

Zastosowanie modeli klasy ARX do krótkoterminowego prognozowania produkcji energii elektrycznej w mikroelektrowniach wiatrowych.

Streszczenie. W referacie przedstawiono metodykę i wyniki badań możliwości zastosowania modeli szeregów czasowych klasy ARX do krótkoterminowego prognozowania produkcji energii elektrycznej w mikroelektrowniach wiatrowych. Badania polegały na doborze istotnych danych wejściowych, identyfikacji jego struktury, estymacji parametrów modelu, oceny dopasowania modelu do danych rzeczywistych oraz oceny poprawności modelu w stosunku do przyjętych założeń

Abstract. The paper deals with both a study method and results of an application of time series models ARIMAX to short-term electric energy production forecasting at wind micro power plants. The performed studies include a selection of significant input data, model structure identification, model parameters estimation, evaluation of model fit to real data as well as a model correctness according to taken assumptions (**Application of ARIMAX models to short-term electric energy production forecasting at wind micro power plants**).

Słowa kluczowe: prognozowanie produkcji energii, elektrownie wiatrowe, modele szeregów czasowych.

Keywords: energy generation forecasting, micro wind power plants, time series models.

doi:10.12915/pe.2014.07.27

Wstęp

Jedną z najbardziej rozpowszechnionych w Polsce technologii OZE w zakresie generacji energii elektrycznej są elektrownie wiatrowe, których moce zainstalowane osiągają od kilku do kilkudziesięciu MW. Przewiduje się także ich duży rozwój w zakresie mikrogeneracji [1].

Optymalne sterowanie pracą mikroinstalacji zawierających źródła wiatrowe wymaga ultra-krótkoterminowych (w horyzoncie do kilku godzin naprzód) prognoz produkcji energii ze źródeł niesterowalnych, w tym mikroelektrowni wiatrowych, których poziom generacji ściśle zależy od warunków pogodowych.

Prognozowanie produkcji energii ze źródeł wiatrowych, w krótkich horyzontach czasu jest w dalszym ciągu przedmiotem wielu badań [2-6]. Podstawowym problemem jest otrzymanie satysfakcjonujących prognoz prędkości i kierunku wiatru oraz innych towarzyszących czynników pogodowych (np. temperatura, ciśnienie, wilgotność powietrza). W tym celu używane są numeryczne prognozy pogody (ang. Numerical Weather Prediction - NWP), które stanowią cały szereg hierarchicznie uporządkowanych modeli. Dane z NWP wykorzystuje się jako wejścia do tzw. „krzywych mocy” (modele zależności generowanej mocy od prędkości wiatru) [2, 6], oraz modeli typu „czarna skrzynka”, wykorzystujące techniki statystyczne [3, 4] i/lub sztucznej inteligencji [5].

W niniejszym referacie autor skupił się na omówieniu i wstępnym przetestowaniu modelu prognostycznego, który charakteryzuje się niskim nakładem pozyskania danych. Zaproponowano wstępnie model klasy ARIMAX, wykorzystujący związki w szeregach czasowych produkcji energii oraz szeregach danych pogodowych pochodzących z lokalnej stacji meteo.

Poprzez prognozowanie krótkoterminowe rozumie się prognozowanie produkcji energii na kilka (przyjęto 6) godzin do przodu (do celów krokowego harmonogramowania pracy źródeł, zasobników i odbiorów sterowalnych w mikroinstalacji).

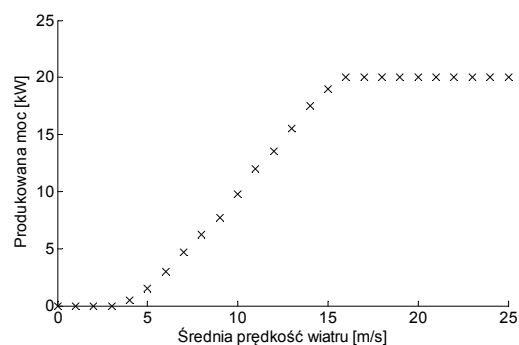
Badany obiekt

Badanym obiektem jest przydomowa mikroturbina wiatrowa (MTW) o mocy znamionowej 20 kW. Rozważana MTW jest zlokalizowana w województwie świętokrzyskim, na płaskim terenie. Teren jest otwarty ze wszystkich stron, a szczególnie z kierunku południowo-zachodniego i

zachodniego. Podstawowe parametry techniczne rozważanej MTW przedstawia tabela 1.

