



Krzysztof Ćwikliński

Ministerstwo Finansów¹
Departament Analiz
Uniwersytet Ekonomiczny we Wrocławiu
Wydział Ekonomii i Finansów
Katedra Ekonometrii i Badań Operacyjnych
krzysztof.cwiklinski@ue.wroc.pl

PROGNOZOWANIE SZEREGÓW CZASOWYCH AKTUALIZACJI JEDNOLITYCH PLIKÓW KONTROLNYCH

Streszczenie: Prognozowanie szeregów czasowych stało się niezbędne w procesie kontrolowania procesów zachodzących w systemach informatycznych Ministerstwa Finansów. Wymierne w sensie finansowym są problemy braku lub niepełnej aktualizacji relacyjnej bazy danych JPK_VAT w akceptowalnym przez prawo terminie. W tym przypadku niezwykle ważną okazuje się umiejętność zastosowania nie tylko klasycznych modeli uwzględniających składniki sezonowe (np. SARIMA), ale także złożone składniki systematyczne (BATS/TBATS). Dokonano analizy szeregów czasowych pod kątem występowania składników systematycznych, postawiono prognozy i przetestowano reszty. Otrzymano i zestawiono wyniki testów wskazujące na konieczność zastosowania modelu TBATS.

Słowa kluczowe: Jednolity Plik Kontrolny, analiza szeregów czasowych, dekompozycja, prognozowanie, BATS/TBATS, SARIMA.

JEL Classification: C32, C53.

Wprowadzenie

W zapisie Ustawy o podatku od towarów i usług [2004, art. 109, ust. 8a]: „Ewidencja, o której mowa w ust. 3, prowadzona jest w formie elektronicznej przy użyciu programów komputerowych”. Sprawne działanie systemu ewiden-

¹ Wszystkie zawarte w niniejszym artykule fakty, badania i wnioski nie reprezentują stanowiska Ministerstwa Finansów lub mojego jako pracownika Ministerstwa Finansów, a jedynie stanowisko osoby prywatnej (autora artykułu).

cjonowania służy m.in. poprawie ściągłości podatku od towarów i usług, sprzyja występowaniu dodatniego salda budżetu państwa, generuje przychody, zmniejsza tendencję do zadłużania, jest stymulantą poziomu życia społeczeństwa. Wymierne w sensie finansowym dla polskiego Ministerstwa Finansów są problemy braku lub niepełnej aktualizacji bazy danych JPK_VAT w akceptowalnym przez prawo terminie. Wymogi dotyczące prawidłowej aktualizacji systemu baz danych JPK_VAT skłoniły autora do przeprowadzenia badań, które wyszły naprzeciw potrzebom regularnego i prawidłowego kontrolowania aktualizacji z wykorzystaniem algorytmów obliczeniowych implementowanych w języku R.

1. Prognozowanie szeregów aktualizacji z uwzględnieniem wahań sezonowych

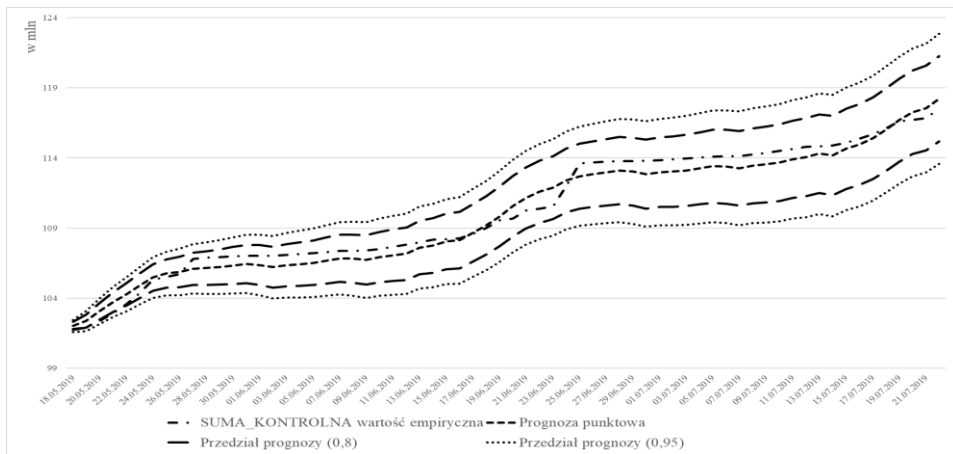
Celem badawczym jest otrzymanie takich prognoz przedziałowych, żeby stanowiły prawidłowy element systemu kontroli aktualizacji tabel JPK_VAT w Ministerstwie Finansów. Prawidłowe prognozy punktowe są uznawane za trafne, gdy wartości błędów *MAPE* (*Mean Absolute Percentage Error*) wynoszą mniej niż 3%. Przedziały prognoz mają być największe spośród uzyskanych z testowanych w artykule modeli. Sformułowano dwie hipotezy służące realizacji celu pracy, jakim jest trafne postawienie prognoz dotyczących aktualizacji tabel bazy danych JPK.

Hipoteza pierwsza przyjęta w niniejszym artykule brzmi: zastosowanie modelu SARIMA do badania aktualizacji tabel JPK_VAT może być dopuszczalne ze względu na brak niewygasającej autokorelacji w resztach modelu, jednak będzie dostarczać mniej dokładnych prognoz ze zbyt szerokimi przedziałami (w porównaniu do prognoz z modeli BATS/TBATS). **Hipoteza druga** to: TBATS jest najlepszy do modelowania szeregów czasowych aktualizacji tabel JPK_VAT, prognozy charakteryzują się niskimi błędami MAPE, a w szeregach reszt nie występuje efekt autokorelacji. W celu weryfikacji lub falsyfikacji hipotez w punkcie drugim przeprowadzone zostaną badania związane z modelowaniem, dekompozycją i prognozowaniem szeregów czasowych aktualizacji tabel JPK_VAT oraz prognozowanie niestacjonarnych szeregów czasowych aktualizacji bazy JPK_VAT.

