

Dekompozycja szeregów czasowych i zastosowanie wskaźników analizy technicznej

Stanisław Olek

Spis treści

1 Dekompozycja szeregów czasowych - usuwanie trendu i sezonowości	1
1.1 Przygotowanie i wstępna analiza danych	1
1.2 Dekompozycja szeregu	4
1.3 Dodanie transformacji Boxa-Coxa	17
1.4 Porównanie z różnicowaniem danych	21
2 Zastosowanie wskaźników analizy technicznej w analizie szeregów czasowych	24
2.1 Wybrane wskaźniki analizy technicznej	24
2.2 Wizualizacja i interpretacja wybranych wskaźników	25

1 Dekompozycja szeregów czasowych - usuwanie trendu i sezonowości

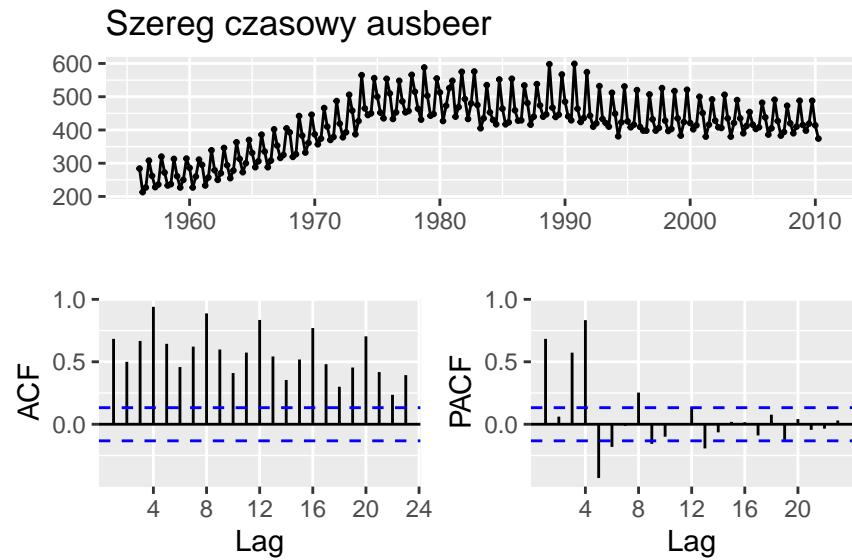
Pierwszym zagadnieniem, jakim się zajmiemy będzie usuwanie trendu i sezonowości w szeregu czasowym. Zobaczmy, na przykładzie, kilka metod dekompozycji szeregu czasowego:

- dekompozycja na podstawie ruchomej średniej,
- dekompozycja na podstawie modelu regresji,
- dekompozycja STL oparta na metodzie loess.

Dodatkowo, do każdej z metod dekompozycji zastosujemy transformację Boxa-Coxa i zobaczymy, czy ma ona istotny wpływ na wyniki. Na koniec porównamy wyniki eliminacji trendu i sezonowości z tymi uzyskanymi przez różnicowanie danych.

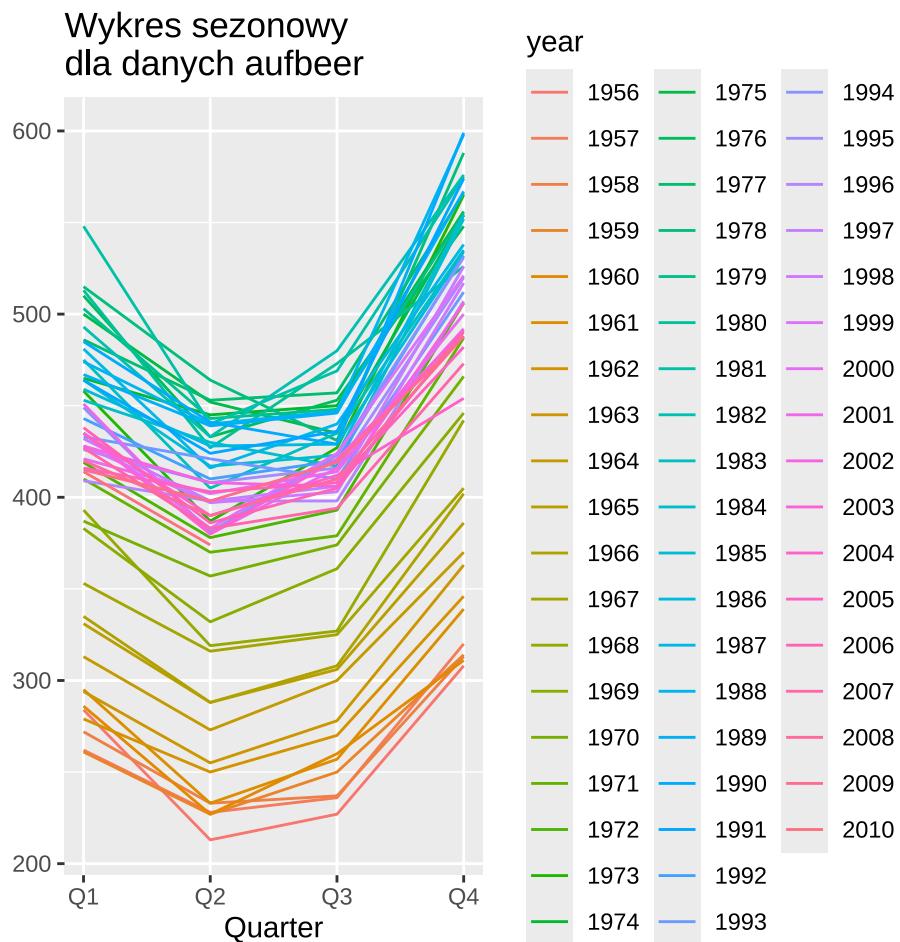
1.1 Przygotowanie i wstępna analiza danych