Tabela 1. Parametry techniczne rozpatrywanej mikroturbiny wiatrowej

Moc znamionowa	20 kW
Prędkość wiatru dla włączenia turbiny	3,5 m/s
Prędkość wiatru dla osiągnięcia mocy znamionowej turbiny	15,8 m/s
Prędkość wiatru dla wyłączenia turbiny	25,0 m/s
Średnica wirnika turbiny	9,6 m
Wysokość wieży	30 m
Typ generatora	asynchr.



Rys.1. Krzywa mocy rozpatrywanej mikroturbiny wiatrowej

Dla rozważanej MTW otrzymano dane o godzinowej produkcji energii elektrycznej z lat 2010 i 2011. Z danych usunięto okresy, w których MTW była wyłączona z powodów awaryjnych (384 h w roku 2010 i 302 h w roku 2011). Ponieważ nie były dostępne dane pogodowe z lokalnej stacji meteo, wykorzystano interpolowane przestrzennie (dla współrzędnych geograficznych rozpatrywanej MTW) pomiary prędkości i kierunku wiatru w stacjach meteorologicznych IMGW. Otrzymane w odstępnie 10 min. chwilowe wartości prędkości i kierunku wiatru zostały uśredniono do pełnej godziny.

Identyfikacja klasy modelu stochastycznego

Produkcja mocy elektrycznej z niesterowalnych OZE jest funkcją czasu, której wartości w każdym momencie są zmiennymi o charakterze losowym. Matematycznym

modelem dogodnym do opisu struktury probabilistycznej procesu jest proces stochastyczny [7]. Niech będzie dana następująca funkcja losowa $Y: T \times \mathcal{E} \rightarrow \mathbb{R}$, gdzie $T = \{t\}$ jest zbiorem liczb rzeczywistych stanowiących oś czasu oraz $\mathcal{E} = \{\xi\}$ jest przestrzenią zdarzeń elementarnych. Jeśli zbiór T stanowi oś czasu, to funkcja losowa $Y(t, \xi)$ jest nazywana procesem stochastycznym (przyjęto dalsze oznaczenie Y_t).

Wśród klasy liniowych modeli regresji wielomianowej rozważa się modele ARIMAX [7].

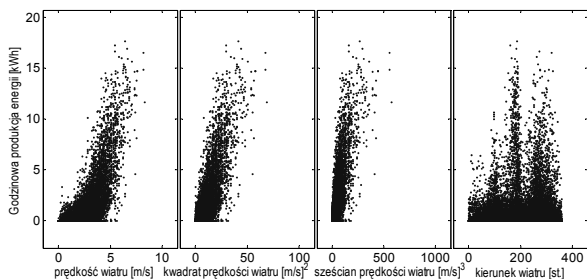
$$(1) \quad A(q)\Delta^d y(t) = \sum_{i=1}^{n_u} B_i(q)u_i(t - n_{ki}) + C(q)e(t)$$

gdzie: $A(q)$, $B(q)$, $C(q)$ – wielomiany skończonego rzędu, q – operator przesunięcia w czasie, $u_i(t)$ – i -te wejście modelu, n_u – liczba wejść modelu $y(t)$ – wyjście modelu, n_{ki} – opóźnienie i -tego wejścia charakteryzujące czas odpowiedzi modelu, Δ^d – operator d -krotnego różnicowania, $e(t)$ – proces białego szumu.

Szacowanie parametrów modelu można przeprowadzić w sposób klasyczny, tj. metodą najmniejszych kwadratów (MNK) lub w sposób rekurencyjny (RMNK) [2, 3].

