

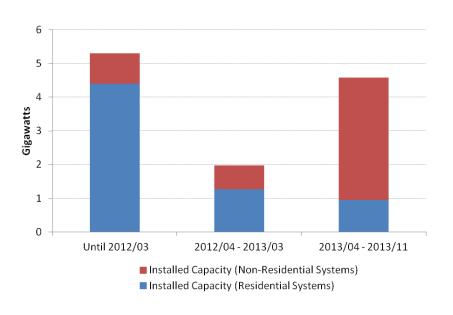
サポートベクター回帰と数値天気予報を利用した 太陽光発電予測研究について

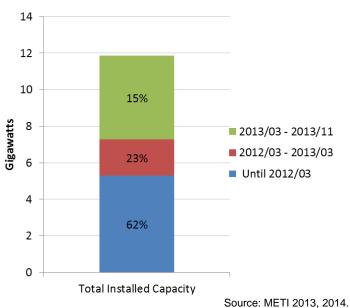
ガリ ダ シルバ フォンセカ ジュニア ジョアン (産業技術総合研究所) 大関 崇(産業技術総合研究所) 荻本 和彦(東京大学)



Introduction

近年、新しい再生可能エネルギー固定価格買取制度の効果で太陽光発電設備の導入量 を伸び続けている。





- 然し太陽光発電量は気象状況による激しく変動するのでそのシステムが高普及率を達し たら電力系統運用に問題を引き起こす恐れがある。
- それで太陽光発電の普及率の増加と共にそのシステムの発電量を高精度で予測出来る 技術を開発する必要性も高まる。



太陽光発電量予測

予測分類

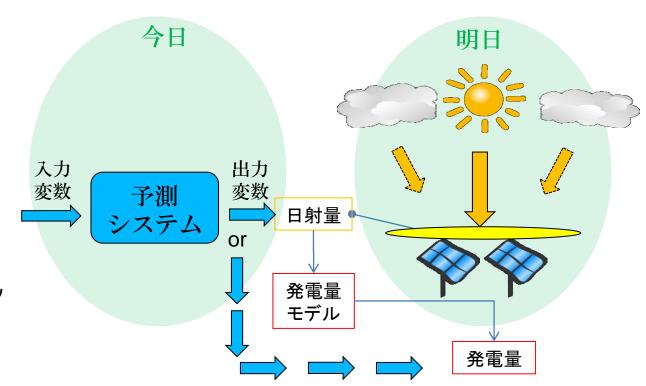
- 間接法
- 直接法

利用データ分類

GPV-MSM (気象庁)

手法

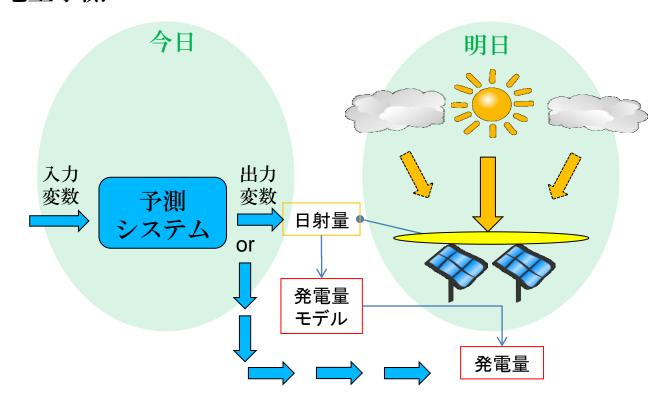
人工知能(ニューラルネットワーク、SVM等)



目的: 事前に(1日前)1日の太陽光発電システムの発電量を予測することです。



太陽光発電量予測



最初に (2010年)

予測分類

間接法でGPV-MSMを利用しながら日射量の予測に集中した。

手法

人工知能(基本的なニューラルネットワークを作った)



入力データ

気象庁のGPV-MSM予報システム

気温

湿度

雲量(下層、中層、上層)



気象庁の地上気象観測再統計値 気象庁のメソ数値予報モデルGPV。

■メソ数値予報モデルGPV(MSM)

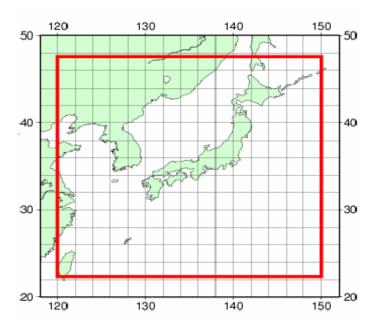
平成18年3月から、計算の格子間隔を5kmメッシュに細分化するとともに、運用回数を1日8回としました。

◇データの概要◇

▽ / 一×∪/成委	· · ·
内容	要素は下記を参照してください。
初期値	00, 03, 06, 09, 12, 15, 18, 21UTC(1日8回)
予報時間	15時間予報(00,06,12,18,UTC)、33時間予報(03,09,15,21UTC) 地上はは時間間隔、気圧面は3時間間隔
データ形式	国際気象通報式FM92 GRIB 二進形式格子点資料気象通報式(第2版)※略称:GRIB2
配信領域	北緯22.4度~47.6度、東経120度~150度(※下図参照)
格子系	等緯度等径度 地上:0.05度×0.0625度(格子数505×481) 気圧面:0.1度×0.125度(格子数253×241)
サイズ	15時間予報:約114MB/回×4回、33時間予報:237MB/回×4回=1404MB/日
技術資料	配信資料に関する技術情報第205号 国際気象通報式・別冊

◇地上物理量

·	1/0-7-						
	海面更正気圧	地上気圧	風	気温	相対温度	時間降水量	雲量
地上	0	0	2	$\overline{}$	0	0	(4)



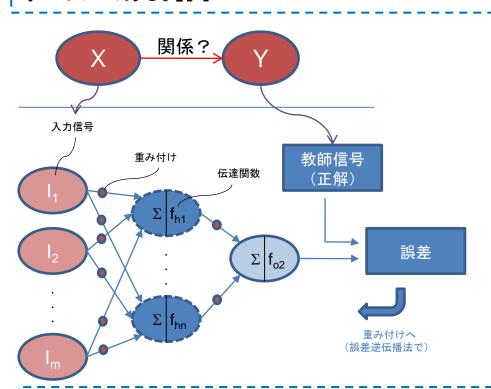
ソース:JMBSCのホームページ



予測手法 1

最初の予測手法 — ニューラルネットワーク(ANN)

