



IDEA

INTERDISCIPLINARY DIVISION
FOR ENERGY ANALYSES

OZE – predykcja generacji ze źródeł punktowych

Konferencja upowszechniająca wyniki projektu: „Rozwój energetyki rozproszonej w klastrach energii KlastER”

(Marcin Blachnik, Sławomir Walkowiak)



Prognozy generacji OZE

- Predykcja generacji OZE
 - Wiatr
 - PV
 - Prognozy średnioterminowe (do 72 godzin do przodu)
 - Prognozy krótkoterminowe (do kilku godzin do przodu)
- Prognozy na podstawie:
 - Historycznych pomiarów generacji (SCADA, ewentualnie inne systemy np. Chronos)
 - Historycznych danych meteorologicznych z pomiarów w lokalizacji farmy OZE
 - Dane z numerycznych modeli prognostycznych (modele mezoskalowe, horyzont prognozy do 100h do przodu)
- Dlaczego?
 - Horyzont średnioterminowy - minimalizacja błędu prognozy generacji oznacza lepszą ocenę zysków ze sprzedaży energii w dniu następnym N+1
 - Horyzont krótkoterminowy - sterowanie instalacjami hybrydowymi (interakcja/synergia PV, wiatru, magazynów energii ewentualnie w przyszłości elektrolizerów, konieczna większa rozdzielczość prognoz i dobre urządzenia pomiarowe)

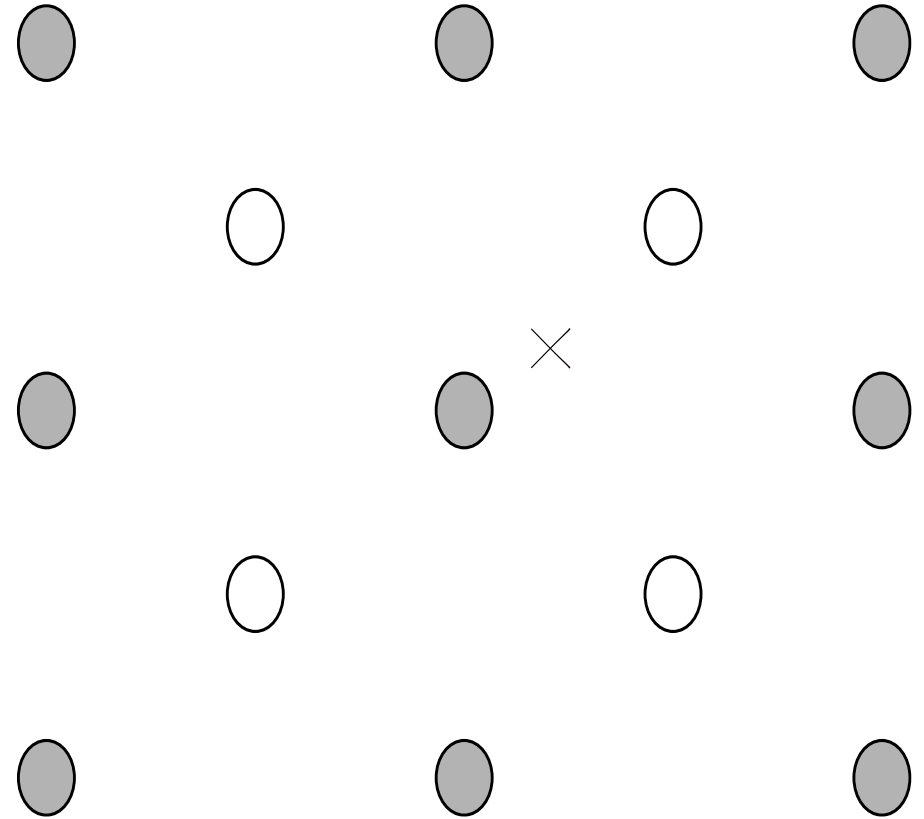
Dane meteorologiczne

- Model o siatce 4 km x 4 km (przykład dla modelu UM)
- 4 prognozy dziennie (0, 6, 12, 18 UTC)
- Prognozy 6 i 18 UTC : 60 h
- Prognozy 0 i 12 UTC: 120 h
- Jeśli wymagamy 72 godzin prognozy generacji - konieczność "sklejania prognoz meteo"

Dane meteorologiczne – wiatr, przykład

- Okres prognozowania 2019-2020
- Farma 1: ~50 MW
- Farma 2: ~90 MW
- Farma 3: ~70 MW

