

Prognozowanie produkcji budowlano montażowej w województwie dolnośląskim. Część I

Magdalena Rogalska

*Katedra Inżynierii Procesów Budowlanych, Wydział Budownictwa i Architektury,
Politechnika Lubelska, e-mail: m.rogalska@tlen.pl*

Streszczenie: Artykuł jest częścią pierwszą cyklu „Prognozowanie produkcji budowlano montażowej w województwie dolnośląskim”. Założono, że wynagrodzenie pracowników będzie jedną ze zmiennych niezależnych do wyznaczenia wielkości produkcji. Prognozowano wynagrodzenia pracowników w sektorze budowlanym metodami regresji wielorakiej i metodą autoregresji średniej ruchomej SARIMA. Przeprowadzono analizę wyników obliczając błędy ME, MAE, MPE, MAPE oraz współczynniki Theila I, I2, I12, I22, I32. Sformułowano wnioski z obliczeń. Wyznaczono równanie regresji wielorakiej z 12 predyktorami wytypowanymi spośród 53 zmiennych niezależnych. Uzyskano dane prognozowane do predykcji produkcji budowlano montażowej.

Słowa kluczowe: prognozowanie, regresja wieloraka, SARIMA, wynagrodzenia pracowników.

1. Wstęp

Podjęto próbę prognozowania produkcji budowlano montażowej w województwie dolnośląskim. Do prognozowania wykorzystano utworzoną bazę danych pogodowych wzbogaconą o dane ekonomiczne w postaci średniego wynagrodzenia w sektorze budowlanym oraz ilości zatrudnionych pracowników w budownictwie w latach 2000-2009 w województwie dolnośląskim. Dane ekonomiczne uzyskano w Dolnośląskim Oddziale Głównego Urzędu Statystycznego we Wrocławiu.

Celem prowadzonych obliczeń jest wykonanie prognozy produkcji budowlano montażowej w okresie od stycznia do września 2010 roku. Prognoza bazować będzie na następujących danych:

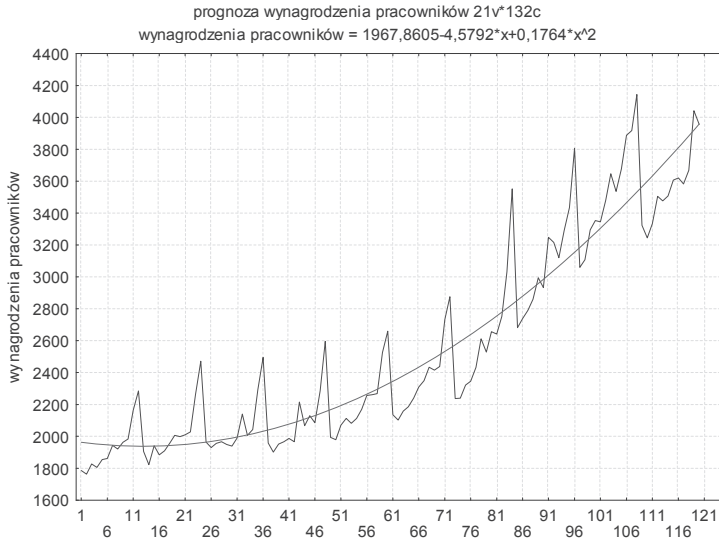
- Produkcja budowlano montażowa przesunięta o 12 okresów
- Prognoza wynagrodzenia pracowników w sektorze budowlanym w województwie dolnośląskim prognozowane na okresy I-IX 2010 (część 1)
- Prognoza liczby pracowników zatrudnionych w budownictwie w województwie dolnośląskim na okresy I-IX 2010 (część 2)
- Prognoza temperatury potencjalnej w województwie dolnośląskim na okresy I-IX 2010 (część 3)

Przeprowadzono analizę wyników obliczając błędy ME, MAE, MPE, MAPE oraz współczynniki Theila I, I2, I12, I22, I32. Sformułowano wnioski z obliczeń.

Wykorzystywano metodę regresji wielorakiej [3,9,10] oraz metodę autoregresji średniej ruchomej SARIMA [1,2,4,5,6,7,8].

2. Prognoza wynagrodzeń w sektorze budowlanym w województwie dolnośląskim I-IX 2010

Dane uzyskano z Głównego Urzędu Statystycznego Oddział Wrocław. Wartości średniego wynagrodzenia w sektorze budowlanym w województwie dolnośląskim w latach 2000-2009 przedstawiono na Rys. 1.



Rys. 1. Wykres zależności średniego wynagrodzenia w sektorze budowlanym w województwie dolnośląskim w badanych okresach od stycznia 2000 do grudnia 2009. Na osi x oznaczono okresy odpowiadające kolejnym miesiącom.

Fig. 1. A plot of the average salary in the construction sector in Lower Silesia in the examined period from January 2000 to December 2009. The x-axis corresponding to the successive periods were determined for each month.

Wyraźnie widoczny jest trend wielomianowy oraz okresowość danych. W arkuszu wyjściowym umieszczono następujące zmienne Tabela 1:

Tabela 1. Zestawienie zmiennych z oznaczeniami do prognozowania wynagrodzeń w sektorze budowlanym w województwie dolnośląskim.

Table 1. List of variables to predict salaries in the construction sector in Lower Silesia.

Zmienna	Nazwa zmiennej	Zmienna	Nazwa zmiennej
V1	data	V9	maj
V2	t	V10	czerwiec
V3	Wynagrodzenia pracowników 2000-2009	V11	lipiec
V4	Wynagrodzenia pracowników 2000-2009 opóźnione o 12 okresów	V12	sierpień
V5	styczeń	V13	wrzesień
V6	luty	V14	październik
V7	marzec	V15	listopad
V8	kwiecień	V16	grudzień

Arkusz wyjściowy (Tabela 2) zawiera zmienne zestawione w tabeli 1. Dekompozycję sezonową zrealizowano metodą (1,0,-1). Metoda polega na tym, że wartość

1 przypisywana jest zmiennej sezonowej wówczas gdy w zmiennej data jest ten sam miesiąc, np. dla przypadku pierwszego gdy zmienna $v1$ to 2000.1 wówczas zmienna $v5$ styczeń przyjmuje wartość 1 a pozostałe zmienne sezonowe (inne miesiące) wartość 0. Przypadki zawierające marzec zawsze dla wszystkich zmiennych $v5-v16$ przyjmują wartość (-1).

Tabela 2. Arkusz wyjściowy do prognozowania wynagrodzenia pracowników.
Table 2. Output sheet to predict salaries of employees.