「脳機能に見られるいくつかの特性を計算機上のシミュレーションによって表現することを目指した数学モデルである。」[1]



利点[2]

- ▶ 線型、非線型問題でも対応出来る。
- ▶ 一般化能力
- ▶ 並列処理能力

欠点

- ▶ 極小値に収束しやすい
- ▶ 大規模な対象に対する精密な計画・設計や制御[3]

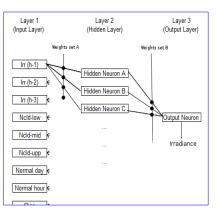
「結果として次元削減されるため、画像や統計など多次元量のデータでかつ線形分離不可能な問題に対して、比較的小さい計算量で良好な解を得られることが多い。」[1]

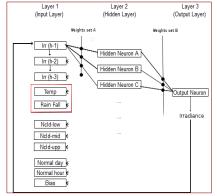


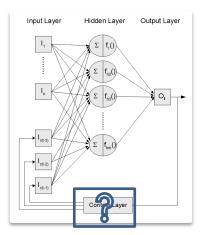
予測手法 1

最初の予測手法 — ニューラルネットワーク(ANN)

自分のANNを作り出来る限り様々な観点から設定入力データなどの適切な設定を検討した。





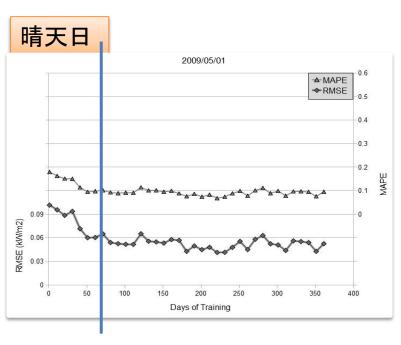


- ① 一年に一つANN 対 一日に一つANN。
- ② Context layerを使うか使わないか。
- ③ 適切な入力データの検討。
- ④ 適切な活性化関数の検討。
- ⑤ 適切な訓練事例量の検討。
- ⑥ 訓練の機関の検討など行った。



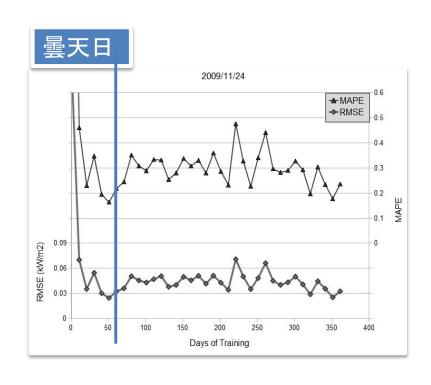
例:訓練事例量の検討

Fonseca et al, 電気学会部門大会B2011



訓練事例量による誤差変動

	51日	61日	181日	361日
MAPE	9.51%	9.87%	7.66%	9.58%
RMSE	0.060	0.060	0.042	0.052



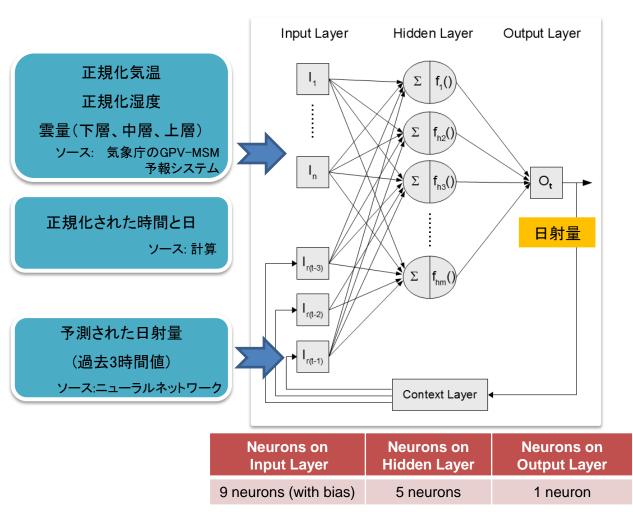
訓練事例量による誤差変動

	51日	61日	181日	361日
MAPE	16.2%	21.7%	27.9%	23.5%
RMSE	0.024	0.032	0.041	0.032



ANNの設定

ニューラルネットワークの設定



収束アルゴリズム

▶誤差逆伝播

活性化関数

>0.52*tanh(X) - 0.51

他の設定パラメーター

▶学習係数: 0.3

▶慣性項: 0.2

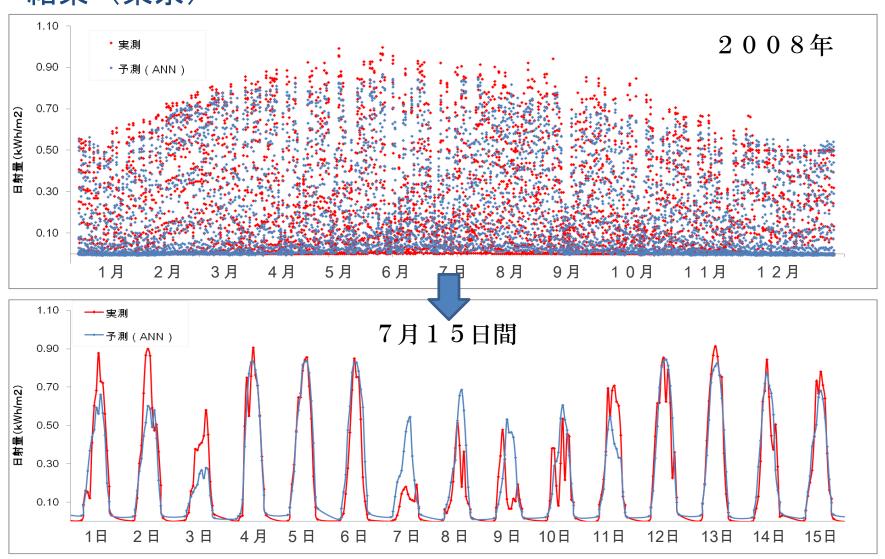
➤訓練:2000 epochs



日射量予測

最初の結果

結果 (東京)