Produkcja energii elektrycznej z turbosespołów wiatrowych zależy silnie od potęgi trzeciej prędkości wiatru. W literaturze wykazano, że dla klasy addytywnych modeli regresji wystarczająca jest nawet aproksymacja tej zależności wielomianem stopnia drugiego [2, 3].

Na rysunku 2 przedstawiono wykres rozproszenia między produkcją godzinową energii, a odpowiednio – prędkością wiatru, kwadratem prędkości, sześcianną prędkości wiatru oraz kierunkiem wiatru. W związku z tym, jako potencjalne zmienne egzogeniczne dobrano kierunek (azymut) wiatru oraz prędkość wiatru w potęgach pierwszej, drugiej i trzeciej.



Rys.2. Wykres rozproszenia potencjalnych zmiennych egzogenicznych względem produkcji energii z mikroturbiny wiatrowej

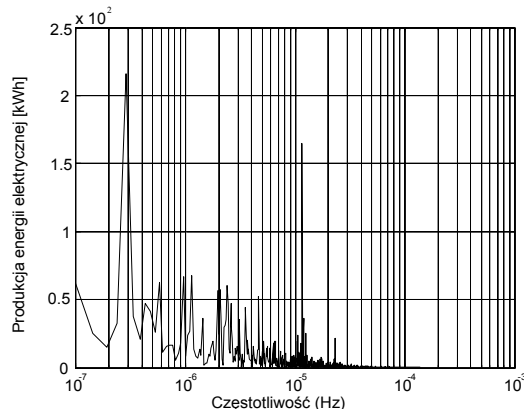
Na rysunku 2 widać, że zwiększanie potęgi funkcji prędkości wiatru, przybliża tę zależność do funkcji liniowej. Obserwuje się także większą produkcję dla wybranych przedziałów kierunku wiatru, tj. nieco poniżej 200 stopni oraz poniżej 300.

Dla otrzymanej realizacji procesu stochastycznego produkcji energii wyznaczono periodogram (rys.3), na którym widoczny jest znaczny udział składowej okresowej ok. sześćdziesięciodniowej ($1/60$ dni = $1,9290 \cdot 10^{-7}$ Hz) i dobowej ($1/24$ h = $1,1574 \cdot 10^{-5}$ Hz). Potwierdzono również istnienie tych składowych w realizacji procesu stochastycznego prędkości wiatru.

Zatem, jako potencjalną zmienną zmienną wejściową dołącza się harmoniczną dobową (składowe: $\cos(2\pi h/24)$ oraz $\sin(2\pi h/24)$).

W tabeli 2 przedstawiono współczynniki korelacji próbkowej Pearsona (r). Postawiono przy tym następujące hipotezy:

- $H_0: r = 0$
- $H_1: r \neq 0$



Rys.3. Krzywa mocy rozpatrywanej mikroturbiny wiatrowej

Tabela 2. Współczynniki korelacji Pearsona otrzymane dla analizowanych zmiennych

–	E	V	v^2	v^3	θ	$\cos(\cdot)$	$\sin(\cdot)$
E	1,00	0,70	0,79	0,80	0,10	-0,14	-0,07
V	0,70	1,00	0,95	0,85	0,04	-0,33	-0,22
v^2	0,79	0,95	1,00	0,97	0,04	-0,28	-0,18
v^3	0,80	0,85	0,97	1,00	0,03	-0,24	-0,14
θ	0,10	0,04	0,04	0,03	1,00	-0,04	0,00
$\cos(\cdot)$	-0,14	-0,33	-0,28	-0,24	-0,04	1,00	0,00
$\sin(\cdot)$	-0,07	-0,22	-0,18	-0,14	0,00	0,00	1,00

Dla poszczególnych współczynników korelacji próbkowej Pearsona otrzymano p-wartości poniżej 0,001, co oznacza odrzucenie hipotezy zerowej na rzecz hipotezy alternatywnej. Tym samym współczynniki korelacji istotnie różnią się od zera.