	1 data	2 t	3 wynagrodzenia pracowników	4 Wynagrodzenia pracowników lag 12	5 styczeń	6 luty	7 marzec	8 kwiecień	9 maj	10 czerwiec	11 lipiec	12 sierpień	13 wrzesień	14 październik	15 listopad	16 grudzień
1	2000.1	1	1786		1	0	0	0	0	0	0	0	0	0	0	0
2	2000.2	2	1763		0	1	0	0	0	0	0	0	0	0	0	0
3	2000.3	3	1827		-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
4	2000.4	4	1806		0	0	0	1	0	0	0	0	0	0	0	0
5	2000.5	5	1855		0	0	0	0	1	0	0	0	0	0	0	0
6	2000.6	6	1861		0	0	0	0	0	1	0	0	0	0	0	0
7	2000.7	7	1944		0	0	0	0	0	0	1	0	0	0	0	0
8	2000.8	8	1921		0	0	0	0	0	0	0	1	0	0	0	0
9	2000.9	9	1963		0	0	0	0	0	0	0	0	1	0	0	0
10	2000.10	10	1984		0	0	0	0	0	0	0	0	0	1	0	0
11	2000.11	11	2163		0	0	0	0	0	0	0	0	0	0	1	0
12	2000.12	12	2285		0	0	0	0	0	0	0	0	0	0	0	1
13	2001.1	13	1906	1786	1	0	0	0	0	0	0	0	0	0	0	0
14	2001.2	14	1922	1763	0	1	0	0	0	0	0	0	0	0	0	0
15	2001.3	15	1941	1827	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
16	2001.4	16	1885	1806	0	0	0	1	0	0	0	0	0	0	0	0
17	2001.5	17	1910	1855	0	0	0	0	1	0	0	0	0	0	0	0
18	2001.6	18	1957	1861	0	0	0	0	0	1	0	0	0	0	0	0
19	2001.7	19	2006	1944	0	0	0	0	0	0	1	0	0	0	0	0
20	2001.8	20	1998	1921	0	0	0	0	0	0	0	1	0	0	0	0
21	2001.9	21	2010	1963	0	0	0	0	0	0	0	0	1	0	0	0
22	2001.10	22	2028	1984	0	0	0	0	0	0	0	0	0	1	0	0
23	2001.11	23	2262	2163	0	0	0	0	0	0	0	0	0	0	1	0
24	2001.12	24	2472	2285	0	0	0	0	0	0	0	0	0	0	0	1

A. Metoda regresji wielorakiej

MODEL R1($v3;v2,v4$)

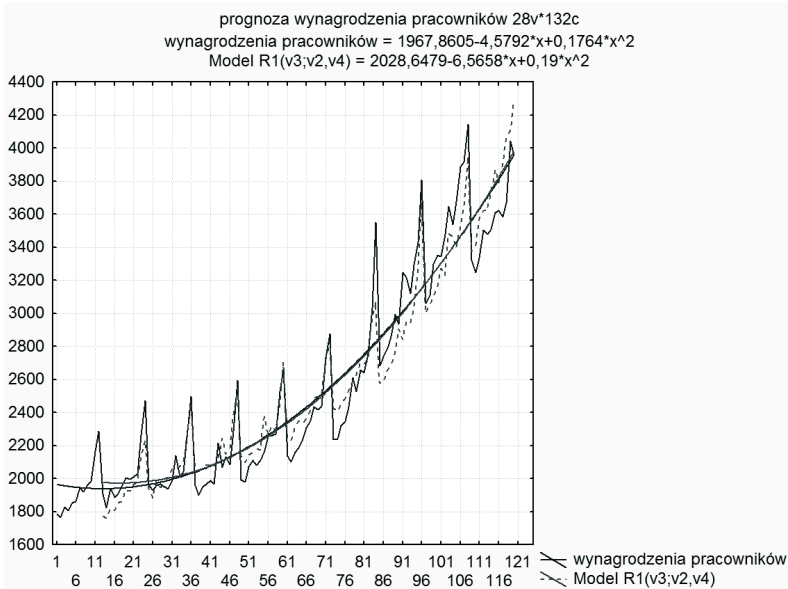
UWAGA: dla potrzeb pracy wprowadzono oznaczenia modeli regresji gdzie R oznacza regresję wieloraką, liczba po R to kolejny model regresyjny, v_i w nawiasie przed średnikiem to zmienna zależna, v_k w nawiasie po średniku to zmienne zależne – predyktory.

Zastosowano metodę regresji wielorakiej wstecznej. Otrzymano równanie regresyjne w postaci (1), podsumowanie regresji zmiennej zależnej $v3$ zamieszczono w Tabeli 3 ,wykres liniowy zależności zmiennej zależnej $v3$ i prognozy $R(v3;v2,v4)$ przedstawiono na Rys. 2.

Tabela 3. Podsumowanie regresji model R1 ($v3;v2,v4$).
Table 3. Summary of regression model R1 ($v3, v2, v4$).

Podsumowanie regresji zmiennej zależnej: wynagrodzenia pracowników (prognoza wynagrodzenia pracowników)						
N=108	R = ,97103845 R^2 = ,94291568 Skoryg. R2 = ,94182836					
	F(2,105) = 867,19 p < 0,0000 Błąd std. estymacji: 156,05					
	b^*	Bł. std. z b^*	b	Bł. std. z b	t(105)	p
W. wolny			297,5985	79,77521	3,73046	0,000310
t	0,304862	0,046243	6,2976	0,95525	6,59261	0,000000
Wynagrodzenia pracowników lag 12	0,695523	0,046243	0,7807	0,05191	15,04062	0,000000

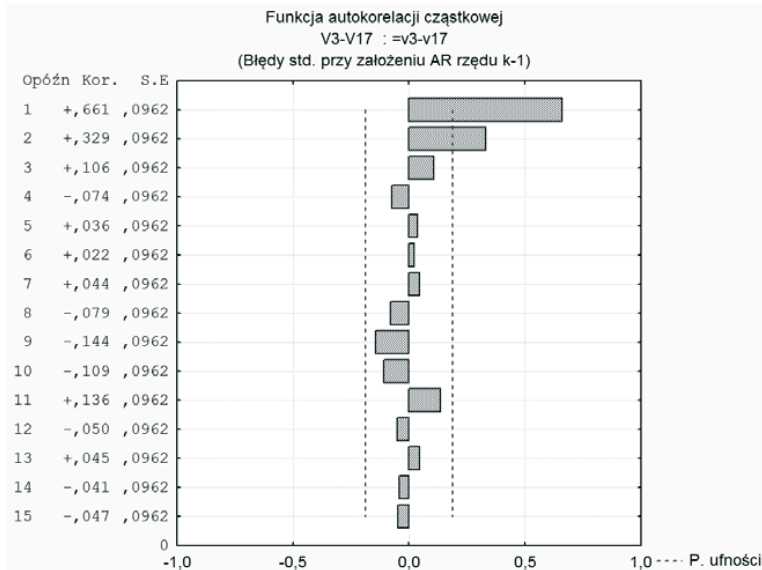
$$v3=297,5985+6,2976v2+0,7807v4 \tag{1}$$



Rys. 2. Wykres zależności średniego wynagrodzenia w sektorze budowlanym w województwie dolnośląskim w badanych okresach od stycznia 2000 do grudnia 2009 oraz prognozy uzyskanej metodą regresji $R1(v3;v2,v4)$.

Fig. 2. A plot of the average salary in the construction sector in Lower Silesia in the examined period from January 2000 to December 2009 and the forecasts obtained by regression $R1(v3;v2,v4)$.

Wniosek: Pomimo zadawalającej wartości R^2 skorygowanego równego 0,9418 oraz $p = 0,0000$ z Rys. 2 widać wyraźnie, że prognoza nie w pełni spełnia oczekiwania. Celem sprawdzenia wykonano wykres autokorelacji reszt Rys. 3.



Rys. 3. Funkcja autokorelacji cząstkowej reszt modelu $R1(v3;v2,v4)$.

Fig. 3. Partial autocorrelation function of the model $R1(v3;v2,v4)$ residuals.