日射量予測 結果の比較

最初の結果

最適な結果を果たす為実際の天 気状態を用いた。

ニューラルネットワーク(ANN) 対 回帰式と天気予報用プログラム

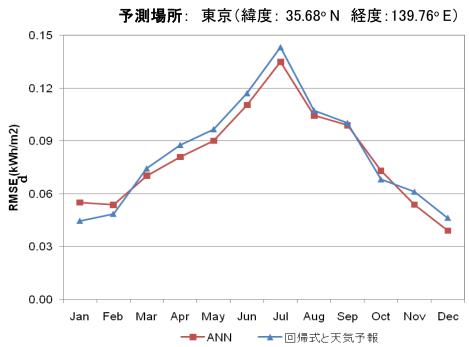


図8 予測日射量:二つ手法の比較 二乗平均平方根誤差。

予測日数:360日(2008年01月~2008年12月)

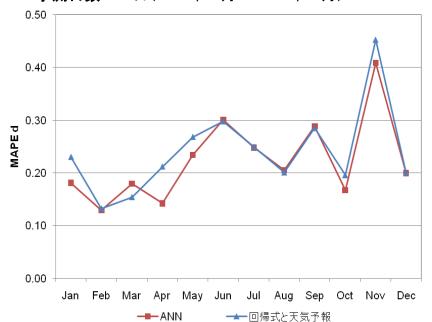


図9 予測日射量:二つ手法の比較 平均是対誤差率。

表4 予測日射量:二つ手法の比較 1年間の誤差。

	誤差値 (1年間)	R	SME (kWh/m²)	MAE (kWh/m²)		MAE _d (kWh/m²)	MAPE _d
	ANN	噩	0.0936	0.0575	噩	0.5318	22.29%
-	回帰式と天気予報	业	0.0963	0.0577 0.35%	Ш	0.6142	23.91% 7.3%



別手法の検討

サポートベクターマシン

「サポートベクターマシン は、教師あり学習を用いる識別手法の一つである。パターン認識や回帰分析へ適用できる。現在知られている多くの手法の中で一番認識性能が優れた学習モデルの一つである」。

ニューラルネットワーク

反復法

学習際最適化問題を解決しない発 見的な、ヒューリスティックな サポートベクターマシン

直接法

学習際最適化問題を解決する



別手法の検討

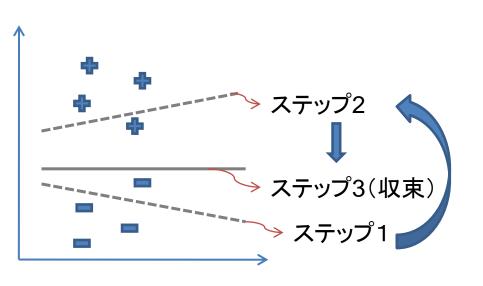
ニューラルネットワーク

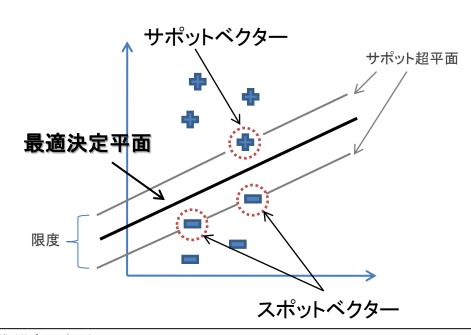
反復法

学習際最適化問題を解決しない発 見的な、ヒューリスティックな サポートベクターマシン

直接法

学習際最適化問題を解決する







サポートベクターマシン

サポートベクターマシンについて (v-SVR)

$$y = f(x) = (w.x) + b$$
 線形表現

同じ問題超平面で表現すると

マップ関数

$$y_i \approx \sum_{j=1}^m w_j \emptyset(x_{ij}) + b$$
 wはラグランジュの乗数で表現できる

$$w = \sum_{j=1}^{n} (\alpha_i - \alpha_i^*) \emptyset(x_t)$$

$$y_i \approx \sum_{j=1}^m (\alpha_i - \alpha_i^*) k(x_i, x_j) + b$$

 $y_i pprox \sum\limits_{j=1}^m (\alpha_i - \alpha_i^*) k(x_i, x_j) + b$ マップ関数カーネル関数*で表現できる *マッピングを定義する為二変数関数。

問題はラグランジュ関数に変換し最適化問題に成り。αを解決する為次の問題を最大にする。

$$-\frac{1}{2} \sum_{i,j=1}^{n} (\alpha_i - \alpha_i^*) k(x_i, x_j) (\alpha_j - \alpha_j^*) - \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) y_i$$

制約条件

$$\sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0$$

$$0 \le \alpha_i^* \le \frac{c}{n}$$

$$\sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \le C.\nu$$

設定パラメーター

カーネル関数

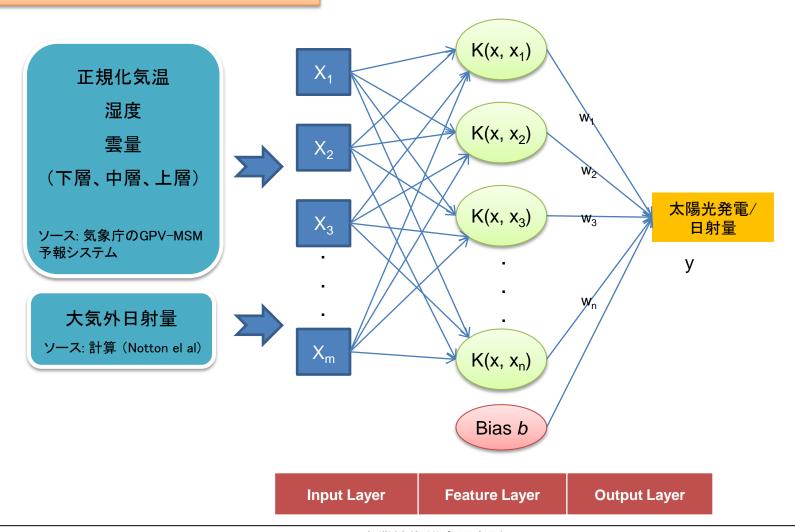
様々形で決定できる

$$k(x_i, x_j) = e^{-\gamma |x_i - x_j|^2}$$



サポートベクターマシン

サポートベクターマシンの設定





手法の比較

1

日射量予測の例 (日本国内の25箇所)

予測手法の比較

ANN

SVM

2

発電量予測の例 (響灘太陽光発電所)