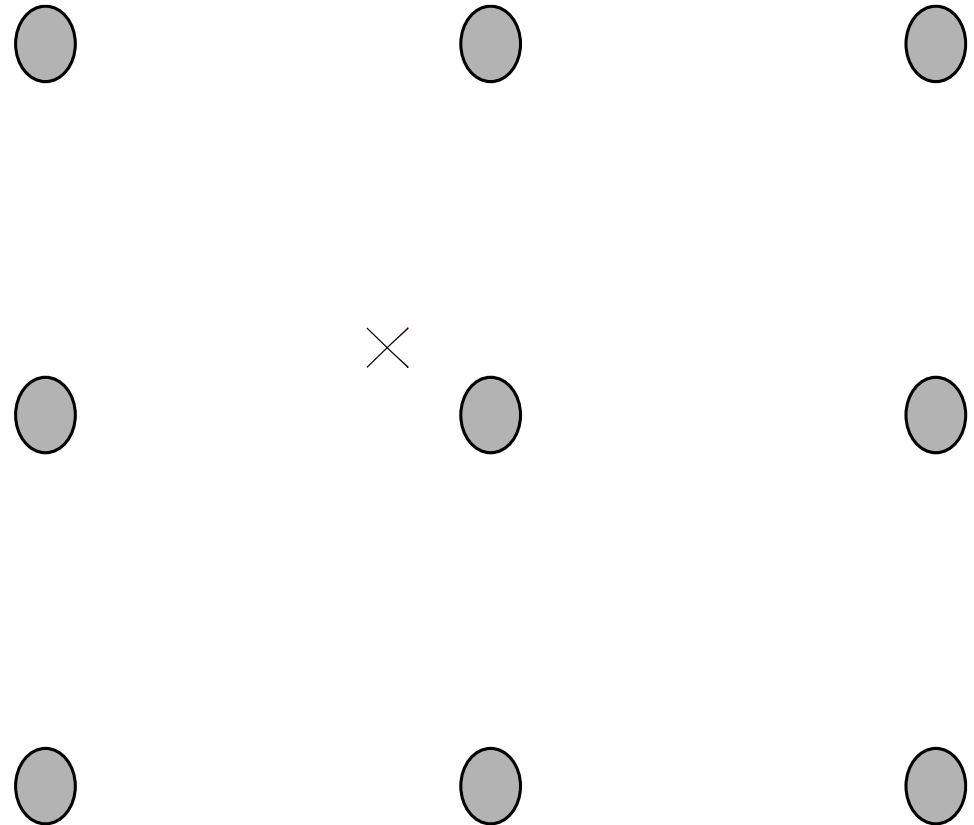
- Rozdzielczość czasowa 1 h
- 90 m składowa U wiatru
- 90 m składowa V wiatru
- Temperatura 1.5 m
- Ciśnienie 1.5 m
- Wilgotność względna 1.5



Dane meteorologiczne – PV, przykład

- Okres prognozowania 2020
- Farma 1: ~0.7 MW
- Farma 2: ~1 MW
- Farma 3: ~1 MW

- Temperatura 1.5 m [K]
 - Rozdzielczość czasowa 1 h
- Całkowity strumień promieniowania krótkofalowego (solflux) [W/m²]
 - Rozdzielczość czasowa 15 min



Kryteria optymalizacyjne

- MSE (Mean Square Error)
- MAE (Mean Absolute Error - jako miara pomocnicza)
- Możemy wyróżnić wyprzedzenia: 2 h, 8 h i 35 h
- Rozkład błędu dla każdej chwili wyprzedzenia (boxplot)
- Dla godzin wyróżnionych - mediana i wariancja błędu nie może się pogarszać wraz ze zmianą modelu

Metodologia oceny modelu



Data prognozy	...	}	Zbiór treningowy cz.1
...	...		
2020.03.12 6:00	...	}	Zbiór testowy
2020.03.12 6:00	...		
2020.03.12 12:00	...	}	Zbiór treningowy cz.2
2020.03.12 12:00	...		
2020.03.12 12:00	...	}	Zbiór testowy
2020.03.12 12:00	...		
2020.03.12 18:00	...	}	Zbiór testowy
2020.03.12 18:00	...		
...	...	}	Zbiór treningowy cz.2
2020.04.21 18:00	...		
2020.04.21 18:00	...	}	Zbiór testowy
2020.04.21 18:00	...		
2020.04.21 18:00	...	}	Zbiór treningowy cz.2
2020.04.22 0:00	...		
2020.04.22 0:00	...	}	Zbiór testowy
2020.04.22 0:00	...		
...	...	}	Zbiór treningowy cz.2
...	...		