Tabela 4. Prognoza modelu $R1(\nu3;\nu2,\nu4)$ wynagrodzenia pracowników od stycznia do października 2010 r. oraz uzyskane w 2011 wartości rzeczywiste.

Table 4. Prediction model $R1(\nu3;\nu2,\nu4)$, staff salaries from January to October 2010, and received in 2011 in real terms

T	Data	Prognoza wynagrodzenia pracowników	Wartości rzeczywiste wynagrodzenia pracowników
121	2010.1	3653,87	3382,46
122	2010.2	3599,28	3281,50
123	2010.3	3675,06	3647,88
124	2010.4	3814,85	3740,27
125	2010.5	3799,29	3600,39
126	2010.6	3829,79	3635,95
127	2010.7	3913,38	3740,96
128	2010.8	3929,83	3854,50
129	2010.9	3907,24	3777,14

Analiza błędów ex ante: Dobroć dopasowania modelu analizowano dla 2009 r. Wyniki obliczeń zestawiono w Tabeli 5.

Tabela 5. Ex ante. Wartości błędów ME, MAE, MPE i MAPE oraz współczynników Theila prognozy $R1(\nu3;\nu2,\nu4)$.

Table 5. Ex ante. Error values ME, MAE, MPE and MAPE and Theil coefficient estimates $R1(\nu3;\nu2,\nu4)$.

Lp	Nazwa i opis	ME	MAE	MPE[%]	MAPE[%]	
1	$R1(\nu3;\nu2,\nu4)$	-208	208	-5,8	5,8	
Lp	Nazwa i opis	I^2	I	I_1^2	I_2^2	I_3^2
1	$R1(\nu3;\nu2,\nu4)$	0,0042	0,0652	0,0033	0,0001	0,0008

Dopasowanie modelu do wartości rzeczywistych dopuszczalne. Wartość MAPE wynosi 5,8% a wartość I jest równa 0,0652.

Analiza błędów ex post: Dobroć predykcji modelu analizowano dla danych uzyskanych w styczniu 2011 r. a dotyczących okresu od stycznia do października 2010 r. Wyniki obliczeń zestawiono w Tabeli 6. Prognozując produkcję budowlano montażową na rok 2010 tych danych w rzeczywistości nie było. Jest to sprawdzenie ex post czyli już po postawieniu prognozy.

Tabela 6. Ex post. Wartości błędów ME, MAE, MPE i MAPE oraz współczynników Theila prognozy $R1(\nu3;\nu2,\nu4)$.

Table 6. Ex post. Error values ME, MAE, MPE and MAPE and Theil coefficient estimates $R1(\nu3;\nu2,\nu4)$.

Lp	Nazwa i opis	ME	MAE	MPE[%]	MAPE	
1	$R1(\nu3;\nu2,\nu4)$	-162	162	-4,59	4,59	
Lp	Nazwa i opis	I^2	I	I_1^2	I_2^2	I_3^2
1	$R1(\nu3;\nu2,\nu4)$	0,0026	0,0511	0,0026	0,0004	0,0005

Wniosek: Błąd MAPE prognozy jest mniejszy niż 5%, model mógłby być dopasowywany celem zmniejszenia błędów ex ante, jednak wartości reszt dla okresów prognozowanych nie będą znane, zatem nie wpłynie to istotnie na wynik prognozy.

MODEL $R2(\nu3;\nu2,\nu5,\nu6,\nu15,\nu16)$

Zastosowano metodę regresji wielorakiej wstecznej. Otrzymano równanie regresyjne w postaci (2), podsumowanie regresji zmiennej zależnej $\nu3$ zamiesz-

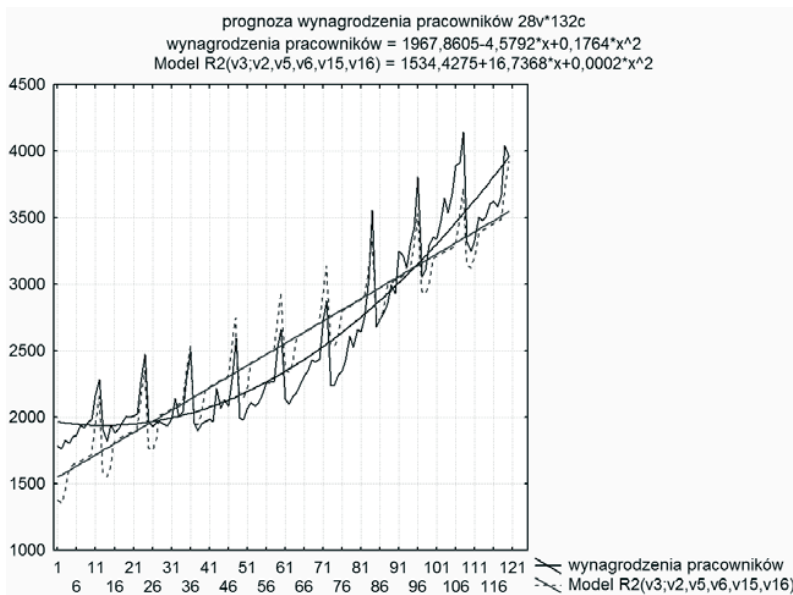
czono w Tabeli 7, wykres liniowy zależności zmiennej zależnej v_3 i prognozy $R_2(v_3;v_2,v_5,v_6,v_{15},v_{16})$ przedstawiono na Rys. 4.

Tabela 7. Podsumowanie regresji model $R_2(v_3;v_2,v_5,v_6,v_{15},v_{16})$.

Table 7. Summary of regression model $R_2(v_3;v_2,v_5,v_6,v_{15},v_{16})$.

Podsumowanie regresji zmiennej zależnej: wynagrodzenia pracowników (prognoza wynagrodzenia pracowników)						
$N = 120$						
$R = ,93360809$ $R^2 = ,87162407$ Skoryg. $R_2 = ,86599355$						
$F(5,114) = 154,80$ $p < 0,0000$ Błąd std. estymacji: 237,67						
	b^*	Bł. std. z b^*	b	Bł. std. z b	$t(114)$	p
W. wolny			1558,479	43,79160	35,58852	0,000000
t	0,876511	0,033687	16,359	0,62875	26,01899	0,000000
styczeń	-0,124960	0,042490	-197,895	67,28980	-2,94093	0,003964
luty	-0,148993	0,042474	-235,954	67,26453	-3,50785	0,000647
listopad	0,127875	0,042497	202,511	67,30154	3,00901	0,003227
grudzień	0,253748	0,042519	401,851	67,33502	5,96794	0,000000

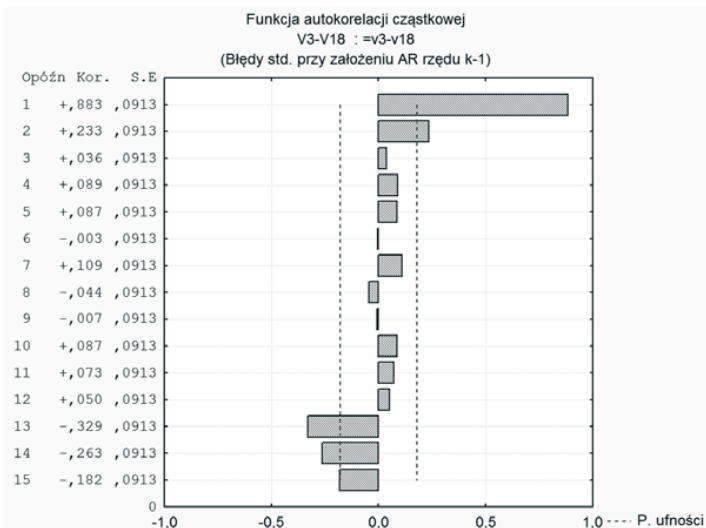
$$v_3 = 1558,479 + 16,359v_2 - 197,895v_5 - 235,954v_6 + 202,511v_{15} + 401,851v_{16} \quad (2)$$



Rys. 4. Wykres zależności średniego wynagrodzenia w sektorze budowlanym w województwie dolnośląskim w badanych okresach od stycznia 2000 do grudnia 2009 oraz prognozy uzyskanej metodą regresji $R_2(v_3;v_2,v_5,v_6,v_{15},v_{16})$.