誤差逆伝播法

多層ネットワーク

専用活性化関数

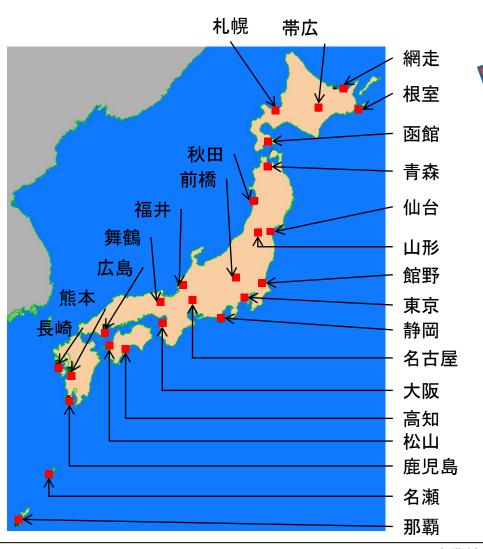
学習データサイズ60日間

予測手法による予測精度が どのように変動するか。





Insolation Forecast for 25 Locations in Japan



予測の条件						
箇所の数	25					
予測期間	1 年間 (2008年) *時間毎 5時から19迄					
学習に利用された データ	60 日間					

^{*}一日の予測は過去の60日間のデータを利用した。

誤差

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (I_{fcs,i} - I_{msd,i})^{2}}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |I_{fcs,i} - I_{msd,i}|$$

/ _{fcs, i} : 予測日射量(kWh/m²) ; /_{msd, i} : 実測日射量(kWh/m²) ;

N: 時間の数

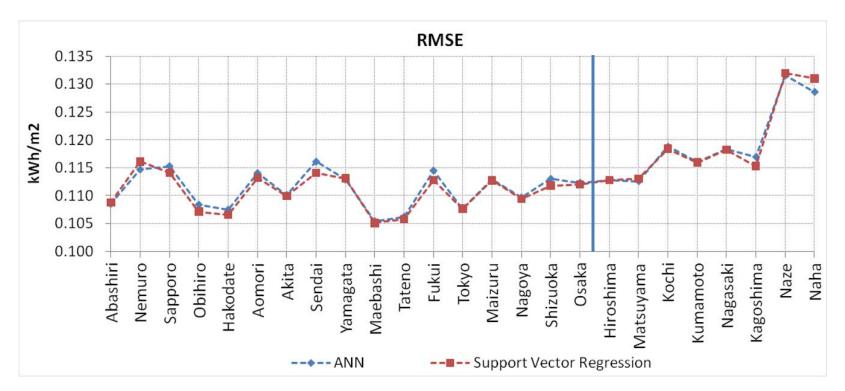




Results



25箇所の一年間のRMSE



Fonseca et al, SWC 2011

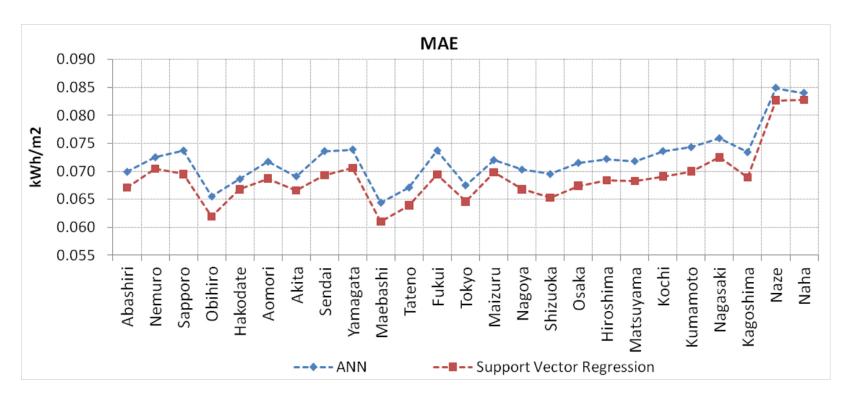




Results



25箇所の一年間の絶対誤差



Fonseca et al, SWC 2011





PV Power Forecast (Hibikinada PV Power Plant)

発電所について

データ	值
緯度	33° 55' N
経度	130° 44' E
最大出力	1 MW(nominal)
太陽光モジュル	5600*
傾斜	20°

^{*}多結晶シリコン

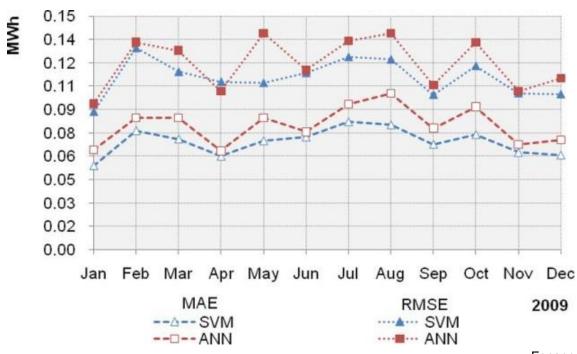


予測条件について

データ	値				
予測期間	1 年 (2009年)				
予測種類	一時間毎(6時から19時迄)				
訓練事例量	60 日				
予測時間	前日の18時				

ソース JPOWER平成19年度 共同研究業務成果報告書

Results



Fonseca et al, IEEE PVSC 2011

	MAE(MWh)	RMSE(MWh)
ANNの予測	0.080	0.1206
SVMの予測	0.069	0.1109
変動(ANNから)	-13%	-8%



Forecast Accuracy

日射量予測の例 (日本国内の25箇所)

	最低 値	最大値	日射計 の精度
RMSE(kWh/m²)	0.105	0.135	
			± 0.03
MAE (kWh/m²)	0.060	0.085	kWh/m ²

発電量予測の例 (響灘発電所)

	最低値	最大出力
RMSE(MWh)	0.116	
		1 MWh
MAE (MWh)	0.069	



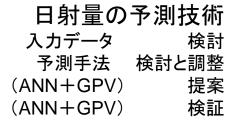
Improving the Forecast Method

予測手法の改善と実用化を向け次の話題を研究した。

- ◆ 学習データ量 (例 15日間から365日間迄)
- ◆ 学習データの種類(例 直前のデータ対類似があるデータ)
- ◆ SVMのカーネル関数の種類 (例 3つ種類)
- ◆ 手法の初期設定 (例 交差検定対単純なグリード検索等)
- ◆ 予測時間の影響 (例 2時間前迄 対 25時間前迄)
- ◆ 入力変数の効果 (例 雲量等使うかどうか)
- ◆ 1日に1SVM 対 1時間に1SVM
- **♦**...