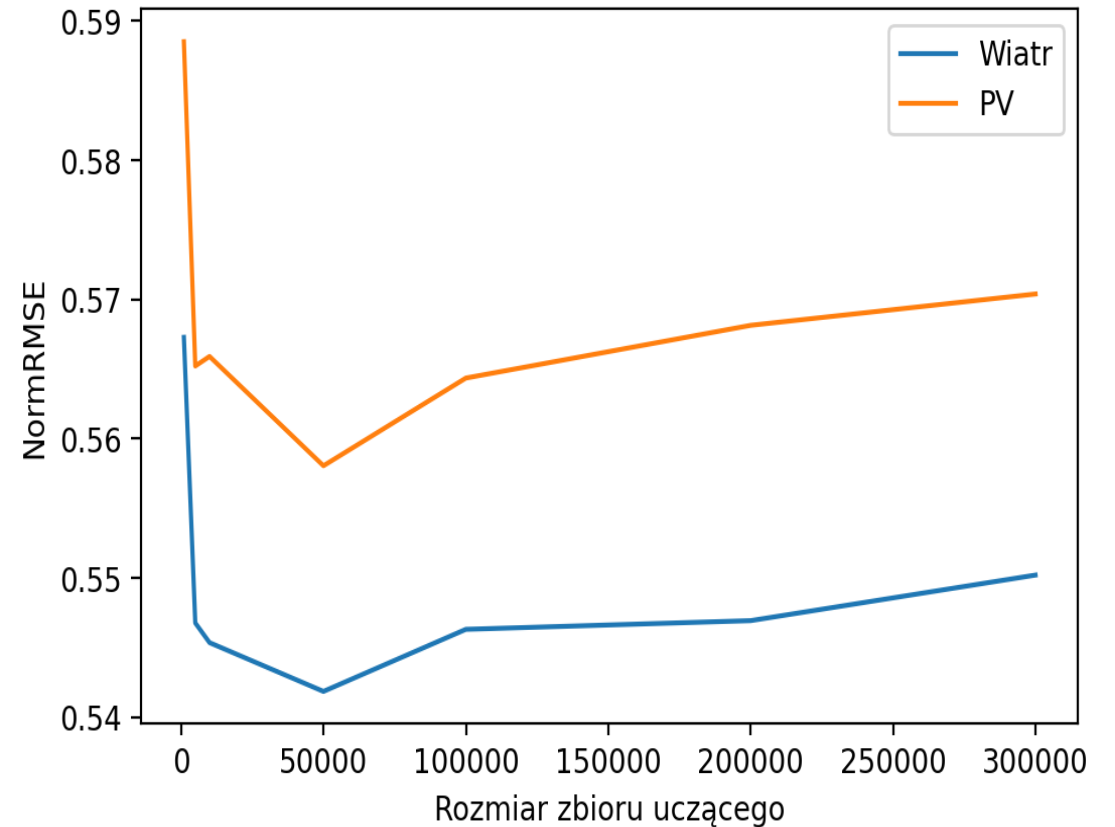
Wybór modelu

- Wybrano algorytm Lasu Losowego (Random Forest)
 - Prostota
 - Odporność na przeuczenie
 - Łatwość interpretacji
 - Wydajność (ważne w perspektywie predykcji dla obszarów lub dużej liczby farm)
- GBT (Gradient Boosted Trees) - wyniki porównywalne do lasu, znaczący spadek wydajności (wzrost czasu obliczeń)

Model	NormRMSE wiatr	NormRMSE PV
Reg. liniowa	0.595	0.581
kNN	0.582	0.578
MLP	0.561	0.586
GBT	0.542	0.578
Las losowy	0.544	0.564

Optymalizacja – krzywa uczenia

- Zależność jakości wyników od rozmiaru zbioru uczącego
- Czas uczenia
- Czas predykcji
- Różnorodność (model diversity) - spadek zdolności predykcyjnych przy zwiększaniu rozmiaru zbioru uczącego



Optymalizacja hiperparametrów - wiatr

- Najczęściej wybierany:
 - las 400 drzew,
 - głębokość 10
- Uniknięcie "wycieku informacji"
- Zagnieżdżenie doboru parametrów w nadrzędnym teście krzyżowym (cross validation)
- 5 iteracji testu krzyżowego
- Przeszukiwanie na siatce
 - 5, 10, 15: głębokość drzewa
 - 50, 100, 200, 300, 400: drzew