Fig. 4. A plot of the average salary in the construction sector in Lower Silesia in the examined period from January 2000 to December 2009 and the forecasts obtained by regression $R_2(v_3;v_2,v_5,v_6,v_{15},v_{16})$.

Wniosek: Wartość R^2 skorygowana wynosi 0,8660 a p wartość maksymalnie 0,0032. Zatem model R_2 również nie jest najlepszy. Celem sprawdzenia wykonano wykres autokorelacji reszt Rys. 5.



Rys. 5. Funkcja autokorelacji cząstkowej reszt modelu $R2(v3;v2,v5,v6,v15,v16)$.

Fig. 5. Partial autocorrelation function of the model $R2(v3;v2,v5,v6,v15,v16)$ residuals.

Tabela 8. Prognoza modelu $R2(v3;v2,v5,v6,v15,v16)$ wynagrodzenia pracowników od stycznia do października 2010 r. oraz uzyskane w 2011 wartości rzeczywiste.

Table 8. Prediction model $R2(v3;v2,v5,v6,v15,v16)$, staff salaries from January to October 2010, and received in 2011 in real terms.

T	Data	Prognoza wynagrodzenia pracowników	Wartości rzeczywiste wynagrodzenia pracowników
121	2010.1	3340,02	3382,46
122	2010.2	3318,32	3281,50
123	2010.3	3400,12	3647,88
124	2010.4	3587,00	3740,27
125	2010.5	3603,35	3600,39
126	2010.6	3619,71	3635,95
127	2010.7	3636,07	3740,96
128	2010.8	3652,43	3854,50
129	2010.9	3668,79	3777,14

Analiza błędów ex ante: Dobroć dopasowania modelu analizowano dla 2009 r. Wyniki obliczeń zestawiono w Tabeli 9.

Tabela 9. Ex ante. Wartości błędów ME, MAE, MPE i MAPE oraz współczynników Theila prognozy $R2(v3;v2,v5,v6,v15,v16)$.

Table 9. Ex ante. Error values ME, MAE, MPE and MAPE and Theil coefficient estimates $R2(v3;v2,v5,v6,v15,v16)$.

Lp	Nazwa i opis	ME	MAE	MPE[%]	MAPE[%]
1	$R2(v3;v2,v5,v6,v15,v16)$	140	140	3,9	3,9

Lp	Nazwa i opis	I^2	I	I_1^2	I_2^2	I_3^2
1	$R2(v3;v2,v5,v6,v15,v16)$	0,0019	0,0442	0,0015	0,0000	0,0004

Dopasowanie modelu do wartości rzeczywistych dopuszczalne. Wartość MAPE wynosi 3,9% a wartość I jest równa 0,0652.

Analiza błędów ex post: Dobroć predykcji modelu analizowano dla danych uzyskanych w styczniu 2011 r. a dotyczących okresu od stycznia do października

2010 r. Wyniki obliczeń zestawiono w Tabeli 10. Prognozując produkcję budowlano-montażową na rok 2010 tych danych w rzeczywistości nie było. Jest to sprawdzenie ex post czyli już po postawieniu prognozy.

Tabela 10. Ex post. Wartości błędów ME, MAE, MPE i MAPE oraz współczynników Theila prognozy $R2(v3;v2,v5,v6,v15,v16)$.

Table 10. Ex post. Error values ME, MAE, MPE and MAPE and Theil coefficient estimates $R2(v3;v2,v5,v6,v15,v16)$.

Lp	Nazwa i opis	ME	MAE	MPE[%]	MAPE[%]	
1	$R2(v3;v2,v5,v6,v15,v16)$	92	101,64	2,48	2,74	
Lp	Nazwa i opis	I^2	I	I_1^2	I_2^2	I_3^2
1	$R2(v3;v2,v5,v6,v15,v16)$	0,0012	0,0357	0,0009	0,0002	0,0007

Wniosek: Błąd MAPE prognozy jest mniejszy niż 3%, model lepszy niż R1.

MODEL $R3(v2;v1,v3,v8,v14,v15,v17,v18,v23,v42,v45,v47,v51)$

Ze względu na istnienie autokorelacji reszt w modelach R1 i R2 należy poszukiwać lepszego modelu. Proponuje się wprowadzenie modelu R3, który będzie posiadał więcej zmiennych zależnych. Zmiennymi zależnymi będą: t i t^2 , 12 zmiennych sezonowych od styczeń do grudzień, 36 zmiennych opóźnionych, wynagrodzenia opóźnione o 1 okres, 2 okresy,....., 36 okresów.

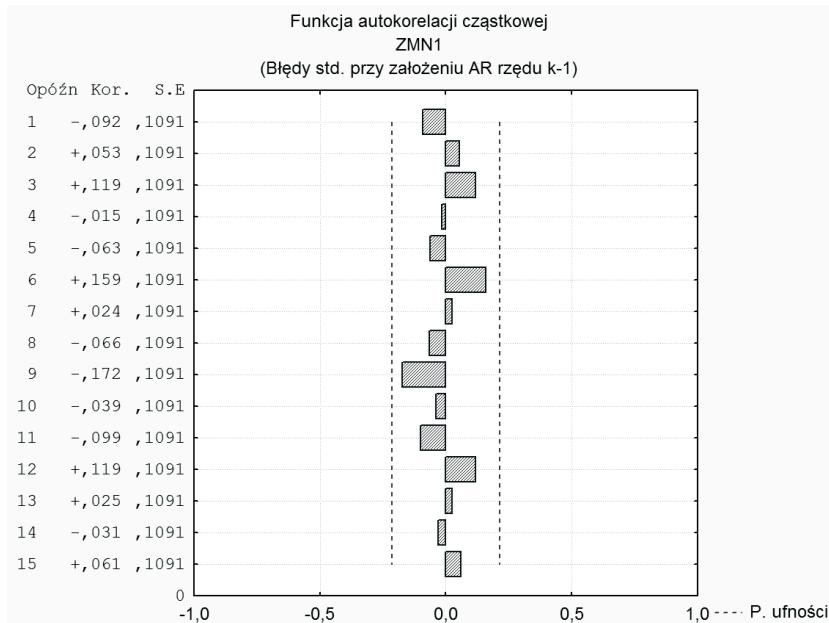
Tabela 11. Podsumowanie regresji model $R3(v2,v1,v3,v8,v14,v15,v17,v18,v23,v42,v45,v47,v51)$.

Table 11. Summary of regression model $R3(v2,v1,v3,v8,v14,v15,v17,v18,v23,v42,v45,v47,v51)$.