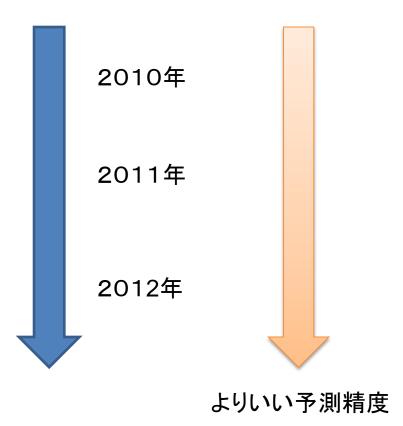
これまでの纏め(前半)



予測手法の改善 ANN→SVR 日射量予測→直接に発電量予測 検討

SVRがガイダンス手法として 検討 分類機対回帰

日射量予測→直接に発電量予測 検証



これまで局地の予測に注目 ⇒ これから広域も

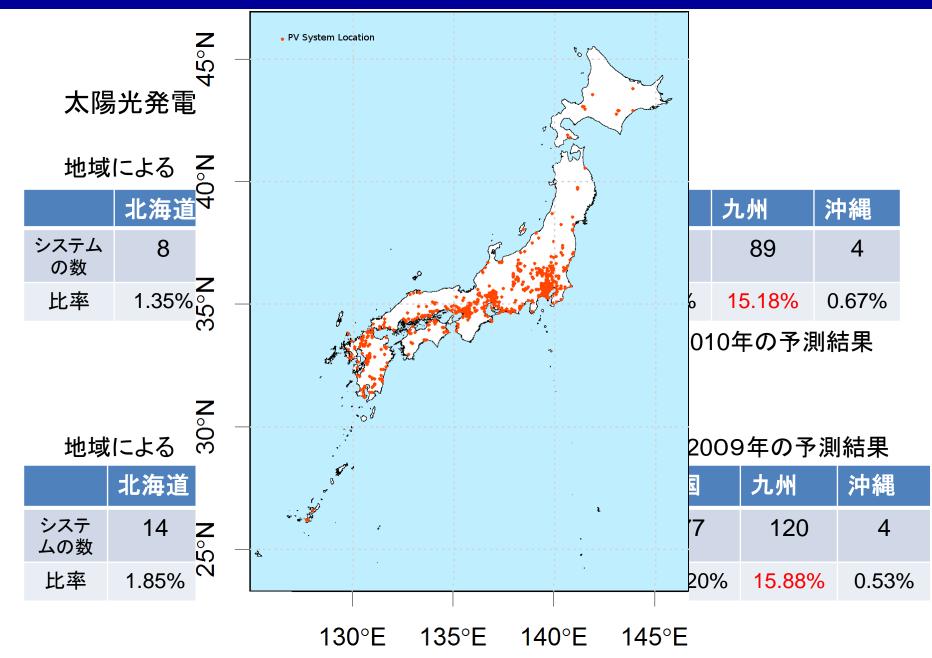


- ▶ 全国の太陽光発電予測(FT-Data) 予測手法の検証
- > 予測値の信頼空間の検討
- ▶ 広域太陽光発電予測

広域予測手法の比較

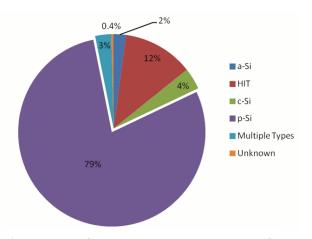
Up-scaling (広域の予測が局地の予測の標本調査で)



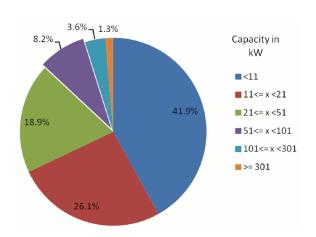




Output Data Description

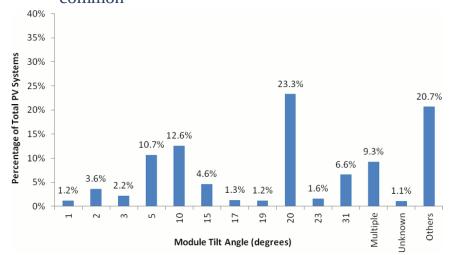


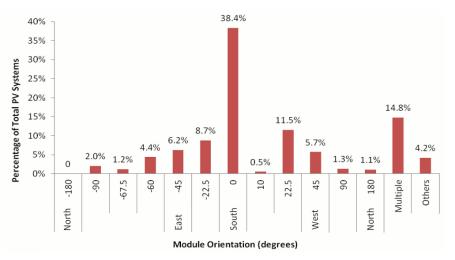
➤ Most of the PV system were of the polycrystalline silicon type.



➤ PV system with less than 10 kW of capacity were common (although there were examples with more than 5 MW).

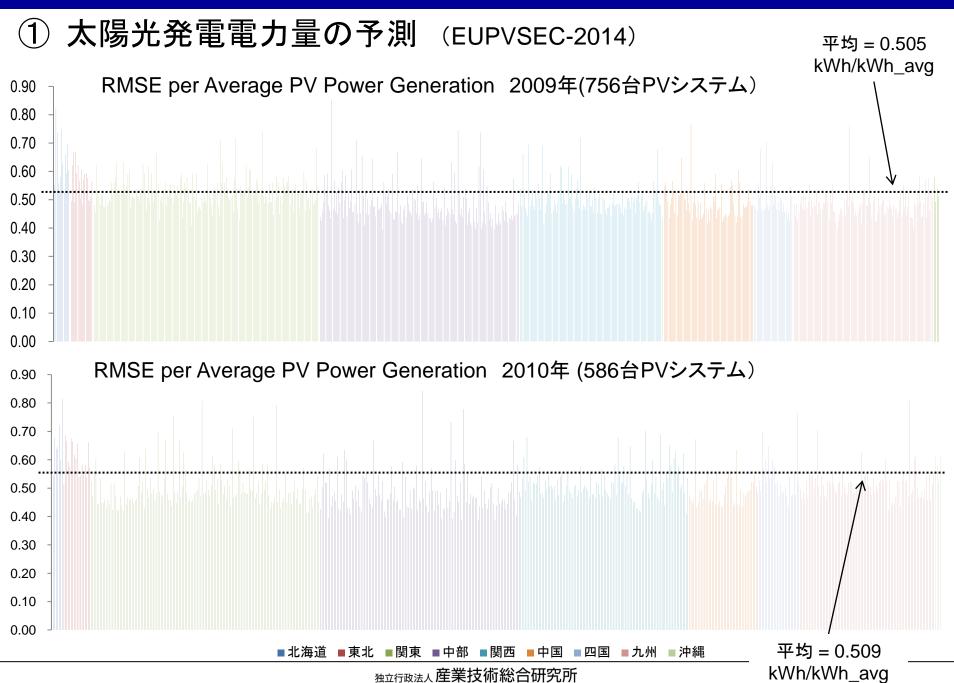
➤ Modules with tilt angles of 10° and 20° were common