Dodatkowo, dokonano sekwencyjnego doboru zmiennych egzogenicznych wykorzystując ocenę przyjętej funkcji kryterialnej (przyjęto współczynnik determinacji R^2). Technika ta polega na dodawaniu/odejmowaniu zmiennych do/z utworzonego już zbioru zmiennych i oceny jak istotnie dodanie/odjęcie zmiennej do/ze zbioru zmienia wartość funkcji kryterialnej [8]. Wykorzystano przy tym 10-krotną walidację krzyżową.

Na podstawie oceny otrzymanych współczynników korelacji próbkowej oraz wyników metody sekwencyjnego doboru zmiennych wybrano jedną zmienną egzogeniczną: sześcianną prędkości wiatru (ostatni dostępny pomiar tej wielkości).

Dokonano estymacji tymczasowego modelu regresji liniowej dla wyprzedzeń prognozy od 1 do 6 h i dla otrzymanych ciągów reszt, sprawdzono niestacjonarność procesu za pomocą rozszerzonego testu Dickeya-Fullera (ADF). Postawiono następujące hipotezy:

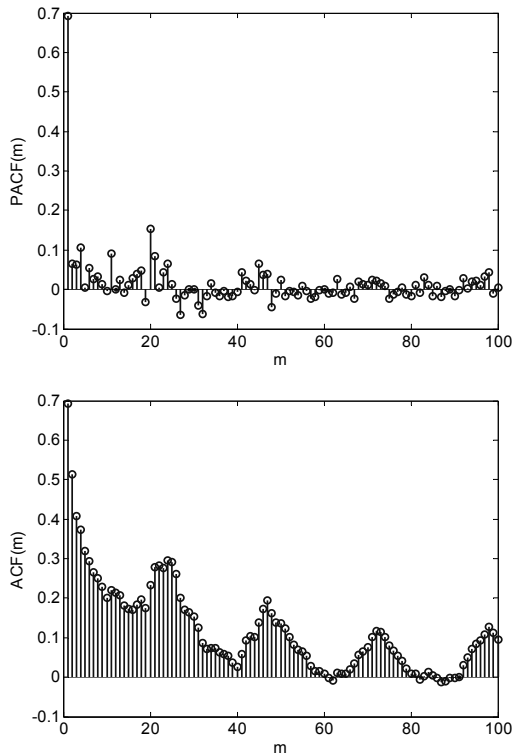
- $H_0: DF = 0$
- $H_1: DF < 0$

Otrzymano wartości statystyki DF dużo poniżej wartości krytycznych odczytanych na poziomie ufności $\alpha = 0,05$. Zatem, dla rozważanego horyzontu prognozy hipoteza zerowa zostaje odrzucona na rzecz hipotezy alternatywnej i tym samym stwierdza się stacjonarność procesu stochastycznego ciągu reszt modelu.

W celu identyfikacji modelu predykcyjnego wyznaczono dla ciągów reszt (tymczasowe modele regresji dla wyprzedzeń prognozy od 1 do 6 h) estymatory funkcji autokorelacji cząstkowej (PACF) i autokorelacji (ACF) (rys.4).

Na podstawie przeprowadzonej analizy PACF i ACF można sądzić, że badany model nie będzie miał części MA (brak urwania funkcji autokorelacji), a będzie miał część AR (urwanie funkcji autokorelacji cząstkowej). Zatem, rozpatruje się modele klasy ARX z jedną zmienną egzogeniczną, tj.:

$$(2) \quad A(q)y(t) = B(q)u(t - n_k) + e(t)$$



Rys. 4. Wartości funkcji autokorelacji cząstkowej i autokorelacji otrzymane z ciągu reszt modelu dla wyprzedzenia prognozy 1 h.