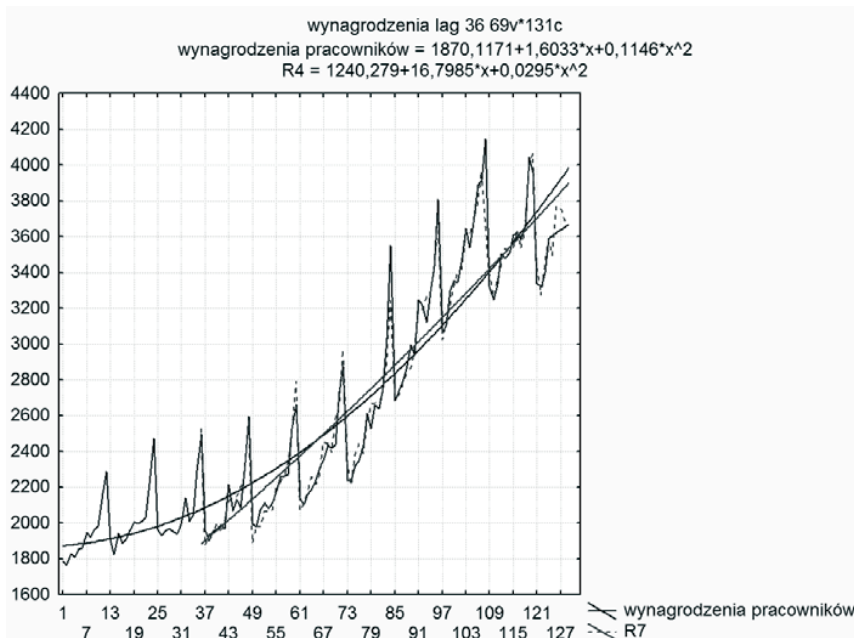
Podsumowanie regresji zmiennej zależnej: wynagrodzenia pracowników (wynagrodzenia lag 36)						
N = 84	R = ,99362697 R ² = ,98729455 Skoryg. R2 = ,98514715					
	F(12,71) = 459,76 p < 0,0000 Błąd std. estymacji: 77,645					
	b^*	Bł. std. z b^*	b	Bł. std. z b	$t(71)$	p
W. wolny			1464,956	456,5466	3,20878	0,002002
t	-0,326051	0,160135	-8,516	4,1825	-2,03611	0,045471
styczeń	-0,173827	0,027749	-269,650	43,0462	-6,26419	0,000000
czerwiec	-0,059187	0,021747	-91,814	33,7352	-2,72160	0,008166
grudzień	0,270948	0,018301	420,310	28,3889	14,80543	0,000000
t^2	0,754775	0,249920	0,124	0,0412	3,02006	0,003510
lag 1	0,381266	0,070233	0,389	0,0716	5,42860	0,000001
lag2	0,201347	0,070607	0,210	0,0737	2,85165	0,005692
lag7	0,227346	0,068520	0,244	0,0736	3,31798	0,001432
lag26	-0,285253	0,051032	-0,469	0,0840	-5,58969	0,000000
lag29	-0,102045	0,032002	-0,185	0,0580	-3,18868	0,002127
lag31	-0,129086	0,048115	-0,249	0,0927	-2,68285	0,009075
lag35	0,204024	0,024931	0,416	0,0508	8,18366	0,000000

$$v2=1464,956-8,516*v1-269,650*v3-91,814*v8+420,310*v14+0,124*v15+0,389*v17+0,210*v18+0,244*v23-0,469*v42+-0,185*v45-0,249*v47+0,416*v51 \quad (3)$$

Obliczono autokorelację reszt, kologram zamieszczono poniżej.



Rys. 6. Funkcja autokorelacji cząstkowej reszt modelu $R3(\nu_3;\nu_2,\nu_4,\nu_5,\nu_6,\nu_{15},\nu_{16})$.
 Fig. 6. Partial autocorrelation function of the model $R3(\nu_3;\nu_2,\nu_4,\nu_5,\nu_6,\nu_{15},\nu_{16})$ residuals.



Rys. 7. Wykres zależności średniego wynagrodzenia w sektorze budowlanym w województwie dolnośląskim w badanych okresach od stycznia 2000 do grudnia 2009 oraz prognozy uzyskanej metodą regresji $R3(\nu_3;\nu_2,\nu_4,\nu_5,\nu_6,\nu_{15},\nu_{16})$.
 Fig. 7. A plot of the average salary in the construction sector in Lower Silesia in the examined period from January 2000 to December 2009 and the forecasts obtained by regression $R3(\nu_3;\nu_2,\nu_4,\nu_5,\nu_6,\nu_{15},\nu_{16})$.

Tabela 12. Prognoza modelu R3 wynagrodzenia pracowników od stycznia do października 2010 r. oraz uzyskane w 2011 wartości rzeczywiste.

Table 12. Prediction model R3 salaries of employees from January to October 2010, and received in 2011 in real terms.

T	Data	Prognoza wynagrodzenia pracowników	Wartości rzeczywiste wynagrodzenia pracowników
121	2010.1	3427,19500	3382,46
122	2010.2	3270,96078	3281,50
123	2010.3	3455,14968	3647,88
124	2010.4	3585,02388	3740,27
125	2010.5	3493,3352	3600,39
126	2010.6	3767,41615	3635,95
127	2010.7	3759,58269	3740,96
128	2010.8	3690,34421	3854,50
129	2010.9	3688,33205	3777,14

Analiza błędów ex ante: Dobroć dopasowania modelu analizowano dla 2009 roku. Wyniki obliczeń zestawiono w Tabeli 13. Dopasowanie modelu do wartości rzeczywistych dopuszczalne. Wartość MAPE wynosi 0,87% a wartość I jest równa 0,0002.

Analiza błędów ex post: Dobroć predykcji modelu analizowano dla danych uzyskanych w styczniu 2011 roku a dotyczących okresu od stycznia do października 2010 r. Wyniki obliczeń zestawiono w Tabeli 13. Prognozując produkcję budowlaną na rok 2010 tych danych w rzeczywistości nie było. Jest to sprawdzenie ex post czyli już po postawieniu prognozy.

Tabela 13. Ex ante i ex post. Wartości błędów ME, MAE, MPE i MAPE oraz współczynników Theila prognozy R3.

Table 13. Ex ante and ex post. Error values ME, MAE, MPE and MAPE and Theil coefficient estimates R3.

Lp	Nazwa i opis ex ante	ME	MAE	MPE[%]	MAPE[%]	
1	R3	-2,26108	30,20258	-0,0869	0,87023	
Lp	Nazwa i opis ex ante	I^2	I	I_1^2	I_2^2	I_3^2
1	R3	0,000193579	0,01391326	0,000001	0,0000	0,0002
Lp	Nazwa i opis ex post	ME	MAE	MPE[%]	MAPE[%]	
1	R3	-25,9608	52,51916	-0,72844	1,484341	
Lp	Nazwa i opis ex post	I^2	I	I_1^2	I_2^2	I_3^2
1	R3	0,000576633	0,02401318	0,000096	0,0001	0,0005

Wniosek: Błąd MAPE prognozy jest najmniejszy spośród proponowanych modeli regresyjnych i wynosi 1,48% model R3 zostanie przyjęty do dalszej prognozy produkcji budowlano montażowej.