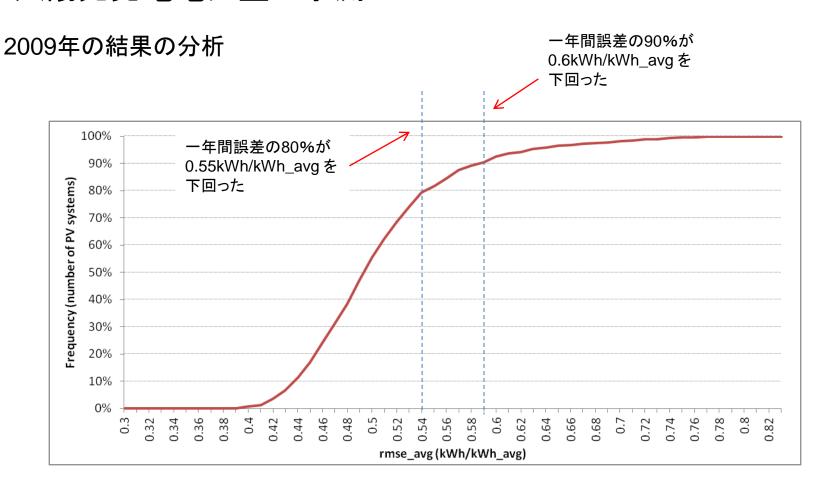
➤ Near 60% of Modules were facing South or near to South directions







① 太陽光発電電力量の予測





SVMの設定について (IEEJ大会2014)

グリッド検索

柼

アンサンブルドの基にしたアプローチ

それぞれ設定パラメータはどこからどこまで変動することを決定(範囲)

決定した範囲から最初の値を選択し SVMを設定する。

一年間時間毎に発電量予測を行う。 一日の予測モデルは直前の60日間で学習する (毎日違う予測モデルを利用する)



予測がお終わったら一年のRMSEを計算してその値を保存する。



設定パラメータの値の組み合わせが終わる迄その値を変動してSVMを設定しもう一回予測を行う。

最低RMSEを生じさせた設定変数値で SVMを設定をする。 決定した範囲から最初の値を選択し SVMを設定する。

直接にその設定で一日に各時間の予測をする

(毎日違う予測モデルを利用する) (一年間の予備データを利用しない)

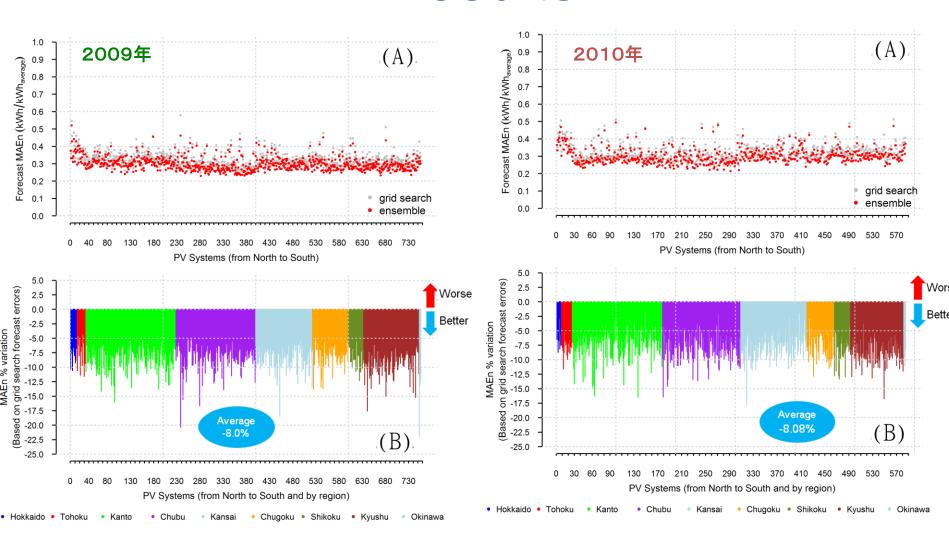
予測値を保存する。

設定パラメータの値の組み合わせが終わる 迄その値を変動してSVMを設定しもう一回予 測を行う。

各時間によるした予測値の分布から中央値を計算してその値は予測値として出力する。



Results

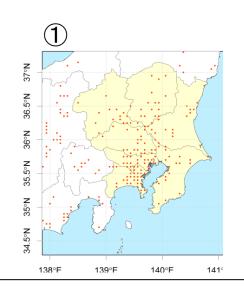


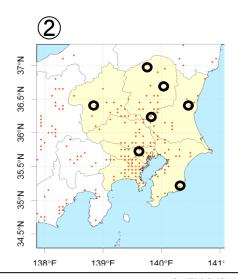


局地予測 ⇒ 広域予測(4つ手法)

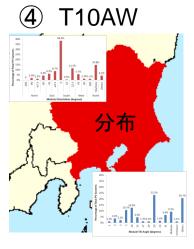
必要な情報

手法			入力データ						
	各 位置	各 容量	各過去 発電量	広域 容量	広域 発電量	方位角 分布	傾斜角 分布	各GPV データ	各大気外 日射量
1- 全局地予測から	100%	100%	100%	Χ	X	X	Χ	0	0
2- 標本調査から	40台	40台	40台	X	X	X	X	0	0
3- 広域発電量から	100%	Χ	Χ	0	0	Χ	X	0	0
4- 日射量予測から	100%	X	X	X	X	0	0	0	0