Farma	Fold	RMSE	MAE	R	R ²	Głęb.	drzewa
1	1	7.096	5.117	0.802	0.644	10	400
	2	7.641	5.803	0.836	0.699	15	400
	3	8.325	6.257	0.880	0.775	10	400
	4	7.916	5.627	0.763	0.582	10	300
	5	6.282	4.459	0.625	0.391	10	400
2	1	12.321	9.203	0.822	0.675	10	300
	2	12.864	9.665	0.840	0.705	10	300
	3	14.312	10.593	0.876	0.767	10	400
	4	12.016	8.218	0.831	0.690	10	300
	5	10.312	7.366	0.810	0.655	10	300
3	1	15.289	11.565	0.724	0.524	15	400
	2	15.509	11.941	0.763	0.583	10	300
	3	18.159	13.501	0.797	0.635	15	400
	4	14.681	10.975	0.736	0.542	10	400
	5	12.727	9.346	0.681	0.464	10	300

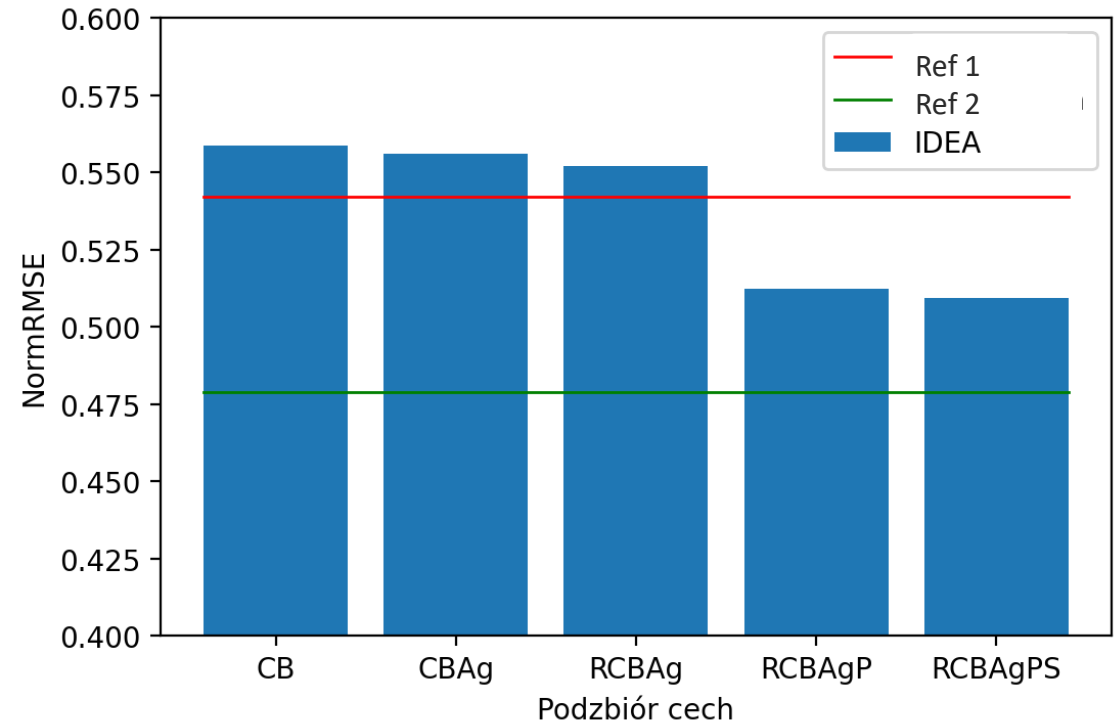
Optymalizacja hiperparametrów - PV

- Najlepszy las: brak jednoznacznej odpowiedzi
- Wybrano jako najbardziej uniwersalny:
 - ilość drzew 300
 - głębokość 10
- Przeszukiwanie na siatce parametrów - jak dla wiatru

Farma	Fold	RMSE	MAE	R	R ²	Głęb.	drzewa
1	1	0.089	0.053	0.861	0.741	5	300
	2	0.092	0.058	0.839	0.704	5	200
	3	0.080	0.049	0.875	0.766	5	300
	4	0.071	0.041	0.865	0.801	5	400
	5	0.162	0.065	0.785	0.617	10	400
2	1	0.134	0.077	0.873	0.763	10	400
	2	0.135	0.081	0.832	0.693	5	300
	3	0.115	0.070	0.883	0.780	5	300
	4	0.115	0.064	0.883	0.780	5	400
	5	0.087	0.047	0.874	0.765	5	300
3	1	0.123	0.070	0.898	0.807	10	400
	2	0.131	0.077	0.851	0.724	10	300
	3	0.130	0.074	0.868	0.753	10	300
	4	0.105	0.058	0.926	0.857	15	400
	5	0.106	0.053	0.841	0.706	10	400

