

Marcin Błażejowski

Wyższa Szkoła Bankowa w Toruniu

ANALIZA PORÓWNAWCZA AUTOMATYCZNYCH PROCEDUR MODELOWANIA I PROGNOZOWANIA

Streszczenie: Problematyka automatycznych procedur modelowania i prognozowania wpisuje się w najnowsze nurty dynamicznego modelowania ekonometrycznego. Badania w tym zakresie skupiają się zarówno nad problemami teoretycznymi algorytmów, jak i nad ich implementacją. W artykule porównane zostaną wyniki analiz empirycznych uzyskanych na podstawie algorytmu Autometrics dostępnego w środowisku PcGive i bazującego na podejściu „od ogólnego do szczególnego” oraz algorytmu CongruentSpecification dostępnego w programie GRETL i bazującego na teorii dynamicznych modeli zgodnych oraz algorytmów automatycznej specyfikacji modeli struktury. Ponadto porównane zostaną także wyniki uzyskane na podstawie popularnych automatycznych procedur wyboru modelu struktury procesu.

Słowa kluczowe: Congruent Specification, Autometrics, procedury automatycznego wyboru modelu, niepewność modelu, prognozy kombinowane.

1. Wstęp

Zagadnienie wyboru modelu (*model selection, model reduction*) poruszane jest w wielu opracowaniach dotyczących modelowania i prognozowania ekonometrycznego i wpisuje się w szerszy kontekst problematyki określanej jako niepewność modelu (*model uncertainty*). Najogólniej rzecz można ująć w ten sposób, iż dopuszcza się, że proces generujący dane (*Data Generating Process – GDP*) może nie zostać w 100% odtworzony przez jeden model aproksymacyjny, lub inaczej, istnieje wiele konkurencyjnych modeli aproksymacyjnych dla jednego (i nieznanego) procesu generującego. Konkurencyjne modele aproksymacyjne mogą powstawać na drodze innych strategii redukcji modelu ogólnego lub innych założeń wstępnych leżących u podstaw poszczególnych podejść do modelowania (modele przyczynowe, modele struktury, transformacje zmiennych itd.). Mając powyższe na uwadze, należy zadać pytanie o wybór modelu prognostycznego spośród konkurencyjnych modeli aproksymacyjnych. Bardzo często proponuje się w takiej sytuacji skonstruowanie prognoz ważonych (kombinowanych). W niniejszym artykule przedstawiona zostanie procedura oceny skuteczności algorytmów prognozowania Congruent Specification dla programu GRETL oraz Autometrics dla środowiska PcGIVE dla modeli przyczynowo-skutkowych oraz X-12-ARIMA i TRAMO/SEATS dla modeli opisu struktury

procesu. Ocena skuteczności zostanie przeprowadzona za pomocą mierników trafności prognoz: RSME, MAE oraz MAPE. Postawiono następującą hipotezę badawczą: „Algorytmy automatycznej budowy modeli przyczynowo-skutkowych i modeli struktury mogą być użytecznymi narzędziami w prognozowaniu ekonomicznych szeregów czasowych”.

2. Algorytmy automatycznej specyfikacji modelu

Algorytmy automatycznej specyfikacji modelu można podzielić na dwie główne kategorie:

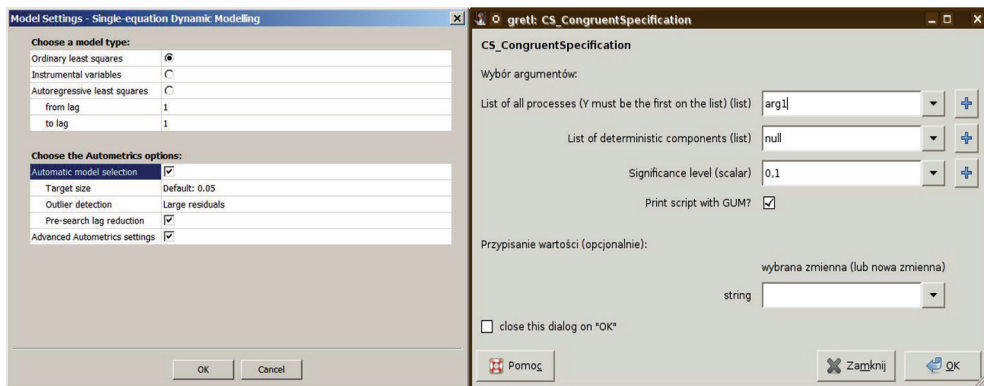
- 1) procedury specyfikacji modeli przyczynowo-skutkowych,
- 2) procedury specyfikacji modeli opisu struktury.

W niniejszym badaniu zostaną wykorzystane oba podejścia.