B. Metoda SARIMA

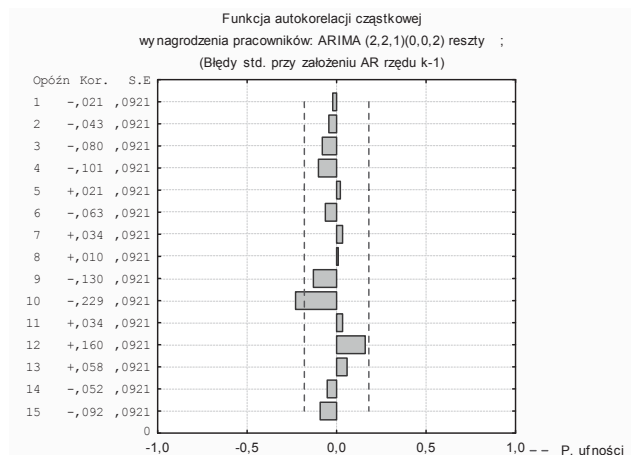
Model SARIMA1 (2,2,1)(0,0,2)

Szereg czasowy średnie wynagrodzenia pracowników w sektorze budowlanym w województwie dolnośląskim w latach 2000-2009 poddano analizie SARIMA. Celem jest wyznaczenie prognozy na okresy od stycznia do października 2010 r.

Tabela 14. Podsumowanie model SARIMA1 (2,2,1)(0,0,2).

Table 14. Summary SARIMA1 model (2,2,1) (0,0,2).

Parametr	Dane: wynagrodzenia pracowników (prognoza wynagrodzenia pracowników) Przekształcenia: 2*D(1) Model: (2,2,1)(0,0,2) Opóź. sezon.: 12 Resztowy MS = 18029,					
	Parametr	Asympt. Błąd std	Asympt. t (113)	p	Dolna gr 95%p.ufn	Górna gr 95%p.ufn
p(1)	-0,317111	0,094392	-3,35952	0,001065	-0,504117	-0,130104
p(2)	-0,199895	0,099611	-2,00675	0,047164	-0,397243	-0,002547
q(1)	0,988259	0,013031	75,84033	0,000000	0,962443	1,014076
Qs(1)	-0,774969	0,092242	-8,40144	0,000000	-0,957718	-0,592220
Qs(2)	-0,656868	0,088138	-7,45275	0,000000	-0,831485	-0,482252



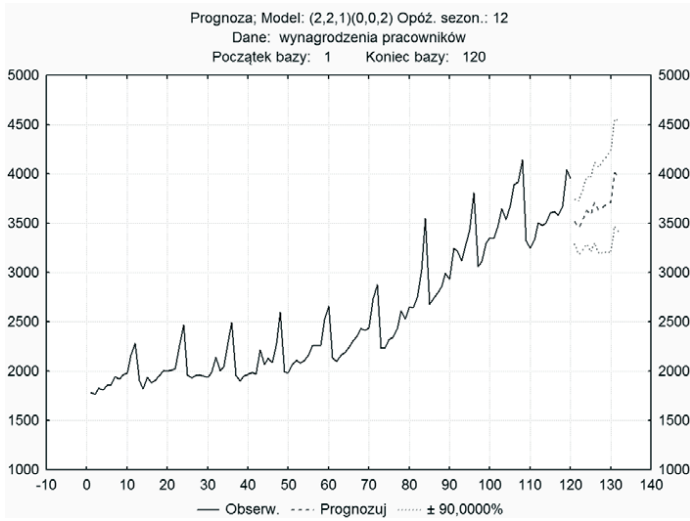
Rys. 8. Funkcja autokorelacji cząstkowej reszt modelu SARIMA1 (2,2,1)(0,0,2).

Fig. 8. Partial autocorrelation function of the model SARIMA1 (2,2,1) (0,0,2) residuals.

Tabela 15. Funkcja autokorelacji (prognoza wynagrodzenia pracowników) modelu SARIMA1 (2,2,1)(0,0,2) oraz wartości Q Boxa i Ljunga do oceny istotności korelacji.

Table 15. Autocorrelation function (workers salary forecast) SARIMA1 (2,2,1) (0,0,2) model and Q values of Ljung Box and evaluating the significance of the correlation.

Opóźn.	Funkcja autokorelacji (prognoza wynagrodzenia pracowników) wynagrodzenia pracowników: ARIMA (2,2,1)(0,0,2) reszty ; (Błędy standardowe to oceny białego szumu)			
	Auto-kor.	Błąd std	Q Boxa i Ljunga	p
1	-0,021155	0,090899	0,05416	0,815973
2	-0,042512	0,090510	0,27477	0,871635
3	-0,077796	0,090119	1,01997	0,796419
4	-0,094483	0,089727	2,12881	0,712082
5	0,033304	0,089332	2,26780	0,810981
6	-0,048152	0,088936	2,56094	0,861580
7	0,049360	0,088538	2,87174	0,896600
8	0,017658	0,088138	2,91188	0,939757
9	-0,128146	0,087737	5,04514	0,830342
10	-0,209376	0,087333	10,79285	0,373921
11	0,048290	0,086928	11,10145	0,434834
12	0,186305	0,086521	15,73810	0,203596
13	0,088719	0,086112	16,79957	0,208702
14	-0,031557	0,085701	16,93515	0,259720
15	-0,103136	0,085288	18,39750	0,242422



Rys. 9. Wykres danych i prognozy SARIMA1 (2,2,1)(0,0,2) wynagrodzenia pracowników w sektorze budowlanym w województwie dolnośląskim w latach 2000-2010.

Fig. 9. Data and forecasts SARIMA1 (2,2,1) (0,0,2) graph the workers salary in the construction sector in Lower Silesia in 2000-2010.

Tabela 16. Podsumowanie model SARIMA1 (2,2,1)(0,0,2).

Table 16. Summary SARIMA1 (2,2,1) (0,0,2) model.

Nr obserwacji	Proгноза; Модел: (2,2,1)(0,0,2) Опóз. сезон.: 12 (prognoza wynagrodzenia pracowników) Dane: wynagrodzenia pracowników Początek bazy: 1 Koniec bazy: 120			
	Proгноза	Dolne 90,0000%	Górne 90,0000%	Błąd standardowy
121	3515,477	3292,794	3738,160	134,2717
122	3457,391	3186,256	3728,526	163,4871
123	3525,428	3222,834	3828,023	182,4562
124	3628,285	3287,384	3969,185	205,5539
125	3586,900	3211,794	3962,007	226,1790
126	3710,867	3305,095	4116,639	244,6691
127	3629,622	3194,168	4065,075	262,5663
128	3660,877	3196,979	4124,775	279,7177
129	3696,812	3205,674	4187,950	296,1426

Tabela 17. Prognoza modelu SARIMA1 (2,2,1)(0,0,2) wynagrodzenia pracowników od stycznia do października 2010 r. oraz uzyskane w 2011 wartości rzeczywiste.

Table 17. SARIMA1 (2,2,1) (0,0,2) prediction model, staff salaries from January to October 2010 and received in 2011 in real terms.