Jako zbiór danych weźmiemy zbiór `ausbeer` z biblioteki `fpp2`. Dane te dotyczą całkowitej kwartalnej produkcji piwa w Australii (w megalitrach) w latach 1956 - 2010. Na początek zobaczymy wykres tych danych, razem z wykresami ACF i PACF.



Rysunek 1: Wykres danych oraz funkcji autokorelacji i cząstkowej autokorelacji

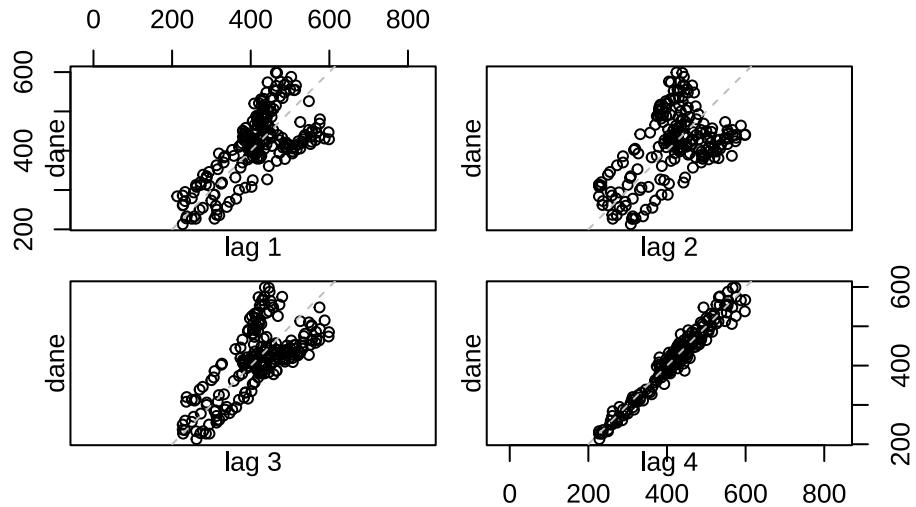
Na wykresach 1 możemy dostrzec trend w latach 1956 - 1975, który nie był tak wyraźny w późniejszym okresie. Prawdopodobnie jest to trend kwadratowy, a w późniejszym okresie możemy mieć do czynienia z trendem liniowym. Wykres ACF pokazuje nam także sezonowość (jest on okresowy). Dodatkowo, możemy stwierdzić, że **ausbeer** nie jest białym szumem (nie jest nawet szeregiem stacjonarnym). Aby dokładniej stwierdzić, jaki okres ma sezonowość dla danych **ausbeer** możemy popatrzeć na wykres sezonowy 2.



Rysunek 2: Wykres sezonowy dla danych aufbeer

Po wykresie sezonowym możemy stwierdzić, że mamy do czynienia z sezonowością o okresie 4 (bo dane są podzielone na kwartały, a nie na miesiące). Jeszcze lepiej pokazuje to wykres 3.

Wykresy rozrzutu dla opóźnienia



Rysunek 3: Wykres rozrzutu dla opóźnienia: widać wyraźnie sezonowość o okresie 4

Dodatkowo, na podstawie wykresu 1 zauważamy, że w danych nie ma obserwacji odstających.