Proces prognozowania z modelu uwzględnia wszystkie oszacowane parametry, których wartość wpływa na trend krótkookresowy i średni poziom badanego zjawiska [De Livera, Hyndman, Snyder, 2010, s. 19-20]. Występujące oddziaływania nieliniowe, w tym składniki sezonowe modelowane przez szeregi Fouriera, wchodzi w skład prognozy punktowej oraz przedziałów prognoz

(w sposób odrębny prognozowana jest wariancja) [De Livera, Hyndman, Snyder, 2010, s. 19-20]. Próbę uczącą stanowiły skumulowane szeregi czasowe aktualizacji tabel JPK_VAT z okresu od 18 maja 2018 r. do 17 maja 2019 r. ($N = 365$). W celu wyboru odpowiedniego modelu ze względu na wartość MAPE dokonano ekstrapolacji na okres próby testowej.

Na rysunkach 1-3 przedstawiono przykładowe wyniki prognozowania dla aktualizacji tabeli SUMA_KONTROLNA, która jest reprezentatywna dla danych w bazie JPK_VAT – zawiera m.in. informacje o liczbie wierszy ewidencji sprzedaży, zakupu, w okresie którego dotyczy JPK_VAT, kwoty podatków należnego i naliczonego (tabele: ZAKUP oraz SPRZEDAZ), a także inne informacje z pozostałych tabel.



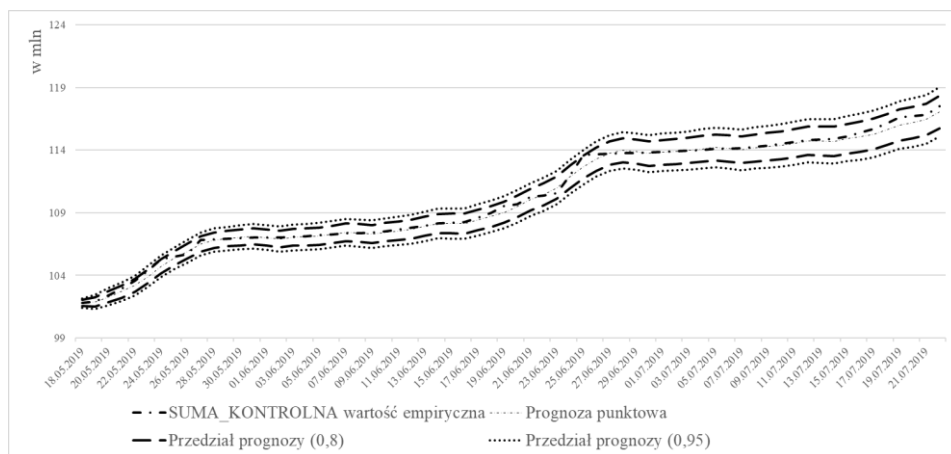
Rys. 1. Model BATS: prognozy i przedziały prognoz skumulowanych aktualizacji

Źródło: Opracowanie własne na podstawie danych z okresu 18 maja 2019 r. – 22 lipca 2019 r.

Analizując rysunek 1, można zauważyć stopniowe poszerzanie przedziałów na zakładanych poziomach ufności. W początkowym okresie próby testowej (18 maja 2019 r. – 22 maja 2019 r.) prognoza była przewartościowana, a wartości empiryczne niemalże pokrywały się z dolną granicą prognozy przedziałowej na poziomie ufności 0,80. Przez cały czas realne wartości aktualizacji nie wykraczały poza przedziały prognoz w okresie próby testowej, choć znaczące niedoszacowania miały miejsce przez większość okresu. Wartość empiryczna zwykle była powyżej wartości prognozowanej.

Okresami przełomowymi, jeśli chodzi o kształtowanie się wartości prognoz, w porównaniu do wartości empirycznych są: 22 maja 2019 r. – 28 maja 2019 r. oraz 19 czerwca 2019 r. – 25 czerwca 2019 r. W tym czasie występuje

wzmózniona zmienność związana z aktualizacją danych w bazie JPK_VAT, która związana jest terminami aktualizacji ewidencji podatku od towarów i usług. Skokowy przyrost wartości empirycznej aktualizacji wystąpił 27 maja 2019 r. – od tego dnia do 16 czerwca 2019 r. wartość empiryczna aktualizacji jest wyższa od teoretycznej. W dniach 17-19 czerwca 2019 r. wartość empiryczna niemalże pokrywa się z prognozami, jednak 20 czerwca 2019 r. aktualizacja nie jest tak wysoka jak wartości prognozy – od tego momentu aż do 24 czerwca 2019 r. wartości empiryczne są niższe od teoretycznych. Drugim okresem, w którym wartości empiryczne znowu stają się wyższe od teoretycznych, jest 25 czerwca 2019 r. – 19 lipca 2019 r. Rysunek 2, analogiczny do poprzedniego, przedstawia prognozy, przedziały prognoz na dwóch zakładanych poziomach ufności w porównaniu do wartości empirycznych na przykładzie modelu TBATS.