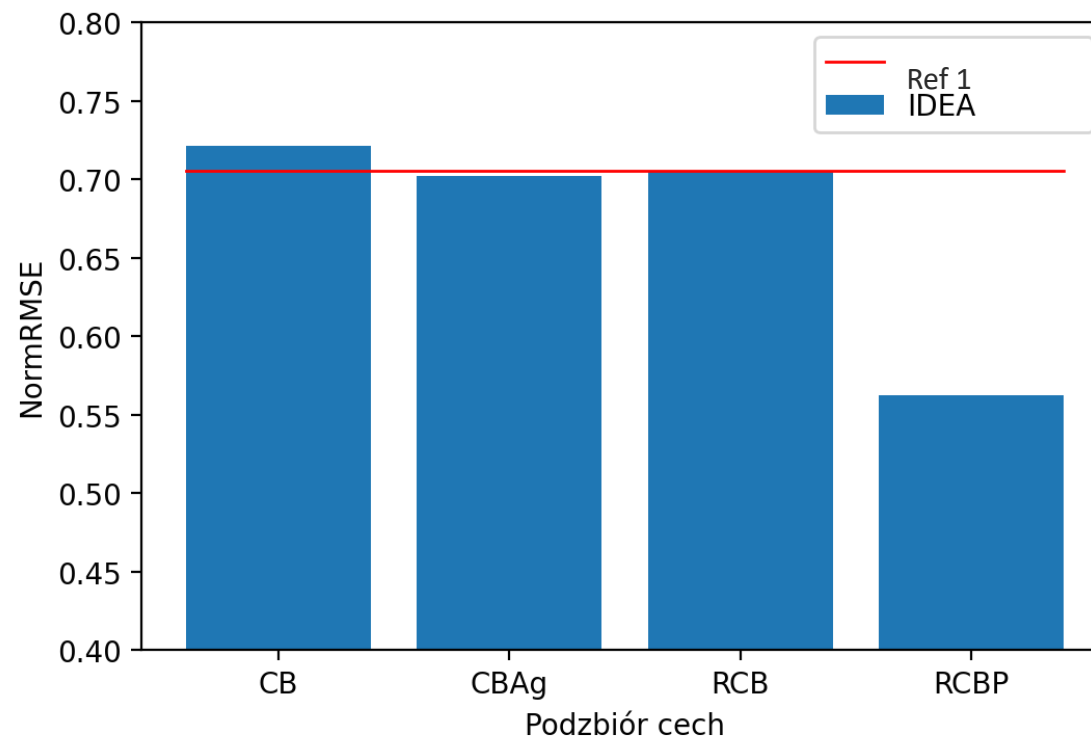
Optymalizacja - selekcja cech, wiatr

- Przykład farma 1
- Cechy bazowe (CB) - siła i kierunek wiatru w najbliższym punkcie
- Cechy bazowe z agregatami (CBAg) - CB oraz temp., ciśnienie i wilgotność względna w postaci agregatów pola pogodowego/otoczenia (średnia i odchylenie standardowe)
- RCBAg - całe pole wiatru (9 punktów, siła i kierunek) oraz CBAg
- RCBAgP – RCBAg oraz zmienność w czasie siły i kierunku wiatru
- RCBAgPS – RCBAgP rozszerzona o podzbiór atrybutów aproksymujących siłę i kierunek wiatru na rozdzielczość 15 min