T	Data	Proгноза wynagrodzenia pracowników	Wartości rzeczywiste wynagrodzenia pracowników
121	2010.1	3515,477	3382,46
122	2010.2	3457,391	3281,50
123	2010.3	3525,428	3647,88
124	2010.4	3628,285	3740,27
125	2010.5	3586,900	3600,39
126	2010.6	3710,867	3635,95
127	2010.7	3629,622	3740,96
128	2010.8	3660,877	3854,50
129	2010.9	3696,812	3777,14

Analiza błędów ex post: Dobroć predykcji modeli analizowano dla danych uzyskanych w styczniu 2011 roku a dotyczących okresu od stycznia do października 2010 roku. Wyniki obliczeń zestawiono w Tabeli 18. Prognozując produkcję budowlano montażową na rok 2010 tych danych w rzeczywistości nie było. Jest to sprawdzenie ex post czyli już po postawieniu prognozy.

Tabela 18. Ex post. Wartości błędów ME, MAE, MPE i MAPE oraz współczynników Theila prognoz SARIMA1(2,2,1)(0,0,2).

Table 18. Ex post. Error values ME, MAE, MPE and MAPE and Theil coefficients SARIMA1 (2,2,1) (0,0,2) predictions.

Lp	Nazwa i opis	ME	MAE	MPE[%]	MAPE[%]	
1	SARIMA(2,2,1)(0,0,2)	27,71	113,00	0,61	3,13	
Lp	Nazwa i opis	P^2	I	I_1^2	I_2^2	I_3^2
1	SARIMA(2,2,1)(0,0,2)	0,0011	0,0341	0,0000	0,0010	0,0006

Wartość MAPE dla modelu SARIMA1(2,2,1)(0,0,2) wynosi 3,13% a wartość I jest równa 0,0341, zatem osiągnięto zdolność predykcji modelu dobrą. Autokorelacja reszt występuje jednak nie jest istotna statystycznie (wartości Q Boxa Ljunga).

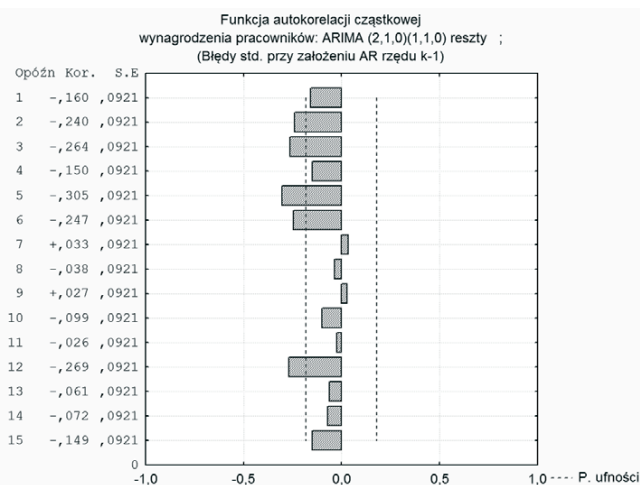
Model SARIMA2 (2,1,0)(1,1,0)

Szereg czasowy średnie wynagrodzenia pracowników w sektorze budowlanym w województwie dolnośląskim w latach 2000-2009 poddano analizie SARIMA. Celem jest wyznaczenie prognozy na okresy od stycznia do października 2010 r.

Tabela 19. Podsumowanie model SARIMA2 (2,1,0)(1,1,0).

Table 19. Summary SARIMA2 (2,1,0) (1,1,0) model.

Parametr	Dane: wynagrodzenia pracowników (prognoza wynagrodzenia pracowników) Przekształcenia: D(1),D(1) Model:(2,1,0)(1,1,0) Opóź. sezon.: 12 Resztowy MS= 17357,					
	Parametr	Asympt. Błąd std	Asympt. t (115)	p	Dolna gr 95%p.ufn	Górna gr 95%p.ufn
p(1)	-1,02711	0,080672	-12,7319	0,000000	-1,18690	-0,867313
p(2)	-0,57928	0,082395	-7,0305	0,000000	-0,74249	-0,416071
Ps(1)	0,98157	0,047090	20,8445	0,000000	0,88829	1,074845



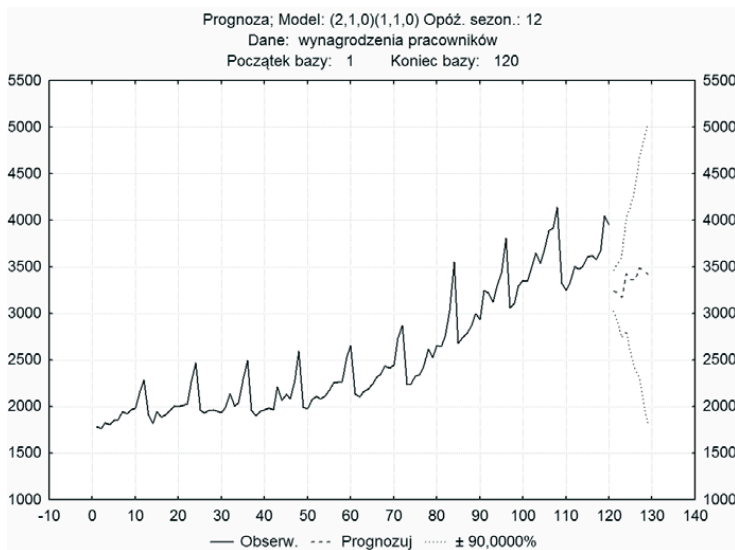
Rys. 10. Funkcja autokorelacji cząstkowej reszt modelu SARIMA2 (2,1,0)(1,1,0).

Fig. 10. Partial autocorrelation function of the model SARIMA2 (2,1,0) (1,1,0) residuals.

Tabela 20. Funkcja autokorelacji (prognoza wynagrodzenia pracowników) modelu SARIMA2 (2,1,0)(1,1,0) oraz wartości Q Boxa i Ljunga do oceny istotności korelacji.

Table 20. Autocorrelation function (workers salary forecast) SARIMA2 (2,1,0) (1,1,0) model and Q values of Ljung Box and evaluating the significance of the correlation.

Opóźn.	Funkcja autokorelacji (prognoza wynagrodzenia pracowników) wynagrodzenia pracowników: ARIMA (2,1,0)(1,1,0) reszty ; (Błędy standardowe to oceny białego szumu)			
	Auto-kor.	Błąd std	Q Boxa i Ljunga	p
1	-0,159774	0,090899	3,08951	0,078808
2	-0,208237	0,090510	8,38272	0,015134
3	-0,162558	0,090119	11,63644	0,008746
4	0,017138	0,089727	11,67292	0,019972
5	-0,118145	0,089332	13,42203	0,019746
6	0,011796	0,088936	13,43962	0,036589
7	0,244608	0,088538	21,07237	0,003672
8	0,023072	0,088138	21,14090	0,006795
9	-0,037587	0,087737	21,32443	0,011304
10	-0,154908	0,087333	24,47065	0,006459
11	0,022268	0,086928	24,53627	0,010674
12	-0,186840	0,086521	29,19959	0,003691
13	0,150030	0,086112	32,23511	0,002225
14	0,115627	0,085701	34,05544	0,002032
15	-0,000083	0,085288	34,05544	0,003357



Rys. 11. Wykres danych i prognozy SARIMA2 (2,1,0)(1,1,0) wynagrodzenia pracowników w sektorze budowlanym w województwie dolnośląskim w latach 2000-2010.

Fig. 11. Data and forecasts SARIMA1 (2,1,0) (1,1,0) graph the workers salary in the construction sector in Lower Silesia in 2000-2010.