Rys. 2. Model TBATS: prognozy i przedziały prognoz skumulowanych aktualizacji

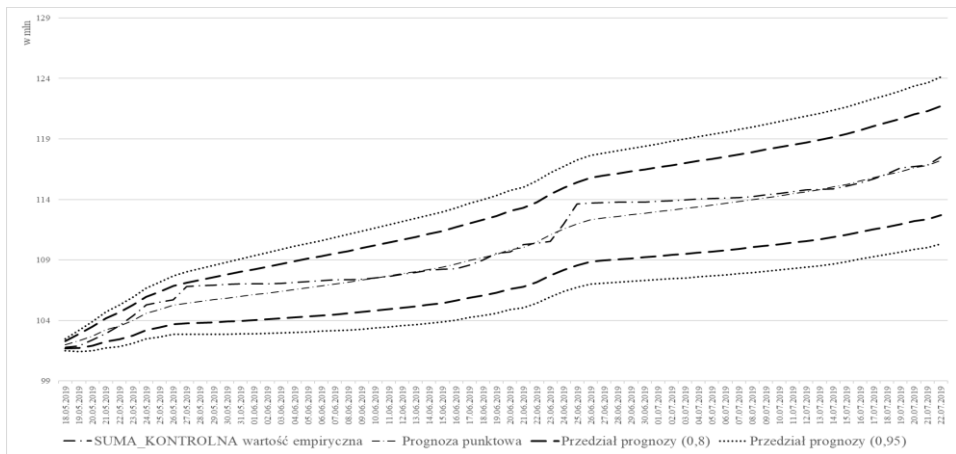
Źródło: Opracowanie własne na podstawie danych z okresu 18 maja 2019 r. – 22 lipca 2019 r.

W przypadku modelu TBATS wartości empiryczne są wyższe od teoretycznych w okresach 20-24 maja 2019 r. oraz 20-27 czerwca 2019 r.

W porównaniu do modelu BATS prognozy przedziałowe modelu TBATS mają znacznie mniejszą rozpiętość, choć z okresu na okres następuje stopniowe poszerzanie przedziałów. Chwilowe przekroczenia przedziałów występują w okresach wzmóznionych wahań aktualizacji: 24-25 czerwca 2019 r., a także 21-27 czerwca 2019 r., gdy wartości empiryczne oraz teoretyczne gwałtownie się zmieniają. Zmiany w kształtowaniu zmienności prognoz występują w okresach nierównowagi lub gdy termin aktualizacji jest wymagany (25. dzień miesiąca) i są związane z występowaniem wahań sezonowych. Okresy: 27 maja 2019 r. – 16 czerwca 2019 r.

oraz 27 czerwca 2019 r. – 14 lipca 2019 r. wskazują na niewielką zmienność wartości empirycznej i teoretycznej. W tych dniach wartości empiryczne bliskie są wartościom prognoz, z którymi niemalże się pokrywają. Jest to stan normalny, ponieważ płatnicy podatku VAT nie czują się zobligowani do aktualizacji ewidencji zbliżającym się prawnie uwarunkowanym terminem.

Stopniowe zwiększanie różnic pomiędzy szeregiem danych empirycznych i teoretycznych występuje, począwszy od 15 lipca 2019 r. do 21 lipca 2019 r. i charakteryzuje analogiczne dni maja oraz czerwca 2019 r. W inny sposób przedziały prognoz obliczane są przez model SARIMA – rysunek 3 całkowicie różni się od analogicznych wykresów powstałych na podstawie wyników modelowania BATS/TBATS.

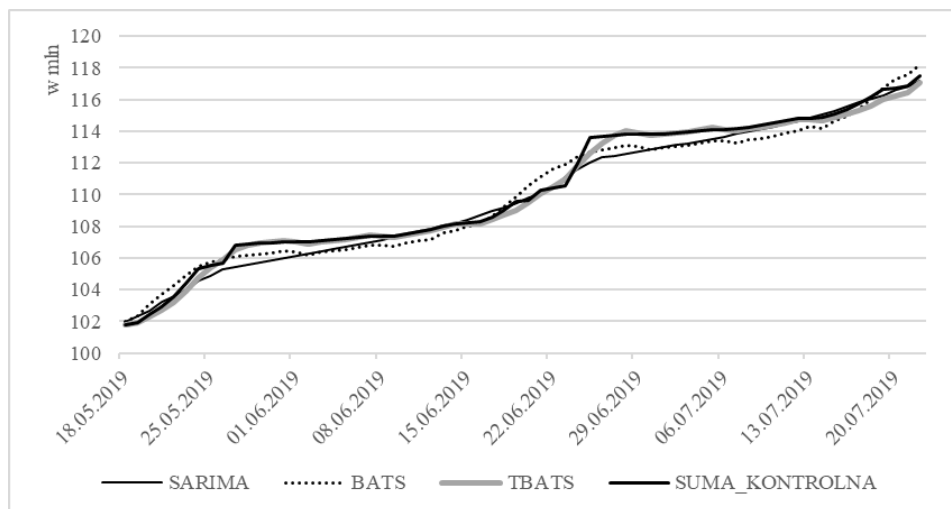


Rys. 3. Model SARIMA: prognozy i przedziały prognoz skumulowanych aktualizacji

Źródło: Opracowanie własne na podstawie danych z okresu 18 maja 2019 r. – 22 lipca 2019 r.

Przede wszystkim słabo dostrzegalne jest zakrzywienie wykresu prognozy punktowej w okresach wzmożonej aktualizacji. Można zaobserwować duże i szybko rosnące rozpiętości przedziałów prognozy, są znacznie szersze od przedstawionych na poprzednich rysunkach. Analiza potwierdza, że wartości empiryczne pozostają zbliżone do wartości teoretycznej zwłaszcza w okresach, które można uznać za stabilne: 8-14 czerwca 2019 r. oraz 8-14 lipca 2019 r. Jest kilka okresów niedowartościowania i przewartościowania prognoz w okresie próby testowej. Krótki okres przewartościowania prognozy występował od 18 do 20 maja 2019 r. Wartość empiryczna również była znacznie wyższa od prognoz w okresach: 22 maja 2019 r. – 9 czerwca 2019 r. oraz 23 czerwca 2019 r. – 11 lipca 2019 r.