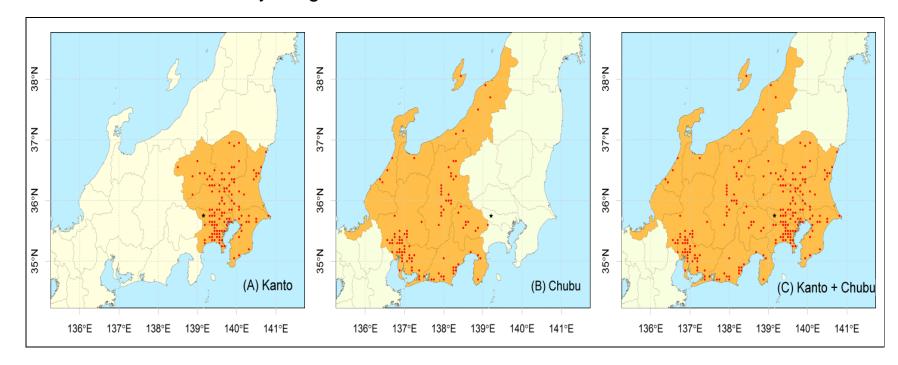


独立行政法人產業技術総合研究所



広域手法の比較(4つ手法)

これまで予測対象は関東(2009年)だったが今回中部と関東と中部で広域の予測を行った (Wiley Progress in Photovoltaic 2014)。

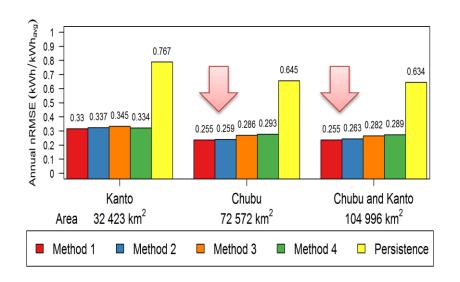


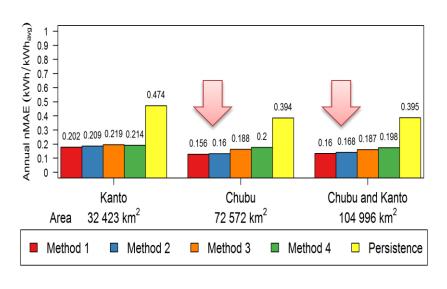
Area	32 423.90 km ²	72 572. 34 km ²	104 996.24 km ²
Installed Capacity	4 409.02 kW	3 694.16 kW	8 103.18 kW
Number of PV systems	143	130	273



広域手法の比較(4つ手法)

結果



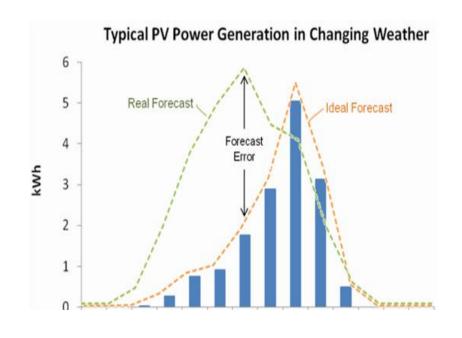


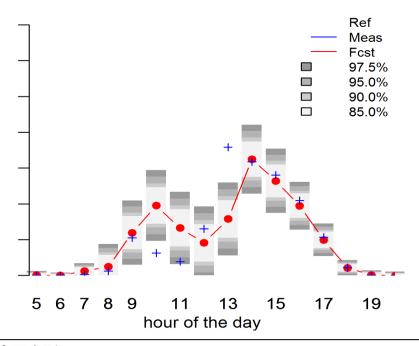
	Kanto	Chubu	Kanto and Chubu
Method with lowest nMAE	Method 1	Method 1	Method 1
Method with highest nMAE	Method 3	Method 4	Method 4
Maximum % Variation*	6%	28%	23.7%

- RMSEとMAEでみると中部や中部と関東の広域予測に手法による無視できない予測 精度の差であった。
- その場合,全局地予測と標本調査の予測手法が予測誤差が最低となった。



- ▶ 本研究で前日太陽光発電予測技術の開発で、様々な検討しながら局地の問題に限らずに広域の予測も対応出来た。
- 国内の予測誤差も特定しさらに徐徐に予測精度も改善してきた。
- ≫ 然しいくら翌日予測精度が改善しても大外れ発生の問題が残る。

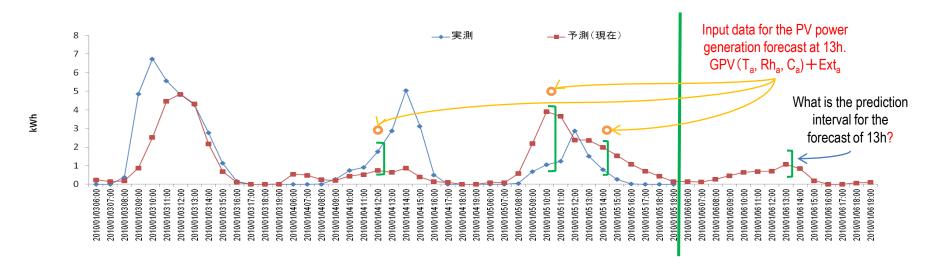






Calculation of the Prediction Intervals (Grand Renewable 2014)

The prediction interval for a target hour are calculated considering the forecast errors that occurred given expected weather conditions, extraterrestrial insolation, season and time of day (morning of afternoon) similar to the target hour's ones.



- ① For a forecast value target, its input data are compared with the input data used in past forecasts (past 60 days, 840 hours).
- 2 The most similar hours (5% quantile, 42 hours) to the target one are selected and gathered in a set.
- ③ It is assumed that the samples in the set are part of a Laplace distribution or a Gaussian distribution.
- 4 If any of the 2 assumptions is true, σ for these distributions can be theoretically calculated from the Maximum Likelihood Estimation Method.