2.1. Procedury automatycznej specyfikacji modeli przyczynowych

W ostatnich latach nastąpił dynamiczny rozwój automatycznych procedur selekcji modelu. W dużej mierze jest to efekt artykułu [Hoover, Perez 1999], w którym autorzy wykazali, że systematyczne (algorytmiczne) wykorzystanie koncepcji modelowania „od ogólnego do szczególnego” prof. D.F. Hendry’ego prowadzi do uzyskania prawidłowych modeli aproksymacyjnych. W efekcie powstał moduł PcGets, a następnie Autometrics dostępny w środowisku PcGive. Moduł ten pozwala na pół-automatyczną specyfikację modelu według koncepcji modelowania „od ogólnego do szczególnego”, tzn. wskazania zmiennej niezależnej, zmiennych zależnych, oraz ustalenie wstępnego rzędu opóźnień (identycznego dla wszystkich szeregów). Następnie algorytm przystępuje do automatycznej redukcji i poszukiwania (lokalnego) modelu aproksymacyjnego, w tym rozszerzania startowej specyfikacji o struktury autoregresyjne. Cała ta procedura jest w 100% realizowana przez program i użytkownik nie ma w zasadzie wpływu na jej przebieg (poza wstępnymi parametrami sterującymi, takimi jak np. poziom istotności czy zestaw testów wykorzystywanych na etapie wyboru modelu). Szczegóły dotyczące procedury Autometrics można znaleźć w [Doornik 2009].

Drugim modulem, będącym wciąż w fazie rozwoju i testów, jest Congruent Specification dostępny dla programu Gretl, który automatyzuje specyfikację modelu według koncepcji modelowania zgodnego w sensie prof. Z. Zielińskiego. Procedura jako wynik podaje specyfikację tzw. modelu pełnego (*General Unrestricted Model* – GUM), tj. kompletną strukturę deterministyczną i autoregresyjną wszystkich procesów. Strategia redukcji modelu oraz wprowadzanie ewentualnych restrykcji w 100% zależy jednak od użytkownika. Szczegóły tego modułu opisane są w [Błażejowski, Kufel, Kufel 2009; 2010]. Rysunek 1 przedstawia przykładowe okno główne modułu Autometrics (po lewej stronie) oraz przykładowe okno główne modułu Congruent Specification (po prawej stronie).



Rys. 1. Okno główne modułów Autometrics i Congruent Specification

Źródło: opracowanie własne.

2.2. Procedury automatycznej specyfikacji modeli struktury

W badaniu wykorzystano także dwa bardzo popularne algorytmy modelowania i prognozowania szeregów czasowych:

- 1) X-12-ARIMA w wersji 0.3 Build 192,
- 2) TRAMO/SEATS w wersji beta (123) z dnia 3.11.2008 r.

Warto w tym miejscu wspomnieć, że wykorzystana wersja TRAMO/SEATS jest ostatnim wydaniem dostępnym w otwartej (źródłowej) formie, które może być wykorzystywane we wszystkich środowiskach operacyjnych (MS Windows, Mac OS X, GNU/Linux). Jednocześnie wykorzystana wersja X-12-ARIMA w momencie przygotowywania tekstu była najnowszą, od 16 sierpnia 2012 r. na stronach US Census Bureau dostępna jest już jednak X-13ARIMA-SEATS w wersji 1.0 Build 149. Jest to pierwsze oficjalne wydanie X-12-ARIMA, które w fazie wyboru modelu bazuje na kodach źródłowych SEATS. Wykorzystana w badaniu wersja 0.3 Build 192 także w fazie wyboru modelu bazuje już na algorytmach SEATS, jednak była to wersja rozwojowa, stąd twórcy przejściowo pozostali przy starej nazwie.

3. Procedura badawcza

Ponieważ ideą badania jest określenie trafności prognoz uzyskanych na podstawie poszczególnych algorytmów, postanowiono przyjąć następującą wieloetapową procedurę oceny ich trafności.

Etap 1. Skonstruowanie prostej prognozy ważonej (kombinowanej).

Na tym etapie zostanie skonstruowana prognoza ważona z jednakowymi wagami dla wszystkich wygenerowanych na podstawie automatycznych procedur specyfikacji modelu prognoz.

Etap 2. Wykorzystanie syntetycznych mierników trafności prognoz.

Na tym etapie jako mierniki wykorzystane zostaną dwie powszechnie stosowane miary trafności prognoz: pierwiastek błędu średniokwadratowego RSME, średni błąd absolutny MAE oraz średni błąd procentowy MAPE. Jeżeli na ich podstawie możliwe będzie wskazanie algorytmu generującego najbardziej trafne prognozy (uwzględniając także prognozę ważoną uzyskaną w etapie 1), badanie się kończy. Jeżeli nie będzie można jednoznacznie wskazać algorytmu generującego najbardziej trafne prognozy, nastąpi przejście do 3 etapu badania.

Etap 3. Wykorzystanie testu weryfikującego trafność analizowanych prognoz.

W przypadku trudności z jednoznacznym wskazaniem w etapie 2 najbardziej trafnych prognoz wykorzystany zostanie test Diebolda-Mariano [Diebold, Mariano 1995] pozwalający rozstrzygnąć, czy trafność dwóch porównywanych ze sobą prognoz, statystycznie rzecz biorąc, nie różni się od siebie (hipoteza zerowa), czy jednak prognozy te charakteryzują się różną trafnością (hipoteza alternatywna). Jednocześnie postanowiono przyjąć jako prognozę porównawczą predykcję z modelu AR(p), który w pewnym sensie można także traktować jako automatyczny algorytm prognostyczny, w którym jedynym parametrem do ustalenia jest rząd autoregresji. Jeżeli w tym etapie uda się wskazać algorytmy generujące wyraźnie bardziej trafne prognozy od predykcji porównawczej, to ostatecznie zostanie skonstruowana prognoza ważona składająca się z prognoz uzyskanych właśnie na podstawie tych algorytmów. Jeżeli na podstawie wyników testu Diebolda-Mariano nie będzie można jednoznacznie wskazać algorytmów generujących bardziej trafne prognozy od innych, to zostanie zaproponowany inny dobór prognoz do prognozy ważonej.