Prognozy dopasowują się do wartości empirycznej, to świadczy również o prawidłowej postaci modelu. Prognozy punktowe z trzech modeli i wartość empiryczną w okresie próby testowej przedstawiono na rysunku 4.



Rys. 4. Prognozy punktowe (BATS/TBATS, SARIMA) i wartość empiryczna

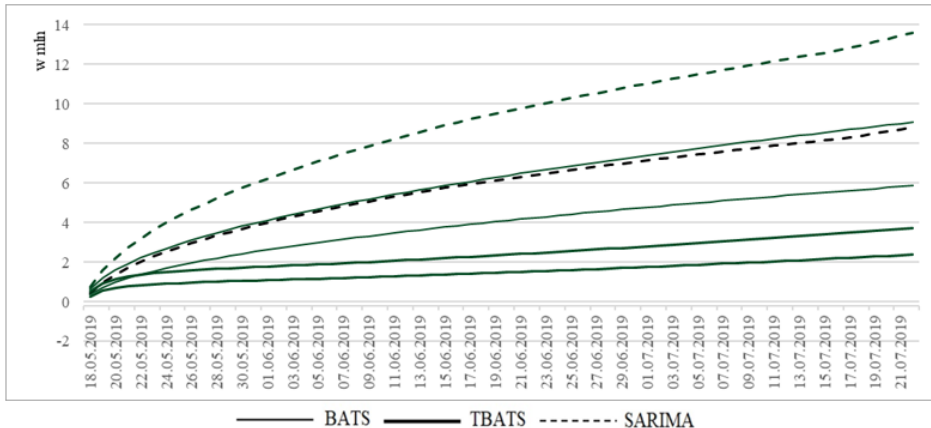
Źródło: Opracowanie własne na podstawie danych z okresu 18 maja 2019 r. – 22 lipca 2019 r.

W modelu TBATS występują dwa dłuższe okresy, gdy prognozy są w niewielkim stopniu oddalone od wartości empirycznych: 26 maja 2019 r. – 16 czerwca 2019 r. oraz 27 czerwca 2019 r. – 13 lipca 2019 r. Występują również okresy, gdy wartości empiryczne wyraźnie odbiegają od teoretycznych, prognoza jest wyraźnie niedowartościowana w dniach: 20-25 maja 2019 r., 17-20 czerwca 2019 r., 23-25 czerwca 2019 r. Znaczące niedowartościowanie lub przewartościowanie prognoz jest cechą charakterystyczną modelu BATS (linia przerywana). Prognozy z tego modelu były wiele razy w wielu okresach przewartościowane, np. w dniach 19-24 maja 2019 r., 19-24 czerwca 2019 r., 19-22 lipca 2019 r.

Znaczące niedowartościowanie prognoz z tego modelu występowało również dla dłuższych okresów: 27 maja 2019 r. – 14 czerwca 2019 r., 24 czerwca 2019 r. – 16 lipca 2019 r. Najmniej dokładne są prognozy modelu SARIMA: wartości teoretyczne są najbardziej oddalone do wartości empirycznych, a prognoza cechuje się najmniejszym dopasowaniem do wahań sezonowych. Przy występowaniu zgrupowań miesięcznych wahań sezonowych prognozy były niedowartościowane.

SARIMA wykazuje stosunkowo słabe możliwości odwzorowania złożonych sezonowości w porównaniu do pozostałych modeli branych pod uwagę

w artykule. Problemy z przedziałami prognoz modelu SARIMA potwierdza rysunek 5, który przedstawia rosnące rozpiętości przedziałów prognoz aktualizacji SUMA_KONTROLNA w miarę upływu okresu prognozowania (próba testowa).

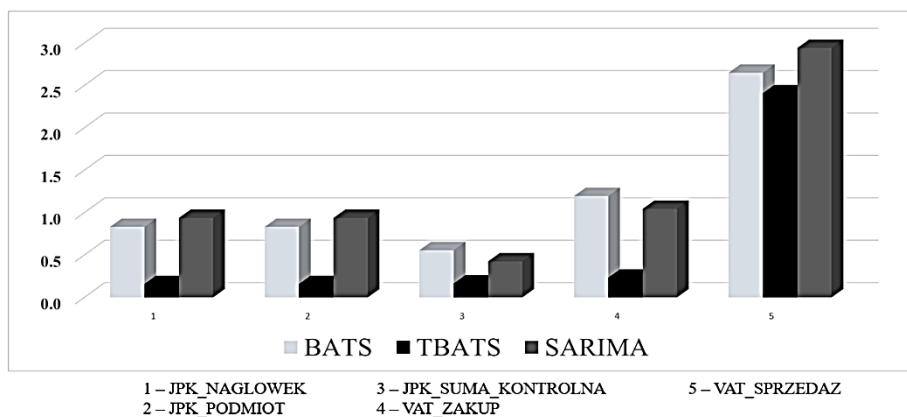


Rys. 5. Rozstęp przedziałów prognozy aktualizacji tabeli SUMA_KONTROLNA

Źródło: Opracowanie własne na podstawie danych z okresu 18 maja 2019 r. – 22 lipca 2019 r.