Calculation of the Prediction Intervals

Laplace Distribution

$$p(z) = \frac{1}{2\sigma} e^{-\frac{|z|}{\sigma}}$$
 (A)

If the set of forecast errors, *err*, follows this distribution σ in Eq. (A) can be calculated via maximum likelihood estimation as:

$$\sigma = \frac{1}{l} \sum_{i=1}^{l} |err_i|$$

Once σ is known Eq. (A) can be integrated and the superior and inferior limits, p_i and p_s , of the predictive intervals can be calculated as:

$$p_i = \sigma \ln(2s)$$
 $p_s = -\sigma \ln(2s)$

Gaussian Distribution

$$p(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{|z^2|}{2\sigma}}$$
 (B)

If the set of forecast errors, *err*, follows this distribution σ in Eq. (B) can be calculated via maximum likelihood estimation as:

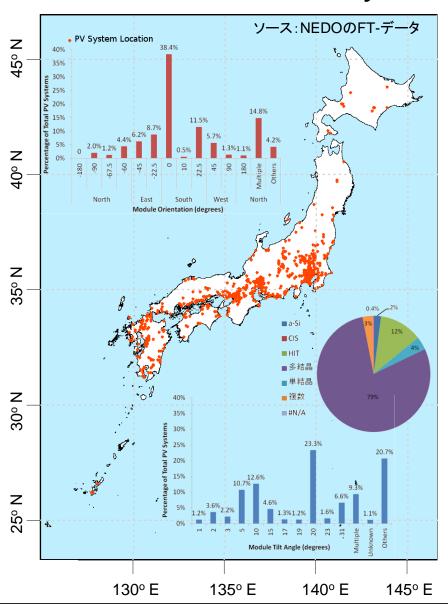
$$\sigma^2 = \frac{1}{l} \sum_{i=1}^{l} |err_i|^2$$

Once σ is known Eq. (A) can be integrated and the superior and inferior limits, p_i and p_s , of the predictive intervals can be calculated as:

$$p_i = -\sigma \Phi^{-1}(1-s)$$
 $p_s = \sigma \Phi^{-1}(1-s)$



PV Systems Description



- ➤ One year of data of 421 PV systems were used (a subset of the systems in the map).
- ➤ These PV systems are in installed in different locations in Japan.
- ➤ They are part of the Field Test Project maintained by NEDO.
- ➤ The set contains a variety of examples regarding installation conditions and rated power.

Table 1. Installed capacity of PV systems per region.

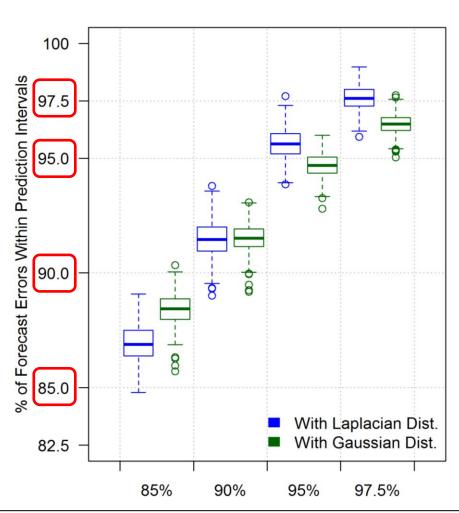
idate it installed superenty of the systems per regions					
	Number of PV systems	Installed Capacity (kW)			
Hokkaido	6	118.9			
Tohoku	12	261.8			
Kanto	109	3 415.8			
Chubu	107	3 438.2			
Kansai	65	2 425.6			
Chugoku	38	1 432.1			
Shikoku	21	896.1			
Kyushu	61	2 198.8			
Okinawa	2	50.4			
Total	421	14 237.7			



Results

Which distribution assumption is better? Four confidence levels.

Predetermined Confidence Level vs. Achieved Annual Forecast Error Coverage



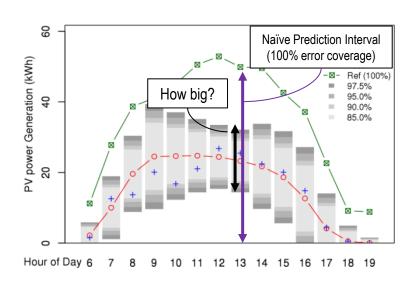
- ➤ Based on the median values in the boxplot it is clear that the assumption of Laplacian distribution to model the forecast error is a better assumption that the Gaussian distribution.
- Moreover, even in the worst forecast error coverage values the confidence level provided good estimations of it.
- ➤ With the Gaussian distribution assumption there is a trend to underestimate the forecast error coverage at low values of confidence levels and overestimate them at high values.



Results

How big are the intervals?

Size of the intervals with the proposed method versus the naïve reference method.



➤ With the Laplacian distribution assumption the proposed method yielded prediction intervals 44% smaller than the naïve method while yielding 97.6% of forecast error coverage.

		Proposed Method with Laplace Distribution			Proposed Method with Gaussian Distribution		
		Forecast errors within prediction intervals*	Prediction intervals relative** size		3	Forecast errors within prediction intervals*	Prediction intervals relative** size
Confidence Level (%)	8 5	86.9%		0.327		88.4%	0.332
	9	91.5%		0.389		91.5%	0.376
	9 5	95.6%		0.484		94.7%	0.439
	9 7. 5	97.6%		0.565		96.5%	0.491

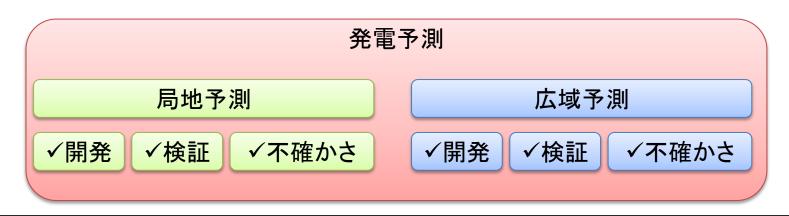
^{*} Median of the forecast error coverages for all PV systems.

^{**} The ratio is based on the average size of the intervals compared with the average size of the intervals from the naïve reference approach.



纏め

- 本研究で日射や太陽光発電の発電予測手法を開発することを目的とした。
- 最初に局地の手法の開発および検証を行い、その後、広域予測にも対応した。
- 広域予測は、どんな情報を使えるのが未定のため4つのシナリオを考慮した予測手法を提案して検証した。
- 得られる情報による発電データを良い精度で予測できることを確認した。
- しかし、結果からPVシステムの完全監視、あるいは標本調査のデータで発電予測を 行った方が良いということが分かった。





ご静聴ありがとうございます