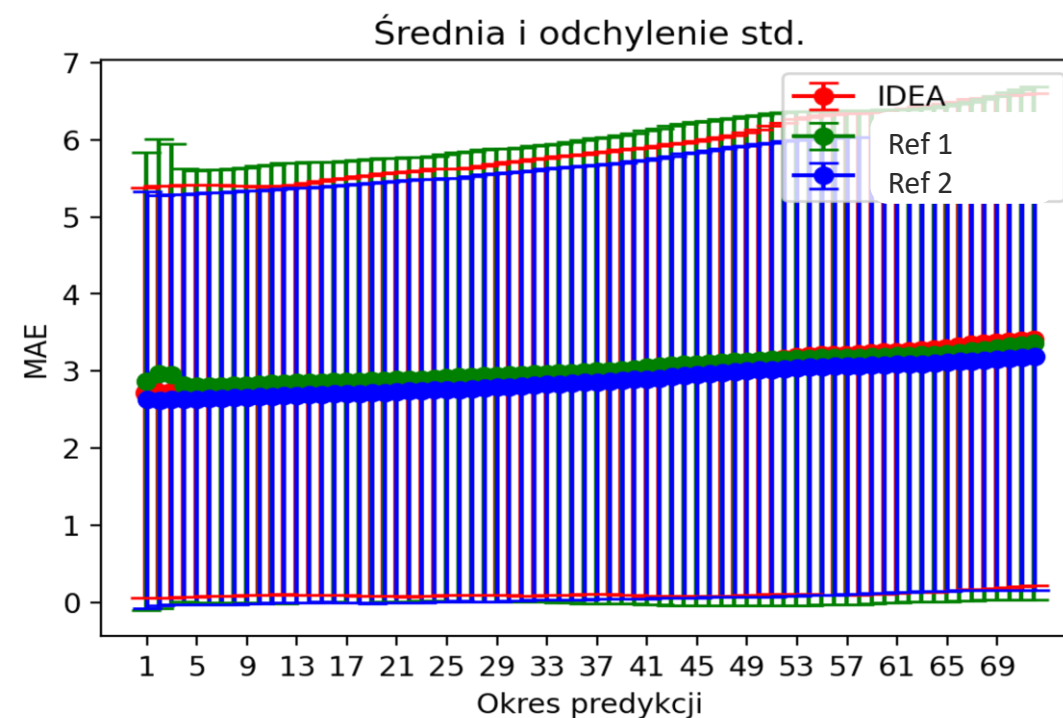
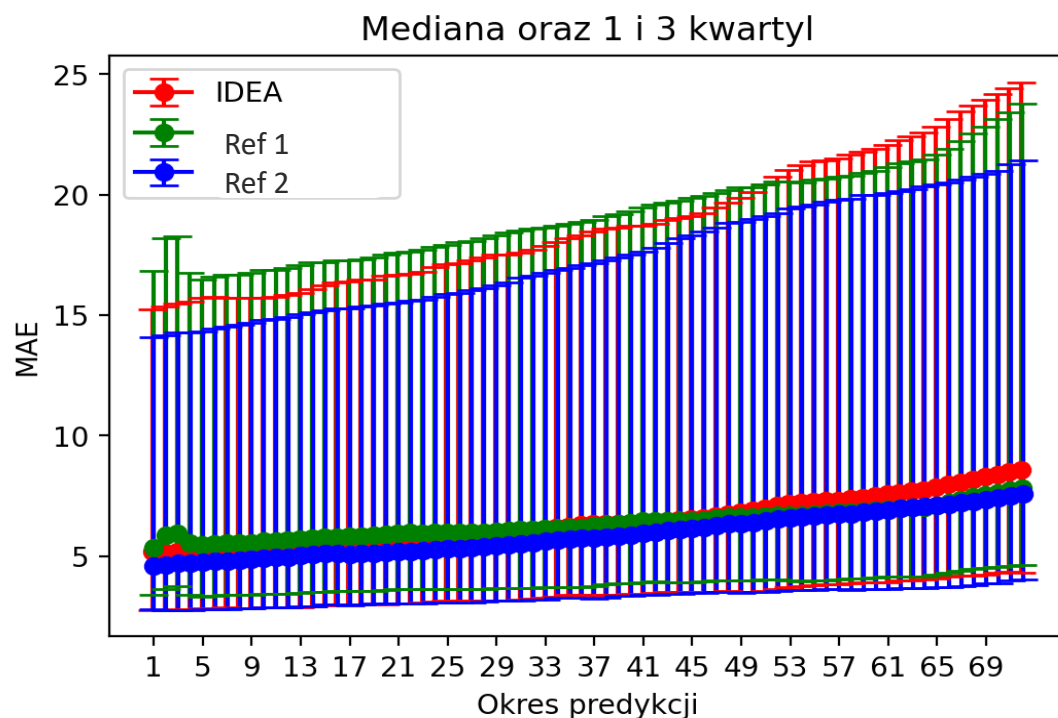
Wyniki szczegółowe - PV

- Przykład farma 1
- Cechy bazowe (CB) - temperatura i solflux w najbliższym punkcie
- Cechy bazowe z agregatami (CBAg) - CB oraz temp., solflux w postaci agregatów pola pogodowego/otoczenia (średnia i odchylenie standardowe)
- RCB – rozszerzone cechy bazowe: całe pole temp. i solflux; po 9 wartości z otoczenia
- RCBP – RCB rozszerzony o aproksymację zmiany w czasie warunków pogodowych