Rozpiętość oszacowano na podstawie minimalnych i maksymalnych wartości określających granice przedziałów – dla każdego przedziału i modelu występują po dwie linie odnoszące się do poziomów ufności. W przypadku SARIMA linie przerywane oznaczają rozstępy przedziału na poziomie ufności 0,80 oraz 0,95. Dla pozostałych modeli: BATS (cieńsze ciągłe linie – rozpiętość przedziału na poziomie ufności 0,80 i 0,95), TBATS (grube ciągłe linie – poziom ufności 0,80 i 0,95) – większy rozstęp jest charakterystyczny dla przedziałów prognozy na poziomie ufności 0,95.

Charakterystyczne dla rysunku 5 są wysokie i szybko poszerzające się rozstępy przedziałów prognoz w przypadku modelu SARIMA. Już od początku analizowanego okresu wyraźnie można dostrzec wyróżniającą się linię przerywaną, która od prognozy na dzień 19 maja 2019 r. wyraźnie góruje nad resztą (wykazuje wartości do kilku razy wyższe od rozpiętości przedziałów prognoz pochodzących z pozostałych analizowanych modeli). Model BATS charakteryzuje się rosnącą rozpiętością przedziałów prognoz, a rozstęp przedziału na poziomie ufności 0,95 przewyższa wartość rozstępu przedziału prognozy z modelu SARIMA na poziomie ufności 0,80. Najniższym rozstępem charakteryzuje się TBATS i to zarówno dla przedziałów na poziomie ufności 0,95, jak i wynoszącym 0,80.



Rys. 6. Średni bezwzględny procentowy błąd prognozy (MAPE) dla prognoz modeli BATS/TBATS, SARIMA

Źródło: Opracowanie własne na podstawie danych z okresu 18 maja 2019 r. – 22 lipca 2019 r.

Rysunek 6 przedstawia błędy prognoz dla trzech metod (BATS – jasnoszary, TBATS – czarny, SARIMA – ciemnoszary) i czterech prognozowanych szeregów czasowych aktualizacji. Przedstawiono MAPE prognoz BATS/TBATS, SARIMA dla pięciu tabel, które w rzeczywistości stanowią jedynie cztery przypadki. Dwie tabele (NAGLOWEK oraz PODMIOT) są identyczne w poziomie aktualizacji. Charakterystyczne są wartości MAPE dla aktualizacji tabeli SPRZEDAZ. Największy MAPE generuje model SARIMA jednak również inne modele mają stosunkowo wysokie, choć dopuszczalne pozostają błędy *ex post*. Średni błąd prognozy mierzony wskaźnikiem MAPE zawiera się w przedziale (0,157%; 2,941%). Modelem, który generował prognozy charakteryzujące się **najniższym błędem *ex post***, był **TBATS** (kolor czarny). W dwóch przypadkach niższe od BATS błędy *ex post* generował SARIMA dla danych tabel: SUMA_KONTROLNA i ZAKUP. Dla NAGLOWEK i PODMIOT oraz SPRZEDAZ niższe MAPE generowały prognozy z modelu BATS.

W celu dalszej analizy sporządzono tabelę 1 składającą się z dwóch części, z której jedna dotyczy błędów MAPE dla prognoz w okresie próby testowej, druga jest z okresu próby uczącej. Obliczono wartości minimalne oraz maksymalne błędów, uwzględniając zarówno wartości w kolumnach, jak i wierszach. Maksimum z wierszy określa, który model miał największy błąd i których danych dotyczył. Wyróżniono czarną pogrubioną czcionką wartości minimalne oraz maksymalne błędów prognoz szeregów aktualizacyjnych tabel JPK_VAT.

Tabela 1. Błędy MAPE dla danych tabel JPK_VAT w okresach próby uczącej i testowej

Błędy MAPE [%] – próba testowa					
Tabela	BATS	TBATS	SARIMA	max	min
NAGLOWEK	0.832	0.157	0.937	0.937	0.157
PODMIOT	0.832	0.157	0.937	0.937	0.157
SUMA_KONTROLNA	0.556	0.163	0.425	0.556	0.163
ZAKUP	1.195	0.232	1.045	1.195	0.232
SPRZEDAZ	2.649	2.408	2.941	2.941	2.408
max	2.649	2.408	2.941		
min	0.556	0.157	0.425		
Błędy MAPE [%] – próba ucząca					
Tabela	BATS	TBATS	SARIMA	max	min
NAGLOWEK	0.290	0.244	0.308	0.308	0.244
PODMIOT	0.290	0.244	0.308	0.308	0.244
SUMA_KONTROLNA	0.296	0.242	0.293	0.296	0.242
ZAKUP	0.342	0.270	0.350	0.350	0.270
SPRZEDAZ	0.310	0.293	0.380	0.380	0.293
max	0.342	0.293	0.380		
min	0.290	0.242	0.293		

Źródło: Obliczenia własne na podstawie danych JPK_VAT z okresów: 18 maja 2018 r. – 17 maja 2019 r. (próba ucząca) oraz 18 maja 2019 r. – 22 lipca 2019 r. (próba testowa).

Maksymalny błąd prognozy modelu SARIMA zarówno w okresie próby testowej, jak i uczącej dotyczył tabeli SPRZEDAZ. Cechą charakterystyczną dla okresu próby uczącej w porównaniu do okresu próby testowej jest niższa skala błędów: maksymalny błąd wynosił dla próby testowej 2,941%, natomiast dla próby uczącej 0,380%.

Zarówno dla próby testowej, jak i uczącej najniższe błędy miały prognozy modelu TBATS, a najwyższe SARIMA. Najniższymi błędami MAPE dla SARIMA w dwóch okresach: próby uczącej (MAPE = 0,293%) oraz testowej (MAPE = 0,425%), były prognozy aktualizacji tabeli SUMA_KONTROLNA. Najniższy błąd, którym charakteryzują się prognozy BATS, w okresie próby testowej wynosi 0,556% i pochodzi z aktualizacji SUMA_KONTROLNA. W przypadku próby uczącej MAPE = 0,290% i dotyczy tabel NAGLOWEK i PODMIOT. Dla próby uczącej oraz testowej najwyższym błędem modelu TBATS jest MAPE prognoz tabeli SPRZEDAZ.