Badania modelu prognostycznego

Rozważano prognozowanie produkcji energii w horyzoncie czasu od 1 do 6 h, gdzie 1 oznacza najbliższą pełną godzinę. Przyjęto oddzielny model na każde wyprzedzenie godzinowe prognozy. Zrealizowano następujący algorytm postępowania:

1. Analiza obserwacji odstających
2. Odjęcie wartości średniej od zmiennej wyjściowej i zmiennej egzogenicznej
3. Podział danych na dane estymacyjne (rok 2010) i walidacyjne (rok 2011) – w proporcji 0,56/0,44
4. Wybór struktury modelu ARX na podstawie kryterium Akaike, opartego na funkcji strat modelu i walidacji krzyżowej;
5. Estymacja parametrów modelu metodą najmniejszych kwadratów (MNK),
6. Ocena dopasowania modeli do danych rzeczywistych produkcji energii
7. Obliczenie błędów ex-post prognoz
8. Diagnostyka modelu (badania normalności rozkładu reszt, autokorelacji reszt, średniej reszt i heteroskedastyczności reszt).

Wyniki otrzymanych struktur modeli ARX przedstawiono w tabeli 3. Pokazano również uzyskane na podstawie estymacji parametrów modelu współczynniki determinancji dla danych estymacyjnych (R_e^2) oraz walidacyjnych (R_v^2).

Tabela 3. Struktura i wskaźniki estymacji parametrów modelu ARX

Wyprzedzenie prognozy	n_a	n_b	n_k	R_e^2	R_v^2
1	8	6	1	59,20%	58,93%
2	8	6	2	43,34%	42,60%
3	8	6	3	32,94%	32,73%
4	8	6	4	25,34%	25,20%
5	8	6	5	19,72%	19,62%
6	8	6	6	15,70%	15,62%

gdzie: n_a – stopień wielomianu $A(q)$, n_b – stopień wielomianu $B(q)$, n_k – opóźnienie zmiennej egzogenicznej (wielomian $B(q)$).

Wyznaczono następujące statystyki błędów prognoz ex-post:

$$(3) \quad nRMSE_1 = \frac{\sqrt{\frac{1}{N} \sum_{h=1}^N (A_h - A_h^*)^2}}{P_n}$$

$$(4) \quad nRMSE_2 = \frac{\sqrt{\frac{1}{N} \sum_{h=1}^N (A_h - A_h^*)^2}}{\frac{1}{N} \sum_{h=1}^N A_h}$$

$$(5) \quad nMAE_1 = \frac{\frac{1}{N} \sum_{h=1}^N |A_h - A_h^*|}{P_n}$$

$$(6) \quad nMAE_2 = \frac{\frac{1}{N} \sum_{h=1}^N |A_h - A_h^*|}{\frac{1}{N} \sum_{h=1}^N A_h}$$

gdzie: A_h – zmierzona godzinowa produkcja energii elektrycznej, A_h^* – prognozowana godzinowa produkcja energii elektrycznej, N – licznosc próby danych.

Należy zaznaczyć, że błędy $nRMSE$ i $nMAE$ typu 1 i 2 różnią się sposobem normalizowania. Wartości obliczonych błędów przedstawiono na rysunku 5.

W badaniach modeli prognostycznych, niezbędnym elementem jest przyjęcie modelu referencyjnego. W tym celu wykorzystano dostępną krzywą mocy MTW (rys.1). Wynik (prognoza) modelu referencyjnego jest niezależny od rozważanego horyzontu prognozy. Obliczone wartości błędów $nRMSE$ i $nMAE$ typu 1 i 2 przedstawiono w tabeli 4.

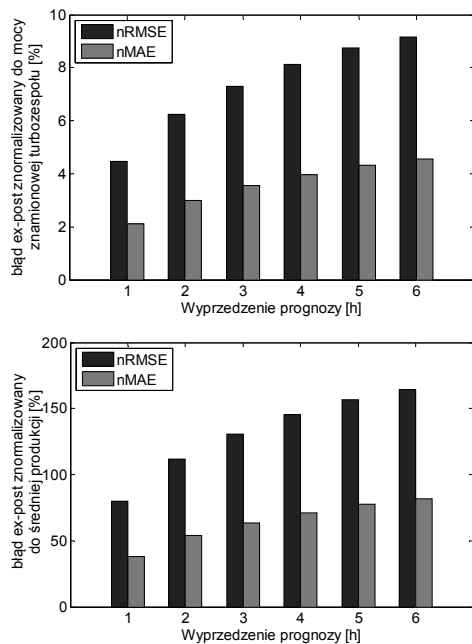
Tabela 4. Błędy ex-post prognoz dla modelu referencyjnego