Etap 4. Ocena trafności skonstruowanej prognozy ważonej.

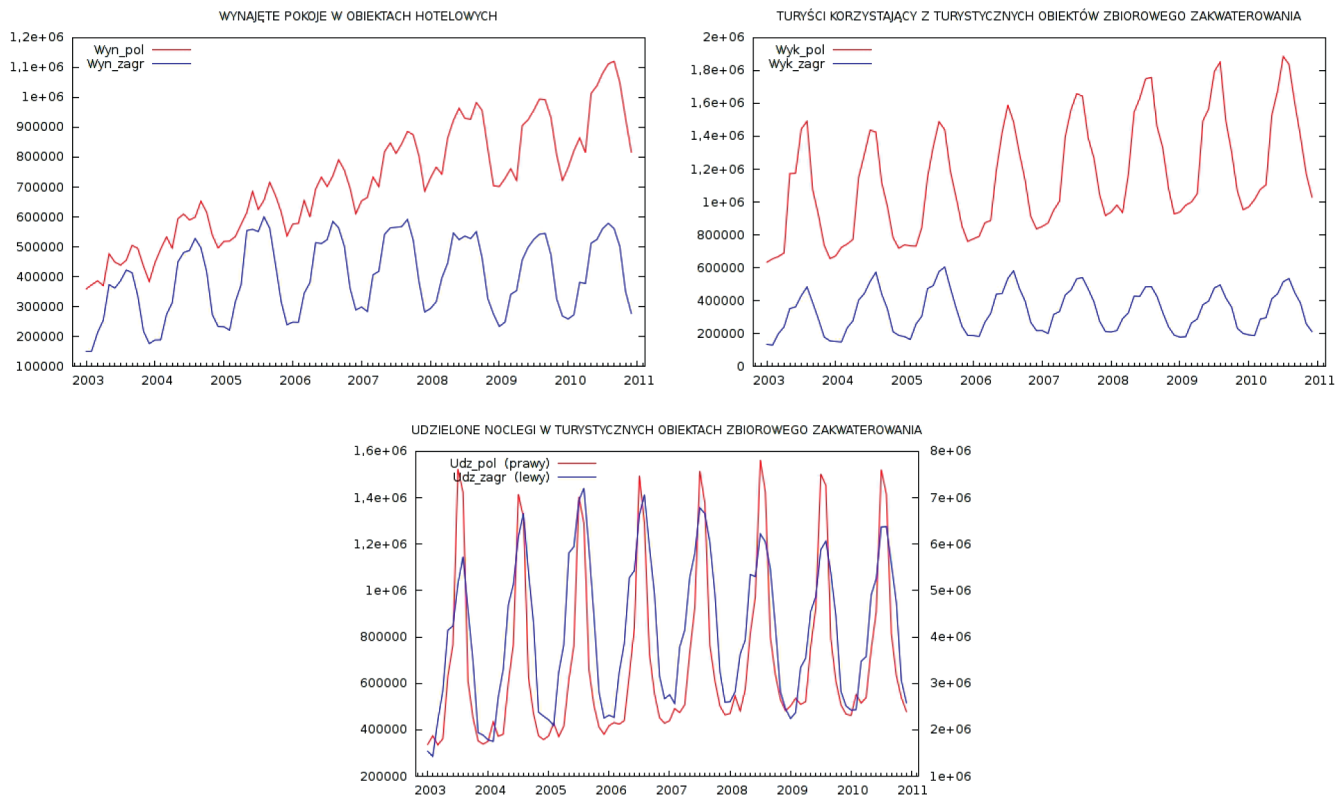
W tym etapie prognoza ważona uzyskana w etapie 3 zostanie oceniona pod kątem trafności za pomocą mierników RSME, MAE i MAPE i uzyskane wyniki zostaną porównane z trafnością prognoz uzyskanych na podstawie analizowanych algorytmów.

Powyższa czteroetapowa procedura pozwoli na zweryfikowanie postawionych w badaniu hipotez.

4. Przykład empiryczny

Prezentowane wyniki badań empirycznych dotyczą miesięcznych szeregów czasowych wykorzystania obiektów zbiorowego zakwaterowania, udzielonych noclegów oraz wynajętych pokoi hotelowych w Polsce z podziałem na turystów zagranicznych i polskich od stycznia 2003 r. do grudnia 2010 r. Dane pochodzą z publikacji Głównego Urzędu Statystycznego dostępnych na jego stronie internetowej. Rysunek 2 przedstawia przebiegi analizowanych szeregów.