Wyniki szczegółowe - wiatr, przykład 2

Błąd predykcji (MAE) w zależności od wyprzedzenia prognozy



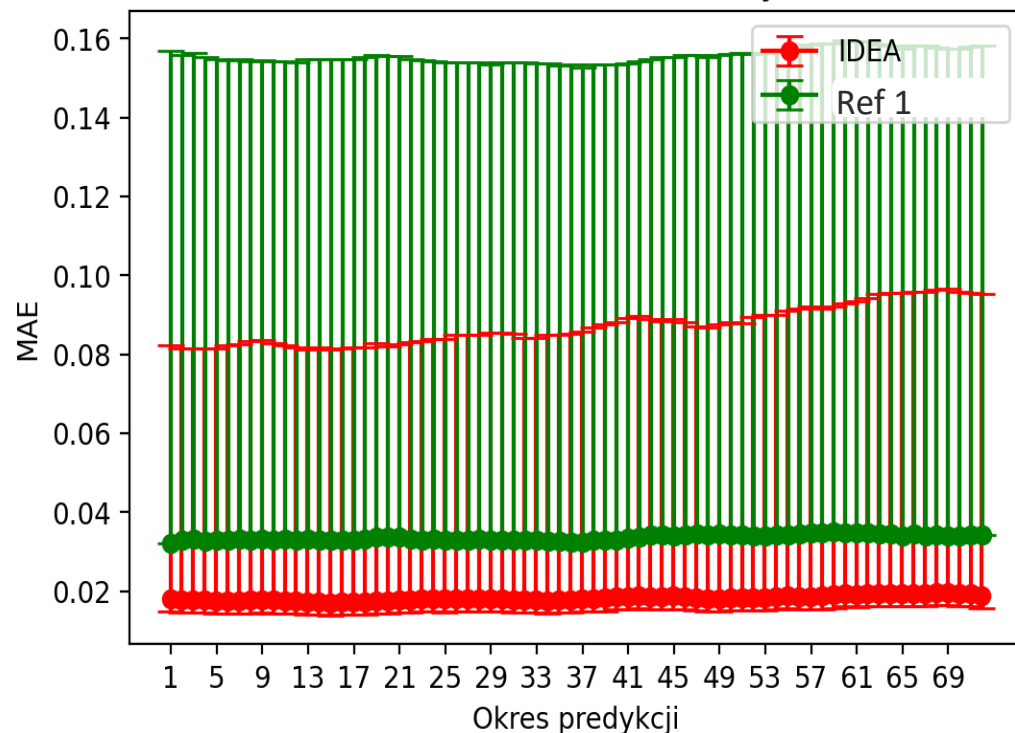
Wyniki uzyskane dla farm wiatrowych

Farma	Model	RMSE	std	MAE	std	R	std	R ²	std
1	IDEA	7.45	0.71	5.45	0.62	0.78	0.09	0.62	0.13
	Ref 1	7.36	0.78	4.66	0.54	0.78	0.09	0.62	0.13
	Ref 2	6.43	0.65	5.22	0.68	0.82	0.08	0.69	0.13
2	IDEA	12.36	1.29	9.01	1.12	0.84	0.02	0.70	0.04
	Ref 1	13.05	0.85	8.29	0.80	0.81	0.03	0.66	0.06
	Ref 2	11.55	0.93	9.17	0.80	0.85	0.02	0.72	0.04
3	IDEA	15.27	1.74	11.47	1.35	0.74	0.04	0.55	0.06
	Ref 1	18.19	2.87	11.97	1.43	0.68	0.06	0.46	0.08
	Ref 2	15.66	1.46	13.24	2.71	0.71	0.06	0.51	0.09