Statystyka błędu	Horyzont prognozy					
	1	2	3	4	5	6
$nRMSE_1$	8,12%	9,42%	9,41%	9,60%	9,64%	9,63%
$nRMSE_2$	4,03%	4,57%	4,57%	4,65%	4,67%	4,66%
$nMAE_1$	151,5%	175,8%	175,6%	179,2	179,9%	179,7%
$nMAE_2$	75,2%	85,3%	85,3%	86,8%	87,2%	87,0%

Otrzymane wartości błędów prognoz ex-post dla modelu referencyjnego są wyższe niż dla badanego modelu klasy ARX, przy czym dla horyzontu prognozy równego 6 h, wartości błędów są zbliżone.

Należy wskazać na duże różnice w wartościach błędów ex-post typu 1 i 2. W rozważanej MTW średnia produkcja energii elektrycznej (ok. 1,1 kWh) jest dużo poniżej wartości mocy znamionowej MTW (20 kW). Z tego względu błędy

typu 1 mają relatywnie małe wartości, co nie oznacza dobrej jakości modelu. Lepszy obraz dokładności otrzymywanych prognoz dają wartości błędów typu 2, które są raczej niesatisfakcjonujące.



Rys.4. Wartości funkcji autokorelacji cząstkowej i autokorelacji otrzymane z ciągu reszt modelu dla wyprzedzenia prognozy 1 h.

Diagnostyka modelu

W celu dokładnej oceny modelu prognostycznego, niezbędna jest analiza własności otrzymanych reszt modelu (błędów prognozy).

W pierwszej kolejności zbadano, czy rozkład reszt pochodzi z rozkładu normalnego. W tym celu zastosowano test Lillieforsa i postawiono następujące hipotezy [7]:

- $H_0: e(t)$ ma rozkład normalny,
- $H_1: e(t)$ ma inny rozkład.

Dla rozważanych modeli (wyprzedzeń prognozy od 1 do 6 h) otrzymano p-wartości poniżej 0,001, co przy założeniu poziomu istotności $\alpha = 0,05$ powoduje odrzucenie hipotezy zerowej i tym samym rozkład reszt modeli prognostycznych nie jest rozkładem normalnym. Ten wniosek nie dyskwalifikuje otrzymanych modeli, jedynie ogranicza stosowanie niektórych testów diagnostycznych modelu [7].

Bardzo istotnym badaniem własności modelu prognostycznego jest analiza autokorelacji reszt. Ponieważ modele ARX nie zawierają wyrazów wolnych oraz otrzymane rozkłady reszt nie są rozkładami normalnymi, sprawdzono istotność autokorelacji reszt za pomocą testu Breuscha-Godfrey'a. Przyjęto następujące hipotezy [7]:

- $H_0: e(t)$ są nieskorelowane
- $H_1: e(t)$ stanowią proces $AR(p)$ lub $MA(p)$.

Dla każdego ciągu reszt modeli oraz wyznaczono ciąg statystyk TR^2 . Największe wartości występują w odstępie 24 h, jednak otrzymane p-wartości są nieco powyżej przyjętego $\alpha = 0,05$, zatem, błędy prognoz z otrzymanych modeli nie są ze sobą skorelowane.

Kolejnym krokiem weryfikacji poprawności modelu jest badanie, czy średnia ciągu reszt modeli prognostycznych jest równa 0. Wykorzystano do tego celu test dla średniej ze statystyką t . Rozważa się następujące hipotezy [7]:

- $H_0: E(e(t)) = 0$
- $H_1: E(e(t)) \neq 0$.