W modelowaniu BATS najwyższym błędem z próby uczącej obciążone były prognozy tabeli ZAKUP (MAPE = 0,342%), a dla próby testowej były to prognozy tabeli SPRZEDAZ. W przypadku modelu BATS najwyższy MAPE równy jest 0,342% i pochodzi z tabeli ZAKUP. Najniższy błąd prognoz z modelu BATS równy jest 0,29% i dotyczy próby uczącej NAGLOWEK i PODMIOT. Maksymalny błąd modelu z tej próby jest ok. siedmiokrotnie niższy od maksymalnego błędu prognoz próby testowej z modelu SARIMA tabeli SPRZEDAZ.

Minimalny błąd modelu prognoz TBATS (SUMA_KONTROLNA) jest ok. dwukrotnie wyższy od minimalnego błędu modelu próby testowej i dotyczy prognozowania aktualizacji tabel NAGLOWEK i PODMIOT.

Reasumując, występują argumenty za wyborem TBATS do systemu prognozowania aktualizacji tabel Jednolitych Plików Kontrolnych. Prognozy TBATS zarówno dla próby uczącej, jak i testowej charakteryzują się najniższymi błędami MAPE. Przy tym prognozy punktowe TBATS charakteryzują trafność i bardzo wysoka dokładność. W przypadku tego modelu nie występuje autokorelacja reszt na żadnym dedykowanym dla testów Ljunga–Boxa i Boxa–Pierce’a opóźnieniu. Rozstęp przedziałów prognozy charakteryzuje się niskimi, wolno rosnącymi wartościami, co czyni model użytecznym do włączenia go do systemu kontrolującego aktualizację tabel JPK_VAT w Ministerstwie Finansów.

W ostatnim rozdziale artykułu podsumowane zostaną wyniki prognozowania aktualizacji w odniesieniu do tez wyrażonych we wstępie, sformułowane zostaną wnioski i zalecenia dotyczące proponowanych kierunków dalszych badań.

2. Konkluzje z przeprowadzonych badań

W podsumowaniu przedstawione zostaną wnioski z artykułu oraz nastąpi ustosunkowanie się do hipotez. Przedstawione badania związane były z potrzebą wykonania analizy szeregów czasowych aktualizacji tabel bazy danych Jednolite Pliki Kontrolne. Główny cel artykułu dotyczył wyboru odpowiedniej metody dekompozycji szeregów czasowych aktualizacji z postawieniem trafnych co najmniej na kilka okresów prognoz. Realizacja celów badawczych wymagała wcześniejszego przeprowadzenia analizy spektralnej dla pozbawionych trendu, zróżnicowanych wartości aktualizacji w okresie próby uczącej. Periodogram wykrył obecność dwóch znaczących okresów wahań: tygodniowych oraz miesięcznych. Istnienie wahań sezonowych w szeregach aktualizacji z okresu próby uczącej potwierdził test g-Fishera, w którym brak podstaw do odrzucenia hipotezy zerowej oznacza brak występowania wahań sezonowych w szeregu.

Wartości *p-value* dla szeregów czasowych pozbawionych długookresowego trendu przedstawiono w tabeli 2.

Tabela 2. Wartości *p-value* dla testu g-Fishera w pomiarze sezonowości

	Test g-Fishera				
	NAGLOWEK	PODMIOT	SUMA_KONTROLNA	ZAKUP	SPRZEDAZ
<i>p-value</i>	1.5971E-30	1.5971E-30	2.2883E-31	3.5352E-24	2.0044E-18

Źródło: Obliczenia własne dla danych zróżnicowanych (okres próby uczącej: 18 maja 2018 r. – 17 maja 2019 r.).

Wartości są bliskie zeru, hipoteza zerowa została odrzucona na rzecz hipotezy o występowaniu wahań sezonowych w szeregach czasowych aktualizacji tabel. Zauważono, że najwyższą wartość *p-value* otrzymano z modelowania aktualizacji tabeli SPRZEDAZ, a nieco niższą w przypadku ZAKUP.

Wartości prawdopodobieństw dla pozostałych trzech tabel są na znacznie niższym poziomie, co potwierdza, że wyniki modelowania tych tabel są zbliżone do siebie w podobieństwie rozkładów. Ostateczne dowody na istnienie składników sezonowych uzyskano po modelowaniu oraz dekompozycji szeregów czasowych, która umożliwiła prezentację składników sezonowych na wykresie [zob. Ćwikliński, 2020, rys. 3]. Uzyskano dodatkowe informacje o skłonności do wysyłania deklaracji przez podmioty gospodarcze w określonych okresach. Wyznaczono dni tygodnia, w których podmioty, płatnicy podatku od towarów i usług w sposób najbardziej intensywny dopełniają ewidencyjnych obowiązków. Dzięki temu wiadomo, że maksymalna wartość aktualizacji występowała w czwartki, a soboty były dniami wolnymi dla większości podmiotów (deklaracje składano najrzadziej). Wyboru odpowiedniej metody dokonano spośród popularnych modeli uwzględniających składniki sezonowe: BATS/TBATS oraz SARIMA. W przypadku modelu BATS wykazano, że w większości przypadków nie nadaje się on do prognozowania, ponieważ generuje reszty, które są skorelowane. Zwłaszcza dla reszt z modelowania aktualizacji tabeli SUMA_KONTROLNA występowała niegasnąca autokorelacja na wszystkich branych po uwagę opóźnieniach. Jedynie w przypadku tabeli SPRZEDAZ reszty modelu BATS były niezależne. Przy tym dla szeregu aktualizacji tabel ZAKUP oraz SPRZEDAZ można było wykazać słabą (zanikającą) zależność reszt na dwóch opóźnieniach, na poziomie istotności 0,1.