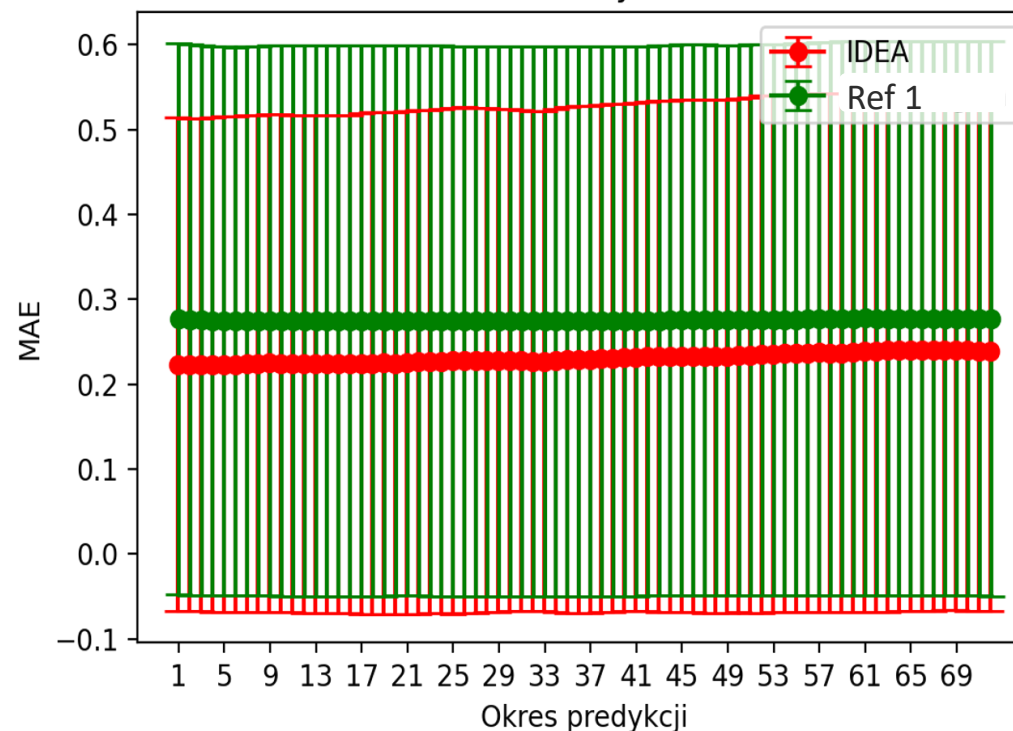
Wyniki szczegółowe - PV, przykład 1

Błąd predykcji (MAE) w zależności od wyprzedzenia prognozy

Mediana oraz 1 i 3 kwartyli



Średnia i odchylenie std.



Wyniki uzyskane dla PV

Farma	Model	RMSE	std	MAE	std	R	std	R ²	std
1	IDEA	0.0990	0.0325	0.0532	0.0084	0.8511	0.0376	0.7258	0.0631
	Ref 1	0.1266	0.0286	0.0758	0.0055	0.7252	0.0442	0.5279	0.0626
2	IDEA	0.1172	0.0176	0.0677	0.0118	0.8693	0.0189	0.7560	0.0324
	Ref 1	0.1620	0.0193	0.1047	0.0180	0.7552	0.0134	0.5705	0.0203
3	IDEA	0.1190	0.0112	0.0663	0.0091	0.8767	0.0314	0.7695	0.0554
	Ref 1	0.1844	0.0295	0.1152	0.0200	0.7300	0.0352	0.5341	0.0508

Podsumowanie wyników

- Wiatr - uzyskano dokładność porównywalną (z reguły lepszą) niż referencyjny (komercyjny) model 1
- Wiatr – dla wyprzedzeń mniejszych niż 48 h uzyskano wyniki porównywalne z referencyjnym algorytmem 2 – przyczyna sklejanie rozkładów różnych prognoz
- Wiatr – potencjalnie zastosowanie sieci konwolucyjnych i dłuższych szeregów danych historycznych może pozwolić na dalszą poprawę wyników
- PV - dokładność porównywalna lub lepsza niż dla referencyjnego algorytmu 2
- PV – bardzo krótkie szeregi danych historycznych, duży spadek jakości predykcji we wrześniu/październiku
- Potencjalnie możliwe dalsze badanie jakości dla modeli meteo o większej rozdzielczości przestrzennej, wykorzystanie danych z profili pionowych wiatru, temp., ciśnienia, rozproszonego strumieniowania krótkofalowego

Podsumowanie zastosowanie

- Wiatr i PV – zbudowano i przetestowano algorytmy predykcyjne
- Model i algorytmy IDEA przygotowane są do skalowania i działania dla dużej liczby farm i urządzeń
 - Wybranie szybkiego, sprawdzonego algorytmu uczenia (las losowy) pozwalającego na uzyskiwanie dobrych wyników przy akceptowalnej (dobrej) jakości wyników
- Prowadzone są prace nad algorytmami predykcji generacji w krótkim horyzoncie czasowym dla Tauron Ekoenergia
- Prowadzone są prace nad wdrożeniem prognoz generacji IDEA na potrzeby PSE S. A.
- Możliwość wykorzystania (po odpowiednim opomiarowaniu) dla lokalnych prosumenckich instalacji hybrydowych



IDEA

INTERDISCIPLINARY DIVISION
FOR ENERGY ANALYSES

Dziękuję za uwagę