Analiza przebiegu procesów widocznych na rys. 2 wskazuje na ich niestacjonarny charakter, w tym szczególnie na występowanie sezonowości. Biorąc jednak

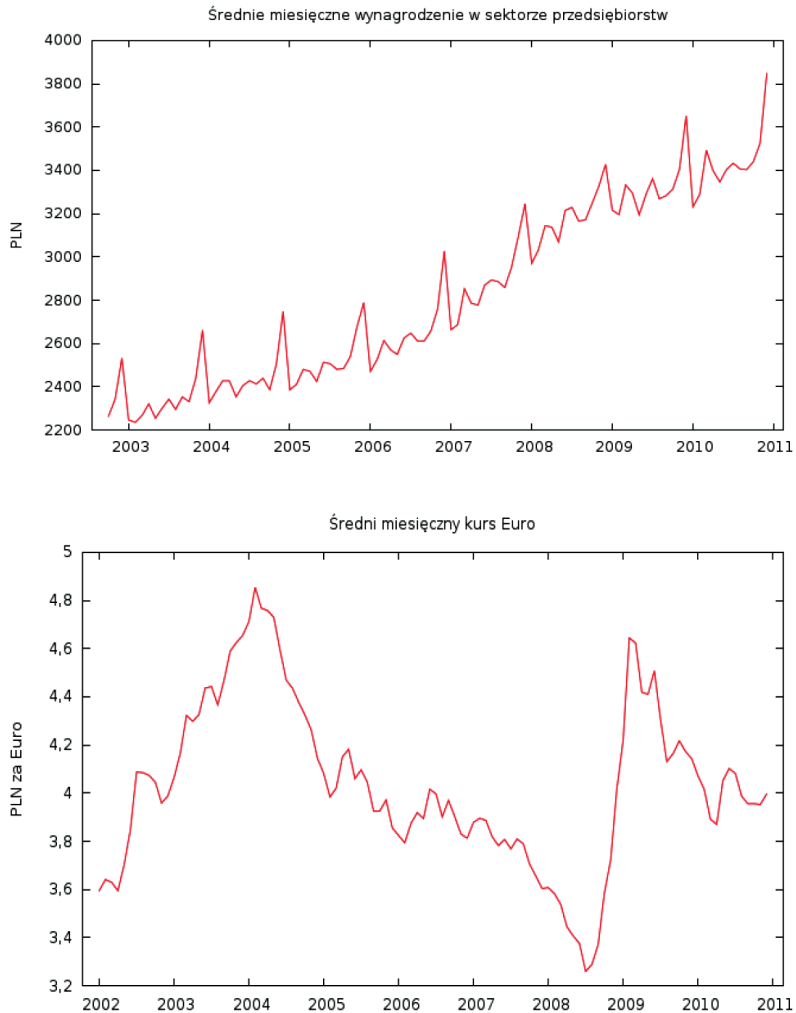


Rys. 2. Przebiegi modelowanych procesów w okresie od stycznia 2003 r. do grudnia 2010 r.

Źródło: opracowanie własne.

pod uwagę kategorię ekonomiczną tych zmiennych, należało się spodziewać takich charakterystyk. Ponieważ w badaniu ocenie będą podlegać także algorytmy automatycznej specyfikacji modeli przyczynowych, to jako zmienne niezależne wykorzystano:

- 1) w przypadku turystów polskich – średnie miesięczne wynagrodzenie w sektorze przedsiębiorstw,
 - 2) w przypadku turystów zagranicznych – średni miesięczny kurs euro.
- Rysunek 3 przedstawia przebiegi tych dwóch kategorii.

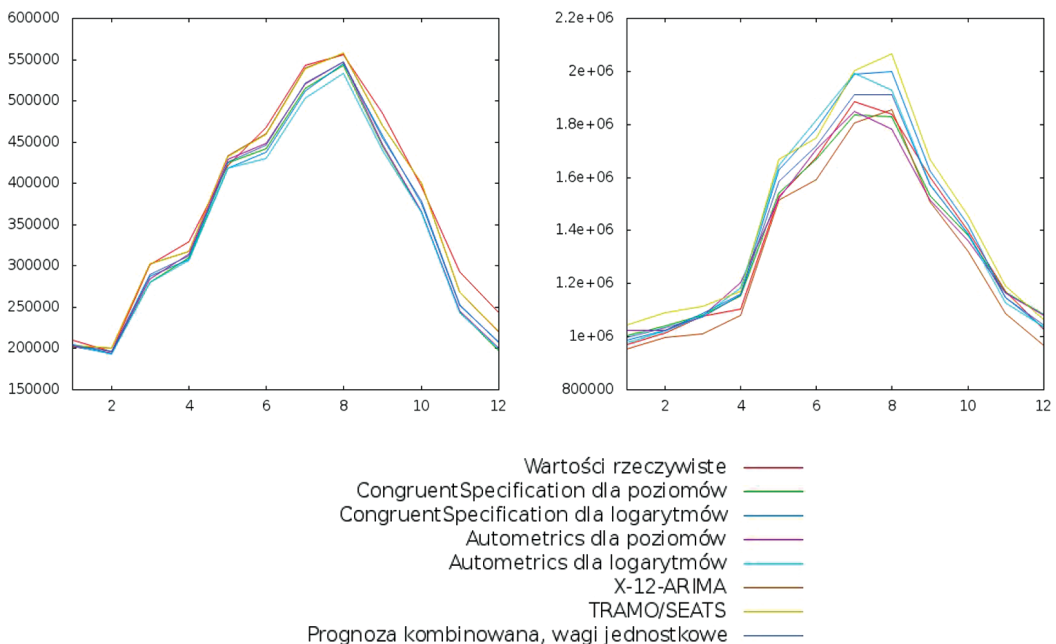


Rys. 3. Przebiegi średniego miesięcznego wynagrodzenia oraz średniego kursu euro w badanym okresie
Źródło: opracowanie własne.