Zgodnie z treścią hipotezy pierwszej model SARIMA miał reszty nieskorelowane w większości przypadków, jednak znacznie poważniejszym problemem były dla tego modelu niezwykle szerokie przedziały prognoz. Rozstęp ten nie miał sobie równych spośród badanych modeli i przyczynił się do potwierdzenia hipotezy pierwszej.

Potwierdzono hipotezę, że zastosowanie modelu SARIMA do badania aktualizacji tabel JPK_VAT może być dopuszczalne ze względu na brak niewygasającej autokorelacji w resztach modelu, będzie dostarczać mniej dokładnych prognoz ze zbyt szerokimi przedziałami (w porównaniu do prognoz modeli BATS/TBATS).

Przedstawiono wyniki modelowania i dzięki nim uzyskano przedziały prognoz na dwóch poziomach istotności, a następnie obliczono rozstęp i poddano analizie odpowiedni wykres. Na rysunku 7 zaprezentowano rozstępy prognoz na

dwóch poziomach istotności, z których najbardziej charakterystyczny był górujący nad pozostałymi łuk rozstępu przedziału prognozy SARIMA.

W tabeli 3 przedstawiono wybrane miary szeregów czasowych rozstępów na poziomie istotności 0,8 oraz 0,95 dla wyników prognozowania modeli, których reszty nie wykazywały autokorelacji na zadanych opóźnieniach (TBATS oraz SARIMA). Minimalną oraz maksymalną wartość rozstępów (jako miary średnie) wybrano z powodów chęci poznania wartości rozstępów przedziału prognozy na początku i na końcu okresu prognozowania. Medianę uznano jako właściwą miarę dla asymetrycznych rozkładów wartości rozstępów przedziałów prognozy. We wszystkich przypadkach mediana rozstępów przedziałów prognoz modelu TBATS jest znacznie niższa od tej wartości z modelu SARIMA.

Tabela 3. Wybrane średnie rozstępów przedziałów prognoz TBATS oraz SARIMA

	α	TBATS			SARIMA		
		min	max	mediana	min	max	mediana
NAGLOWEK	0.80	136 014	778 727	493 091	183 432	3 635 982	2 751 268
	0.95	208 015	1 190 960	754 118	280 535	5 560 755	4 207 701
PODMIOT	0.80	136 014	778 727	493 091	183 432	3 635 982	2 751 268
	0.95	208 015	1 190 960	754 118	280 535	5 560 755	4 207 701
SUMA	0.80	478 740	2 580 477	1 677 098	628 892	9 020 229	6 390 598
	0.95	732 170	3 946 499	2 564 901	961 808	13 795 251	9 773 576
ZAKUP	0.80	11 218 718	54 824 266	37 224 413	15 262 871	247 836 399	186 512 903
	0.95	17 157 548	83 846 479	56 929 827	23 342 546	379 033 064	285 246 870
SPRZEDAZ	0.80	44 343 267	389 207 744	273 346 835	57 541 066	849 985 189	598 010 990
	0.95	67 817 174	595 241 879	418 047 909	88 001 467	1 299 940 168	914 578 886

Źródło: Obliczenia własne na podstawie modelowania szeregów czasowych aktualizacji tabel JPK_VAT (próba testowa: 18 maja 2019 r. – 22 lipca 2019 r.).

Znacząca różnica pomiędzy wartościami statystyk opisowych obliczonych na podstawie rozstępu przedziałów prognoz występuje zwłaszcza dla wartości maksymalnych i mediany. Mediana dla przedziału prognozy na poziomie ufności 0,95 w przypadku SUMA_KONTROLNA wynosi ok. 2,5 mln z modelu TBATS i jest znacznie niższa w porównaniu do SARIMA, gdzie wynosi ok. 9,7 mln. Wartość maksymalnego rozstępu dla przedziału prognozy na poziomie ufności 0,95 (model SARIMA) sięgała ok. 13,79 mln i dla aktualizacji tabeli SUMA_KONTROLNA była ponad trzykrotnie wyższa niż z modelu TBATS. Analogicznie w tabeli NAGLOWEK oraz PODMIOT maksymalny rozstęp przedziału prognozy na poziomie ufności 0,95 wynosił 1,19 mln w przypadku TBATS, natomiast SARIMA generowała rozstęp o wartości ponad 5,56 mln. Natomiast wartość minimalna z reguły jest wyższa o parę- lub kilkadziesiąt pro-

cent w przypadku rozstępu przedziału prognozy z modelu SARIMA. Wartość minimalna rozstępu przedziału prognozy (poziom ufności 0,95) w przypadku SUMA_KONTROLNA wyniosła ok. 732 tys. w przypadku modelu TBATS, a dla modelu SARIMA wartość tej statystyki wynosiła 961 tys.

Jednak należy podkreślić, że minimalne wartości rozstępu prognoz w każdym z przypadków były niższe, jeśli pochodziły z modelu TBATS. Również prognozy punktowe z modelu SARIMA można określić jako trafne na trzyprocentowym poziomie istotności. Analiza błędów MAPE zarówno dla próby uczącej (2), jak i testowej (8), a także po ich ilościowym porównaniu (tabela 1) wykazała, że maksymalnym błędem prognoz dla tabeli SPRZEDAZ był błąd modelu SARIMA. Również dla danych z dwóch innych tabel błąd modelu SARIMA był najwyższy (NAGLOWEK i PODMIOT).