Dla rozpatrywanych 6. modeli ARX, otrzymano p-wartości w granicach 0,8÷0,9, co oznacza, że nie ma podstaw do odrzucenia H_0 , a zatem oczekiwana wartość błędów prognozy jest równa 0.

Ostatnim z rozważanych badań reszt modeli ARX jest testowanie ich heteroskedastyczności. W tym celu wykorzystano test Breuscha-Pagana i rozpatrzono następujące hipotezy [7]:

- $H_0: \text{Var}(e(t)) = \text{const}$
- $H_1: \text{Var}(e(t)) \neq \text{const}$.

Dla każdego z modeli otrzymano p-wartości w granicach 0,5÷0,7, co oznacza, że nie ma podstaw na poziomie istotności $\alpha = 0,05$ do odrzucenia H_0 , a zatem stałość wariancji jest zachowana.

Podsumowanie i wnioski

W artykule przedstawiono sposób i wyniki badania modelu prognozowania produkcji energii z przydomowej elektrowni wiatrowej (mikroturbiny wiatrowej) dla różnych horyzontów czasu (od 1 do 6 h). Zaproponowano wstępnie model stochastyczny klasy ARIMAX wykorzystujący zmienną egzogeniczną, której wartości są brane z lokalnego pomiaru prędkości wiatru (np. ze przydomowej stacji meteo). W trakcie badań wybrano ostatecznie zbiór modeli klasy ARX.

Pomimo poprawności modelu prognostycznego (uzyskane wyniki diagnostyki modelu), jakość uzyskiwanych prognoz nie jest satisfakcjonująca. Należy zaznaczyć, że błędy prognoz ex-post dla horyzontu 5 i 6 h są zbliżone do modelu referencyjnego.

Uzyskane wyniki badań potwierdzają, że bardzo trudno jest uzyskać satisfakcjonującą dokładność prognoz bez numerycznych prognoz pogody.

W dalszych badaniach, autor zamierza sprawdzić zasadność stosowania tego typu modeli stochastycznych w problemie prognozowania krótkoterminowego (do celów bieżącego sterowania źródłami, zasobnikami i odbiornikami energii elektrycznej), przy większej rozdzielczości czasowej i krótszym horyzoncie prognozy.

Praca sfinansowana z projektu badawczego „System wspomaganie udziału w rynku energii elektrycznej dla operatorów OZE” INNOTECH-K1/H11/21/158535/NCBR/12

LITERATURA

- [1] Paska, J., Wytwarzanie rozproszone energii elektrycznej i ciepła. Oficyna Wydawnicza Politechniki Warszawskiej (2010).
- [2] Foley A.M., Leahy P.G., Marvuglia A., McKeogh E.J., Current methods and advances in forecasting of wind power generation, *Renewable Energy* 37 (2012), no. 1, 1-8.
- [3] Kavasseri R.G., Seetharaman K., Day-ahead wind speed forecasting using ARIMA models, *Renewable Energy* 34 (2009), no. 5, 1388-1393.
- [4] Popławski, T., Szeląg P., Wykorzystanie własności podobieństwa procesów do prognozowania mocy przez turbiny wiatrowe. *Rynek Energii*, 1 (2011), 103-107.
- [5] Piotrowski, P., Analiza statystyczna danych mających wpływ na produkcję energii elektrycznej przez farmę wiatrową oraz przykładowe prognozy krótkoterminowej, *Przełęcz Elektrotechniczny*, 88 (2012), nr.3a, 161-164.
- [6] Lange M., Focken U., Physical approach to short-term wind power forecast, Springer (2005).
- [7] Ljung, L., System Identification: Theory for the User, Prentice-Hal PTR (1999).
- [8] Ćwik J., Koronacki J., "Statystyczne systemy uczące się." Akademicka Oficyna Wydawnicza EXIT (2008).

Autor: dr inż. Jacek Wasilewski, Politechnika Warszawska, Instytut Elektroenergetyki, ul. Koszykowa 75, 00-662 Warszawa, E-mail: jacek.wasilewski@ien.pw.edu.pl