W badaniu empirycznym wykorzystano wszystkie cztery omówione algorytmy automatycznej specyfikacji modelu w następujących konfiguracjach:

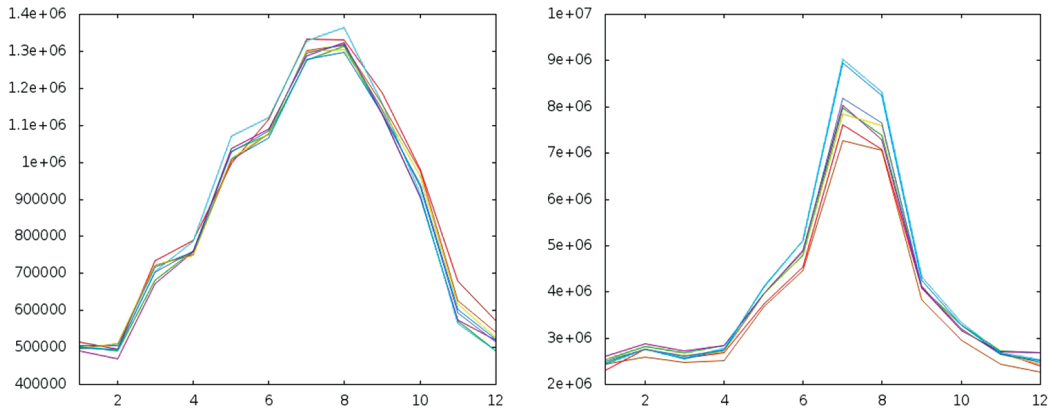
- 1) Congruent Specification dla zmiennych w poziomach,
- 2) Congruent Specification dla zmiennych zlogarytmowanych,
- 3) Autometrics dla zmiennych w poziomach,
- 4) Autometrics dla zmiennych zlogarytmowanych,
- 5) X-12-ARIMA w wariancie pełnej automatycznej analizy,
- 6) TRAMO/SEATS w wariancie pełnej automatycznej analizy.

W celu oceny trafności prognoz próba została skrócona o 12 miesięcy, dla których wygenerowano prognozy ze wszystkich sześciu konkurencyjnych modeli. Dodatkowo skonstruowano prognozę ważoną z jednakowymi wagami zgodnie z etapem 1 procedury badawczej. Rysunek 4 przedstawia przebiegi wygenerowanych prognoz wykorzystania obiektów zbiorowego zakwaterowania przez turystów zagranicznych (rysunek po lewej stronie) oraz polskich (rysunek po prawej stronie). Rysunek 5 przedstawia analogiczne prognozy udzielonych noclegów, natomiast rysunek 6 przedstawia prognozy wynajętych pokoi hotelowych w Polsce. Tabela 1 przedstawia syntetyczne mierniki trafności tych prognoz.



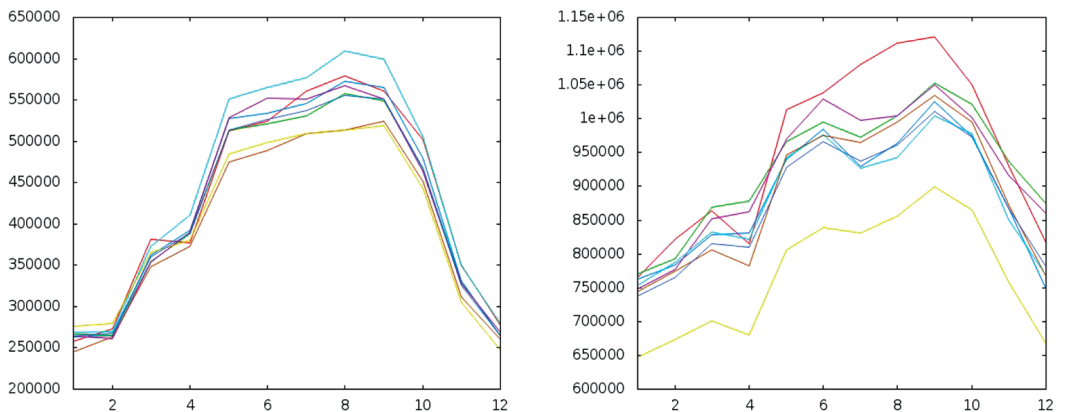
Rys. 4. Prognozy wykorzystania obiektów zbiorowego zakwaterowania przez turystów zagranicznych i polskich od stycznia do grudnia 2010 r.

Źródło: opracowanie własne.