Tabela 21. Podsumowanie model SARIMA2 (2,1,0)(1,1,0).

Table 21. Summary SARIMA2 (2,1,0) (1,1,0) model.

Nr obserwacji	Prognoza; Model: (2,1,0)(1,1,0) Opóz. sezon.: 12 (prognoza wynagrodzenia pracowników) Dane: wynagrodzenia pracowników Początek bazy: 1 Koniec bazy: 120			
	Prognoza	Dolne 90,0000%	Górne 90,0000%	Błąd standar- dowy
121	3242,507	3024,045	3460,970	131,7458
122	3225,306	2920,512	3530,099	183,8087
123	3173,533	2738,412	3608,655	262,4041
124	3425,248	2812,253	4038,242	369,6719
125	3367,569	2595,772	4139,366	465,4393
126	3355,778	2390,941	4320,614	581,8537
127	3489,158	2318,544	4659,771	705,9495
128	3464,538	2084,200	4844,876	832,4256
129	3420,989	1809,864	5032,115	971,6043

Tabela 22. Prognoza modelu SARIMA2 (2,1,0)(1,1,0) wynagrodzenia pracowników od stycznia do października 2010 r. oraz uzyskane w 2011 wartości rzeczywiste.

Table 22. SARIMA2 (2,1,0) (1,1,0) prediction model, staff salaries from January to October 2010, and received in 2011 in real terms.

T	Data	Prognoza wynagrodzenia pracowników	Wartości rzeczywiste wynagrodzenia pracowników
121	2010.1	3242,507	3382,46
122	2010.2	3225,306	3281,50
123	2010.3	3173,533	3647,88
124	2010.4	3425,248	3740,27
125	2010.5	3367,569	3600,39
126	2010.6	3355,778	3635,95
127	2010.7	3489,158	3740,96
128	2010.8	3464,538	3854,50
129	2010.9	3420,989	3777,14

Analiza błędów ex post: Dobroć predykcji modeli analizowano dla danych uzyskanych w styczniu 2011 roku a dotyczących okresu od stycznia do października 2010 r. Wyniki obliczeń zestawiono w Tabeli 23. Prognozując produkcję budowlano montażową na rok 2010 tych danych w rzeczywistości nie było. Jest to sprawdzenie ex post czyli już po postawieniu prognozy.

Tabela 23. Ex post. Wartości błędów ME, MAE, MPE i MAPE oraz współczynników Theila prognoz SARIMA2 (2,1,0)(1,1,0).

Table 23. Ex post. Error values ME, MAE, MPE and MAPE and Theil coefficients SARIMA2 (2,1,0) (1,1,0) predictions.

Lp	Nazwa i opis	ME	MAE	MPE[%]	MAPE[%]	
1	SARIMA(2,2,1)(0,0,2)	277,3804	277,3804	7,525015	7,525015	
Lp	Nazwa i opis	I^2	I	I_1^2	I_2^2	I_3^2
1	SARIMA(2,2,1)(0,0,2)	6,9E-03	7,8E-03	5,6E-04	1,1E-03	8,3E-02

Wartość MAPE dla modelu SARIMA2 (2,1,0)(1,1,0) wynosi 7,52% a wartość I jest równa 7,8E-03, zatem osiągnięto zdolność predykcji modelu dopuszczalną. Istniejące autokorelacje reszt są istotne statystycznie.

Podsumowanie: Do prognozowania produkcji budowlano montażowej przyjęto wartości szeregu z prognozy R^2 zestawione w Tabeli 24.

Tabela 24. Prognoza wynagrodzenia pracowników od stycznia do października 2010 r. oraz uzyskane w 2011 wartości rzeczywiste.

Table 24. Salaries of employees prediction from January to October 2010, and received in 2011 the actual values.

T	Data	Prognoza wynagrodzenia pracowników	Wartości rzeczywiste wynagrodzenia pracowników
121	2010.1	3340,02	3382,46
122	2010.2	3318,32	3281,50
123	2010.3	3400,12	3647,88
124	2010.4	3587,00	3740,27
125	2010.5	3603,35	3600,39
126	2010.6	3619,71	3635,95
127	2010.7	3636,07	3740,96
128	2010.8	3652,43	3854,50
129	2010.9	3668,79	3777,14

3. Wnioski

Po przeprowadzeniu obliczeń, testowania i weryfikacji jako najlepszy uzyskany model uznano równanie regresyjne oznaczone jako R3. Spełnia ono warunki maksymalizacji skorygowanego współczynnika regresji, minimalizacji błędów oraz braku autokorelacji cząstkowej reszt. Wyznaczone równanie posiada 12 predyktorów wyeliminowanych spośród 53 zmiennych niezależnych. Nie jest więc równaniem najprostszym lecz jedynym spełniającym wszystkie wymagania. Wyznaczone równanie regresyjne będzie używane do predykcji produkcji budowlano-montażowej w województwie dolnośląskim.

Literatura

- [1] Box, G.E.P., Pierce, D.A., *Distribution of residual autocorrelations in autoregressive-integrated moving average time series models*, Journal of the American Statistical Association 65 (1970) 1509-26.
- [2] Christodoulos Ch., Michalakelis Ch., Varoutas D., *Forecasting with limited data: Combining ARIMA and diffusion models*, Technological Forecasting and Social Change 77 (2010) 558-565.
- [3] Cieślak M., *Prognozowanie gospodarcze : metody i zastosowanie*, Wydaw. Naukowe PWN, Warszawa 2001.
- [4] Dickey D.A., Fuller W.A., *Likelihood ratio statistics for autoregressive time series with a unit root*, Econometrica 49(4) (1981) 1957-72.
- [5] Ediger V.E., Akar S., *ARIMA forecasting of primary energy demand by fuel in Turkey*, Energy Policy 35(3) (2007) 1701-1708.
- [6] Faruk D.O., *A hybrid neural network and ARIMA model for water quality time series prediction*, Engineering Applications of Artificial Intelligence 23(4) (2010) 586-594.
- [7] Gilbert K.C., Chatpattananan V., *An ARIMA supply chain model with a generalized ordering policy*, Journal of Modelling in Management 1(1) (2006).
- [8] Gilbert K.C., *An ARIMA supply chain model*, Management Science 51(2) (2005) 305-10.
- [9] Kot S., Jakubowski J., Sokołowski A., *Statystyka*, Difin, Warszawa 2007.
- [10] Stanisz A., *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny*, T 1. StatSoft Polska Sp. z o.o., Kraków 2006.

Prediction of construction and assembly production in the province of Lower Silesia. Part I

Magdalena Rogalska

*Department of Production Engineering and Construction Management,
Faculty of Civil Engineering and Architecture, Lublin University of Technology,
e-mail: m.rogalska@tlen.pl*

Abstract: The article is the first part of the series „Prediction construction and assembly production in Lower Silesia.” It was assumed that salary of employees will be one of the independent variables to determine the volume of production. Salaries of employees was predicted, using multiple regression and autoregressive moving average SARIMA methods. An analysis of the results was carried out. The errors ME, MAE, MPE, MAPE and Theil coefficients I, I2, I12, I22, I32 were calculated. Multiple regression equation with 12 predictors was set. Predictors were selected from among the 53 independent variables. Forecasted data were obtained for construction and assembly production prediction.

Keywords: prediction, multiple regression, SARIMA, salaries of employees.

