・冬季ともに二度間欠空調を設けている空調時間6:00~12:00,18:00~24:00の場合でも予測精度が低下する傾向はみられなかった。

表 5 説明変数Aによるパターンごとの予測精度の比較(夏)

空調条件		重回帰分析			DL
空調時間	設定温度	(a)式誤差	(b)式誤差	(c)式誤差	DL誤差
24時間空調	26°C	14.8%	15.6%	19.4%	16.9%
	28°C	17.7%	17.5%	19.5%	20.5%
6:00~18:00	26°C	7.7%	14.9%	20.1%	22.6%
	28°C	9.6%	14.0%	19.6%	22.0%
18:00~6:00	26°C	3.9%	6.0%	18.0%	24.0%
	28°C	4.1%	6.1%	18.4%	26.2%
9:00~21:00	26°C	11.0%	14.5%	19.0%	21.8%
	28°C	12.8%	13.8%	18.6%	19.1%
21:00~9:00	26°C	4.3%	8.1%	19.4%	25.2%
	28°C	4.3%	8.0%	20.0%	27.7%
6:00~12:00,18:00~24:00	26°C	7.1%	19.9%	23.1%	22.9%
	28°C	8.6%	17.8%	22.6%	26.8%
9:00~15:00,21:00~27:00	26°C	6.3%	16.8%	21.9%	23.1%
	28°C	8.2%	15.7%	22.4%	26.6%
各式の誤差平均値		8.6%	13.5%	20.1%	23.3%

表 6 説明変数Aによるパターンごとの予測精度の比較(冬)

空調条件		重回帰分析			DL
空調時間	設定温度	(a)式誤差	(b)式誤差	(c)式誤差	DL誤差
24時間空調	20°C	19.7%	28.3%	24.2%	48.2%
	22°C	17.5%	17.5%	15.2%	29.3%
6:00~18:00	20°C	17.4%	29.0%	33.2%	60.4%
	22°C	15.2%	15.0%	19.1%	36.7%
18:00~6:00	20°C	12.6%	19.4%	41.4%	78.9%
	22°C	5.8%	7.5%	21.7%	41.5%
9:00~21:00	20°C	30.1%	47.8%	49.8%	85.6%
	22°C	27.2%	24.7%	28.6%	47.1%
21:00~9:00	20°C	7.7%	14.6%	30.1%	56.2%
	22°C	5.2%	7.7%	17.8%	34.3%
6:00~12:00,18:00~24:00	20°C	19.0%	31.9%	30.0%	57.7%
	22°C	16.0%	17.4%	16.7%	33.6%
9:00~15:00,21:00~27:00	20°C	23.4%	39.5%	37.4%	65.8%
	22°C	17.5%	19.4%	19.3%	35.5%
各式の誤差平均値		16.7%	22.8%	27.4%	50.7%

### 4.3.2 説明変数ごとによる予測結果比較

説明変数Aと説明変数Bによる予測結果を表7、表8に示す。説明変数ごとに比較すると、僅かに説明変数Bのほうが高い傾向がみられるものの大きな差はなかった。これは、4.2.1で示したように、設定温度・空調条件ごとの式(a)においては高い決定係数が出ていることから、予測対象日の前日・当日どちらの気象条件を説明変数としても予測精度への影響は少ないためと考えられる。

## 5. まとめ

本研究では、機械学習を用いて、簡易的に翌日の日積算熱負荷を目的変数とした熱負荷予測モデルを、気象条件を中心とした入手可能な各種パラメータを説明変数として用いることで構築し、予測精度の評価を行った。また、モデルの構築を重回帰分析とDLでそれぞれ行ったことや、構築時に予測対象日の前日・当日のパラメータを用いたことによる予測精度の差に関しても検討した。以下に得られた知見を示す。