Rys. 5. Prognozy noclegów udzielonych turystom zagranicznym i polskim od stycznia do grudnia 2010 r.

Źródło: opracowanie własne.



Rys. 6. Prognozy pokoi hotelowych wynajętych turystom zagranicznym i polskim od stycznia do grudnia 2010 r.

Źródło: opracowanie własne.

Analiza zarówno przebiegów prognoz, jak i syntetycznych mierników ich trafności nie pozwala na jednoznaczne wskazanie algorytmu automatycznej specyfikacji, który systematycznie generuje bardziej trafne prognozy od pozostałych procedur. W związku z tym postanowiono przejść do etapu 3 procedury badawczej i wykorzystać test Diebolda-Mariano do zweryfikowania hipotez mówiących o braku istotnych różnic w trafności poszczególnych prognoz. Jako prognozę porównawczą wykorzystano predykcję z modelu AR(12). Wyniki testu Diebolda-Mariano zaprezentowane są w tab. 2.

Tabela 1. Syntetyczne mierniki trafności wygenerowanych prognoz analizowanych szeregów czasowych

	Turyści zagraniczni			Turyści polscy		
	RSME	MAE	MAPE	RSME	MAE	MAPE
Trafność prognoz wykorzystania obiektów zbiorowego zakwaterowania						
Congruent Specification dla poziomów	28 291	23 793	6,90%	35 507	28 180	2,22%
Congruent Specification dla logarytmów	23 891	20 137	5,71%	73 562	54 089	3,52%
Autometrics dla poziomów	25 905	21 448	6,29%	49 649	39 446	3,04%
Autometrics dla logarytmów	31 493	27 083	7,54%	72 250	53 692	3,55%
X-12-ARIMA	12 373	9 713	3,13%	61 984	53 352	3,98%
TRAMO/SEATS	12 193	9 703	3,08%	99 013	82 503	5,83%
Prognoza kombinowana, wagi jednostkowe	21 481	17 857	5,20%	38 281	32 526	2,35%
Trafność prognoz noclegów udzielonych turystom						
Congruent Specification dla poziomów	56 268	47 941	5,84%	212 402	181 688	5,06%
Congruent Specification dla logarytmów	42 464	36 668	4,22%	555 271	341 049	6,57%
Autometrics dla poziomów	52 166	45 753	5,78%	234 192	191 934	5,45%
Autometrics dla logarytmów	51 844	37 800	4,85%	585 671	365 954	7,14%
X-12-ARIMA	27 632	23 946	2,95%	179 322	157 623	4,78%
TRAMO/SEATS	33 932	30 572	3,74%	200 153	140 365	3,33%
Prognoza kombinowana, wagi jednostkowe	41 203	35 075	4,29%	269 223	185 423	4,12%
Trafność prognoz pokoi hotelowych wynajętych turystom						
Congruent Specification dla poziomów	18 900	15 823	3,77%	58 594	47 604	4,82%
Congruent Specification dla logarytmów	13 784	12 501	3,10%	81 418	68 384	6,80%
Autometrics dla poziomów	19 332	16 762	4,01%	53 713	45 164	4,60%
Autometrics dla logarytmów	24 360	18 833	4,01%	87 256	71 639	7,05%
X-12-ARIMA	37 118	32 607	7,20%	70 218	64 363	6,54%
TRAMO/SEATS	37 643	32 405	7,32%	188 657	183 859	19,08%
Prognoza kombinowana, wagi jednostkowe	17 650	14 557	3,48%	84 515	73 157	7,29%

Źródło: opracowanie własne.

Tabela 2. Wartości p uzyskane w teście Diebolda-Mariano weryfikującym istotność różnic trafności konkurencyjnych prognoz

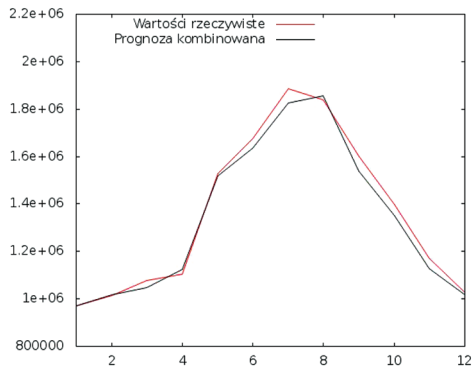
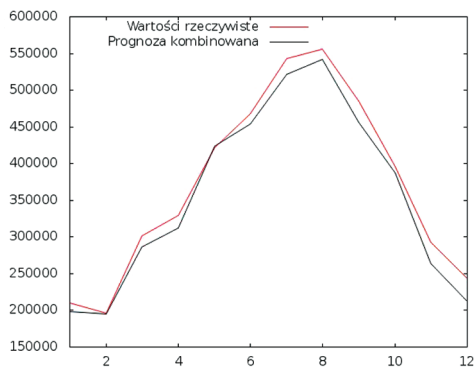
	Wykorzystanie obiektów zbiorowego zakwaterowania		Udzielone noclegi		Wynajęte pokoje	
	zagraniczni	polscy	zagraniczni	polscy	zagraniczni	polscy
Turyści						
Congruent Specification dla poziomów	68,6%	92,9%	74,8%	56,6%	79,6%	66,0%
Congruent Specification dla logarytmów	85,8%	57,0%	99,6%	61,8%	64,3%	99,4%
Autometrics dla poziomów	79,8%	67,1%	81,4%	48,5%	80,6%	57,8%
Autometrics dla logarytmów	51,7%	51,0%	86,3%	61,6%	99,7%	93,4%
X-12-ARIMA	47,9%	37,8%	59,8%	80,0%	60,7%	79,5%
TRAMO/SEATS	46,7%	53,0%	76,0%	75,8%	65,0%	7,5%
Prognoza kombinowana, wagi jednostkowe	99,9%	82,3%	96,8%	64,5%	76,6%	96,1%

Źródło: opracowanie własne.