Potwierdzono hipotezy postawione w artykule *Modelowanie i dekompozycja szeregów czasowych aktualizacji Jednolitych Plików Kontrolnych* z niniejszego zeszytu „Studiów Ekonomicznych. Zeszytów Naukowych Uniwersytetu Ekonomicznego w Katowicach”. **Przede wszystkim potwierdzono hipotezę, że najlepszymi właściwościami charakteryzuje się model TBATS**, który dzięki zastosowaniu trygonometrycznego modelowaniu składnika sezonowego generował szeregi niezależnych reszt dla wszystkich tabel.

Błędy MAPE z modelu TBATS były najniższe w okresie próby uczącej i średnio wynosiły 0,259% oraz 0,624% w okresie próby testowej. Pozostałe modele dawały wyższe wartości błędów MAPE. W przypadku SARIMA błędy w okresie uczącej wynosiły średnio 0,328%, a w okresie próby testowej 1,257%. Z kolei BATS, którego reszty nie były niezależne, generował dodatkowo błędy prognoz w okresie próby uczącej na poziomie średnio 0,306%, a w okresie próby testowej 1,213%.

Potwierdzono hipotezę o przewadze modelu TBATS nad pozostałymi branymi pod uwagę w tych badaniach. Model TBATS wykazał się najlepszymi właściwościami: brakiem autokorelacji w szeregach reszt, najniższymi błędami prognoz MAPE, najwęższymi przedziałami prognoz na dwóch stopniach istotności przy najmniejszej różnicy ich rozpiętości.

Podsumowanie

Reasumując niniejsze rozważania, należy podkreślić, że występują również inne, nieprezentowane w tym artykule podejścia prognostyczne, które można zastosować w przyszłości. Zwłaszcza gdy uznamy, że składnik sezonowy zmie-

nia się w czasie, przydatne mogą okazać się modele oparte na dekompozycji STL+ETS [Guibao i in., 2016, s. 2-3]. Zastosowanie modyfikacji już istniejących modeli (np. SARIMAX) będzie konieczne, gdy zaistnieją również inne przyczyny wpływające na zmienność przedziału prognozy [Melih, 2018, s. 2604-2606]. Inne wyzwania stawia przed prognozą wykorzystanie niestandardowych ustawień algorytmów i poleceń języka R w celu wyboru ostatecznej postaci modeli do prognozowania. Nie należy także zapominać o dodatkowych kryteriach (nie tylko informacyjnych), które mogą stanowić pomoc przy wstępnej lub ostatecznej selekcji.

Wiele potencjalnych wyborów zależy od możliwości prognozującego, w tym wiedzy, narzędzi, jakimi dysponuje, a także rodzajów praktycznych i teoretycznych zadań. Ciekawa wydaje się kontynuacja badań i zastosowanie kolejnych testów statystycznych, z których jednymi z najważniejszych są testy stabilności parametrów strukturalnych: Chowa oraz Nybloma i Hansena [Doman, Doman, 2009, s. 92-93]. Pierwszy z nich jest doskonale znany i zależy od subiektywnego podejścia prognosty, dwa kolejne nie wymagają i nie zależą od uznaniowego podziału próby – czyni to interesującymi z punktu widzenia badacza i możliwości wykonania dodatkowych porównań, skonfrontowania wyników [Doman, Doman, 2009, s. 92].

Literatura

- Ćwikliński K. (2020), *Modelowanie i dekompozycja szeregów czasowych aktualizacji jednolitych plików kontrolnych*, „Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach”, nr 390, podseria „Informatyka i Ekonometria”, nr 15, s. 60-75 (w druku – niniejszy zeszyt).
- De Livera A.M., Hyndman R.J., Snyder R.D. (2010), *Forecasting Time Series with Complex Seasonal Patterns Using Exponential Smoothing*, Department of Econometrics and Business Statistics, Working Paper, No. 15/09, <http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers> (dostęp: 6.05.2019 r.).
- Doman M., Doman R. (2009), *Modelowanie zmienności i ryzyka*, Oficyna a Wolters Kluwer business, Kraków.
- Guibao K., Yao H., Xin H., Xuan P., Min L., Chaoli H., Li G., Ping X., Dehua Y. (2016), *Epidemiological Analysis of Hemorrhagic Fever with Renal Syndrome in China with the Seasonal Trend Decomposition Method and the Exponential Smoothing Model*, “Scientific Reports”, No. 6, <https://www.nature.com/articles/srep39350.pdf> (dostęp: 1.05.2019).

Melih Y. (2018), *Forecasting Monthly Sales of White Goods Using Hybrid Arimax and Ann Models*, "Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi Aralık", Vol. 22(4), s. 2603-2617, <https://dergipark.org.tr/en/download/article-file/621691> (dostęp: 12.06.2019 r.).

Ustawa z dnia 11 marca 2004 r. o podatku od towarów i usług, Dz.U. z 2004 r., nr 54, poz. 535.

FORECASTING OF STANDARD AUDIT FILES FOR TAX (SAF-T) UPDATES

Summary: The forecasting of different time series became necessary process at the Ministry of Finance IT systems. The problems with lack of information and actual updates of Standard Audit Files for Tax are known. Capabilities to choosing right predicting model of time series with complex seasonal patterns are crucial in some cases. In the article, author made the decomposition of time series with complex seasonal patterns. The results of modeling and testing indicated the best predicting (according to Mean Absolute Percentage Error) and time series decomposition method – TBATS.

Keywords: Standard Audit Files for Tax (SAF-T), analyze of time series, decomposition, predicting, BATS, TBATS, SARIMA.