(1) 少ないデータ数で同一の説明変数を用いて予測した場合、重回帰分析のほうがDLよりも精度が高くなる。

(2) 説明変数として予測対象日の各値を用いた場合と前日の各値を用いた場合では予測精度に大きな違いは見られない。

(3) 重回帰分析結果より、パターンを統合していくにつれてモデルが線形性を失うことにより予測精度が低下する傾向がみられる。

今後の課題として、より多くの学習データを用いてモデルの構築を行うこと、空調条件の変化や変数同士の関連性を考慮して説明変数の個数や種類を調整すること、時間間隔を短くした詳細な予測をすることなどがあげられる。

表 7 説明変数Aと説明変数Bによる予測精度の比較(夏)

空調条件		重回帰式(c)誤差		DL誤差	
空調時間	設定温度	説明変数A	説明変数B	説明変数A	説明変数B
24時間空調	26°C	19.4%	19.9%	16.9%	15.2%
	28°C	19.5%	18.3%	20.5%	16.9%
6:00~18:00	26°C	20.1%	20.2%	22.6%	21.3%
	28°C	19.6%	17.0%	22.0%	22.7%
18:00~6:00	26°C	18.0%	21.8%	24.0%	23.6%
	28°C	18.4%	20.1%	26.2%	23.3%
9:00~21:00	26°C	19.0%	20.1%	21.8%	19.9%
	28°C	18.6%	16.9%	19.1%	21.6%
21:00~9:00	26°C	19.4%	22.4%	25.2%	25.1%
	28°C	20.0%	20.2%	27.7%	21.2%
6:00~12:00,18:00~24:00	26°C	23.1%	22.2%	22.9%	21.9%
	28°C	22.6%	19.3%	26.8%	22.8%
9:00~15:00,21:00~27:00	26°C	21.9%	21.9%	23.1%	20.8%
	28°C	22.4%	20.3%	26.6%	23.0%
各式の誤差平均値		20.1%	20.0%	23.3%	21.4%

表 8 説明変数Aと説明変数Bによる予測精度の比較(冬)

空調条件		重回帰式(c)誤差		DL誤差	
空調時間	設定温度	説明変数A	説明変数B	説明変数A	説明変数B
24時間空調	20°C	24.2%	21.0%	48.2%	45.7%
	22°C	15.2%	12.7%	29.3%	28.6%
6:00~18:00	20°C	33.2%	33.8%	60.4%	57.0%
	22°C	19.1%	20.1%	36.7%	35.9%
18:00~6:00	20°C	41.4%	40.0%	78.9%	72.8%
	22°C	21.7%	20.4%	41.5%	39.9%
9:00~21:00	20°C	49.8%	42.0%	85.6%	78.2%
	22°C	28.6%	21.8%	47.1%	44.1%
21:00~9:00	20°C	30.1%	28.7%	56.2%	54.0%
	22°C	17.8%	17.1%	34.3%	31.7%
6:00~12:00,18:00~24:00	20°C	30.0%	28.2%	57.7%	55.7%
	22°C	16.7%	15.9%	33.6%	32.6%
9:00~15:00,21:00~27:00	20°C	37.4%	32.6%	65.8%	60.8%
	22°C	19.3%	17.4%	35.5%	34.6%
各式の誤差平均値		27.4%	25.1%	50.7%	48.0%

### 参考文献

- 1) 谷川悠:ディーブラーニングを用いた熱負荷予測アルゴリズムの開発に関する研究, 2021.2
- 2) Ozaki A., Kagawa H.: Simulation Software to Describe theHygrothermal Environment of Whole Buildings Based on detailedPhysical Models, Proc. of the 7th International Conference of SystemSimulation in Buildings, pp03 (24 pages), 2006
- 3) 国土技術政策総合研究所: 自立循環型住宅設計技術資料 - 評価法および要素技術の設計手法に関する研究成果 -, 2012.7
- 4) 浦野明, 田中拓也: ニューラルネットワークを用いた室内環境予測, 空調調和・衛生工学会論文集, pp97~100, 2017