Analiza wyników przedstawionych w tab. 2 nie pozwala na jednoznaczne wskazanie algorytmów generujących bardziej trafne prognozy od pozostałych. Najmniejsza wartość p została osiągnięta dla prognozy pokoi hotelowych wynajętych przez polskich turystów uzyskanej na podstawie procedury TRAMO/SEATS i wynosiła 7,5%. Taki wynik nie pozwala na wskazanie algorytmów, na podstawie których powinno się budować prognozy ważone. W związku z tym zaproponowano skonstruowanie prognozy ważonej na podstawie następujących modeli (wagi 1/3):

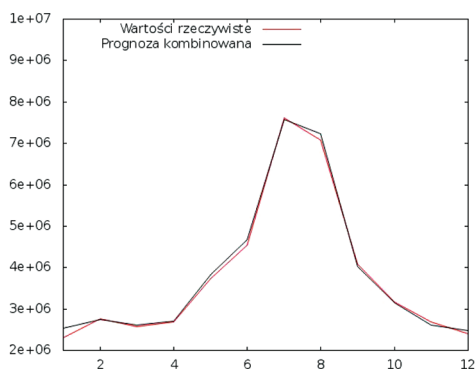
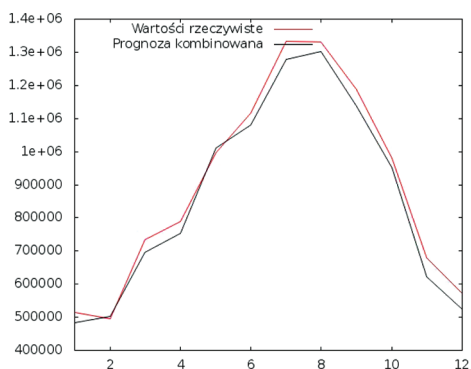
- 1) autoregresyjny model struktury AR(12),
- 2) specyfikacja zgodna uzyskana na podstawie procedury Congruent Specification dla poziomów zmiennych,
- 3) specyfikacja na podstawie procedury X-12-ARIMA w wersji pełnej automatycznej analizy.

Taki wybór podyktowany był powszechną dostępnością procedur Congruent Specification oraz X-12-ARIMA, które są darmowymi modułami. Z kolei model autoregresyjny, który w tym kontekście może być traktowany nawet jak prognoza naiwna, w rzeczywistości często generuje trafne prognozy. Rysunki 7-9 przedstawiają przebiegi prognoz ważonych uzyskanych zgodnie z zaproponowaną kombinacją, przy czym rysunki po lewej stronie dotyczą turystów zagranicznych, a te po prawej stronie turystów polskich. Tabela 3 prezentuje syntetyczne mierniki trafności tych prognoz.



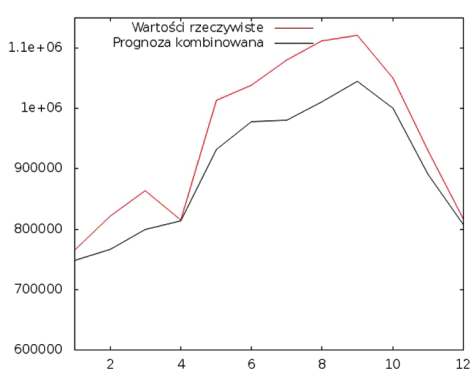
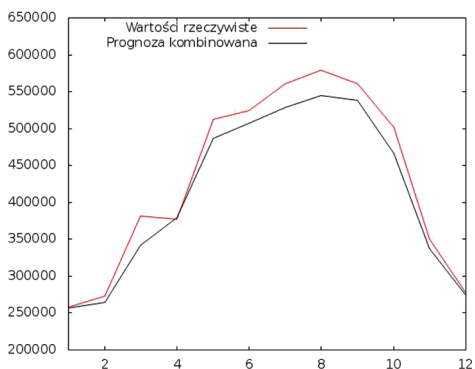
Rys. 7. Prognozy ważone wykorzystania obiektów zbiorowego zakwaterowania przez turystów zagranicznych i polskich od stycznia do grudnia 2010 r.

Źródło: opracowanie własne.



Rys. 8. Prognozy ważone noclegów udzielonych turystom zagranicznym i polskim od stycznia do grudnia 2010 r.

Źródło: opracowanie własne.



Rys. 9. Prognozy ważone pokoi hotelowych wynajętych turystom zagranicznym i polskim od stycznia do grudnia 2010 r.

Źródło: opracowanie własne.

Tabela 3. Syntetyczne mierniki trafności prognoz ważonych analizowanych szeregów

	Wykorzystanie obiektów zbiorowego zakwaterowania		Udzielone noclegi		Wynajęte pokoje	
	zagraniczni	polscy	zagraniczni	polscy	zagraniczni	polscy
Turyści						
RSME	18 902	35 326	38 463	99 603	23 503	63 220
MAE	16 239	28 525	35 514	78 356	19 392	54 684
MAPE	4,78%	2,00%	4,28%	2,36%	4,16%	5,48%

Źródło: opracowanie własne.

Analiza wyników zaprezentowanych w tab. 3 wskazuje na wysoką trafność prognoz kombinowanych z trzech modeli, tj. AR(12), specyfikacji uzyskanej dzięki algorytmom Congruent Specification oraz X-12-ARIMA, ponieważ największa wartość miernika MAPE wyniosła 5,48% i została zanotowana dla prognoz liczby pokoi hotelowych p wynajętych turystom polskim. Warto jednak zauważyć, że wartości MAPE wyliczone dla prognoz tej kategorii wygenerowanych przez poszczególne algorytmy automatycznej specyfikacji wahały się od 4,6 do 19,08% (por. tab. 1), co wskazuje na stosunkową trudność uzyskania trafnych prognoz tej kategorii z wykorzystaniem narzędzi automatycznego wyboru modelu, a wartość MAPE na poziomie 5,48% stawia uzyskaną prognozę kombinowaną na drugim miejscu pod względem trafności (bardziej trafne prognozy generowała tylko procedura Autometrics dla logarytmów zmiennych). Porównując trafność prognoz dla pozostałych szeregów, należy zauważyć, że zaproponowana kombinacja trzech prognoz generowała najmniejsze błędy MAPE dla prognoz wykorzystania obiektów zbiorowego zakwaterowania oraz udzielonych noclegów dla turystów polskich. W przypadku turystów zagranicznych zaproponowana kombinacja generowała prognozy, które pod względem wartości MAPE plasowały się na miejscu trzecim (wykorzystanie obiektów zbiorowego zakwaterowania), czwartym (udzielone noclegi) oraz piątym (liczba wynajętych pokoi hotelowych).

5. Podsumowanie

Podsumowując wyniki przeprowadzonego badania, należy stwierdzić, że pozytywnie weryfikują one sformułowane na wstępie hipotezy badawcze:

1. Algorytmy automatycznej specyfikacji modelu są użytecznym narzędziem prognozowania ekonomicznych szeregów czasowych.

2. Ze względu na różnorodność wewnętrznych struktur modelowanych szeregów prognozy kombinowane uzyskane na podstawie:

1) modelu autoregresyjnego,

2) modelu przyczynowo-skutkowego wyspecyfikowanego zgodnie z koncepcją modelowania zgodnego na podstawie procedury Congruent Specification,

3) modelu struktury procesu wybranego przez procedurę X-12-ARIMA, polepszają trafność generowanych prognoz.

Literatura

- Błażejowski M., Kufel P., Kufel T., *Automatic Procedure of Building Congruent Dynamic Model in Gretl*, [w:] *Econometrics with Gretl. Proceedings of the Gretl Conference 2009*, I. Díaz-Emparanza, P. Mariel, M.V. Esteban (red.), Universidad Del Pais Vasco, Bilbao 2009.
- Błażejowski M., Kufel P., Kufel T., *Automatic Procedure of Building Specification of Dynamic Congruent Model in GRETL Package*, Optimum, Studia Ekonomiczne nr 4(48), Wydawnictwo Uniwersytetu w Białymstoku, Białystok 2010.
- Diebold F.X., Mariano R.S., *Comparing predictive accuracy*, "Journal of Business & Economic Statistics" vol. 13, no 3 (Jul., 1995).
- Doornik J.A., *Autometrics*, [w:] *The Methodology and Practice of Econometrics*, J.L. Castle, N. Shephard (red.), Oxford University Press, 2009.
- Hoover K.D., Perez S.J., *Data mining reconsidered: encompassing and the general-to-specific approach to specification search*, "Econometrics Journal, Royal Economic Society" 1999, vol. 2(2).
- Szmuksta-Zawadzka M., Zawadzki J., *O miernikach dokładności prognoz ex post w prognozowaniu zmiennych o silnym natężeniu sezonowości*, *Metody Ilościowe w Badaniach Ekonomicznych = Quantitative Methods in Economics* 2012, t. 13/1.

COMPARATIVE ANALYSIS OF AUTOMATIC MODELLING AND PREDICTION PROCEDURES

Summary: Automatic model selection issues lie in modern dynamic econometric research area and concern theoretical aspects of such an approach as well as the implementation of algorithms. In the paper we compare empirical results for two automatic model selection procedures: Autometrics for PcGets based on "general to specific" approach and Congruent-Specification for Gretl based on dynamic congruent modelling approach. We also compare results from the common automatic time-series model selection procedures.

Keywords: CongruentSpecification, Autometrics, model selection procedures, model uncertainty, combined forecasts.