1.2 Dekompozycja szeregu

Zastosujemy teraz kilka metod dekompozycji szeregu, aby zobaczyć, która z nich daje najlepszy rezultat. Analizować będziemy 3 metody:

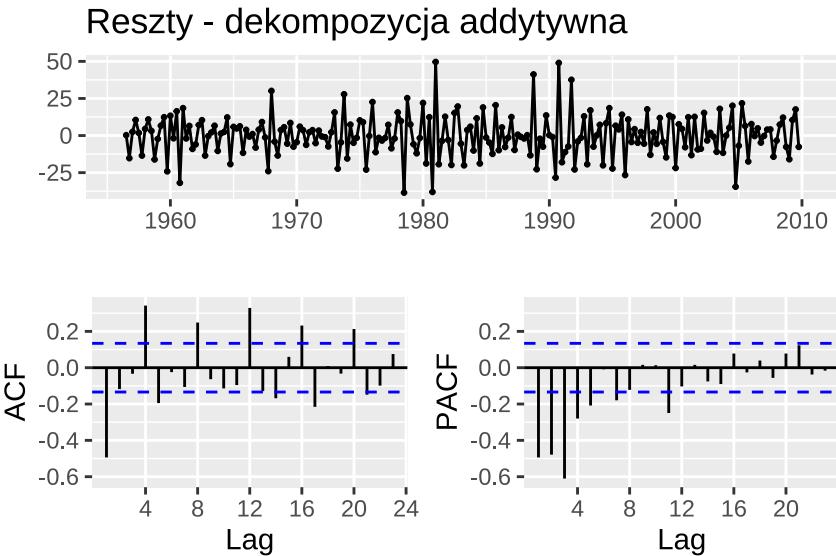
1. dekompozycja na podstawie ruchomej średniej 1.2.1
2. dekompozycja na podstawie modelu regresji 1.2.2
3. dekompozycja STL opartą na metodzie loess 1.2.3

Najpierw zajmiemy się dekompozycją na podstawie ruchomej średniej.

1.2.1 Dekompozycja na podstawie ruchomej średniej

W przypadku tej dekompozycji zajmiemy się dwoma rodzajami - dekompozycją addytywną i mnożycielną. Do dekompozycji na podstawie ruchomej średniej wykorzystujemy funkcję `decompose`.

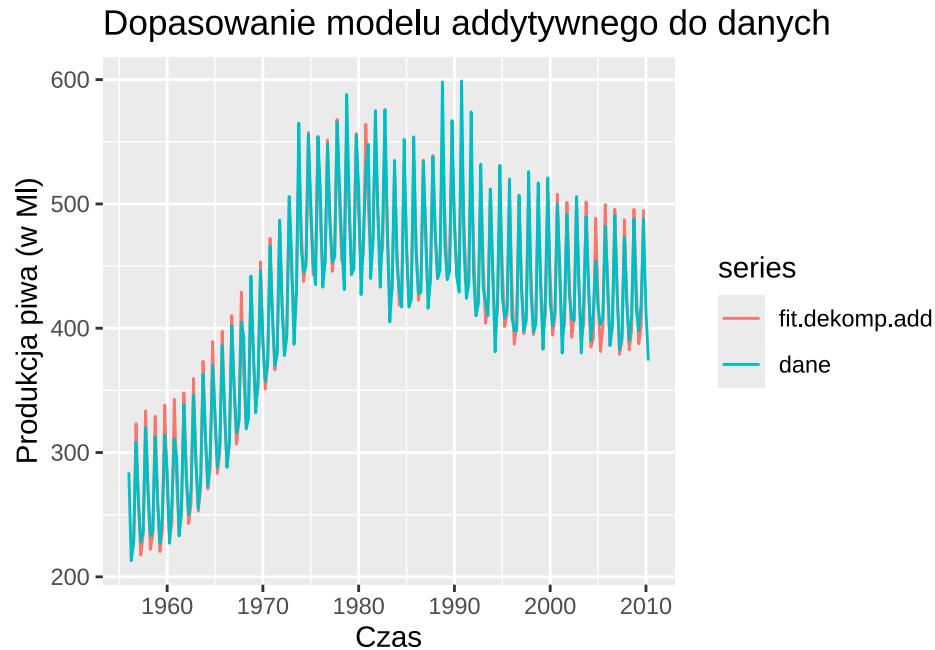
Sprawdzimy szereg reszt z obu metod, aby stwierdzić, czy po usunięciu trendu i sezonowości mamy do czynienia z szeregiem stacjonarnym.



Rysunek 4: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji addytywnej

Z wykresu 4 możemy wywnioskować, że udało nam się pozbyć trendu i sezonowości (ACF nie zanika i nie wygląda na okresową). Dodatkowo, wartości odstające nie wyróżniają się na tyle, żebyśmy mogli stwierdzić, że wariancja nie jest jednorodna. Stąd możemy stwierdzić, że szereg reszt ma stacjonarny charakter. Nie jest to jednak biały szum, gdyż za dużo obserwacji funkcji ACF odstaje poza zaznaczony przedział ufności.

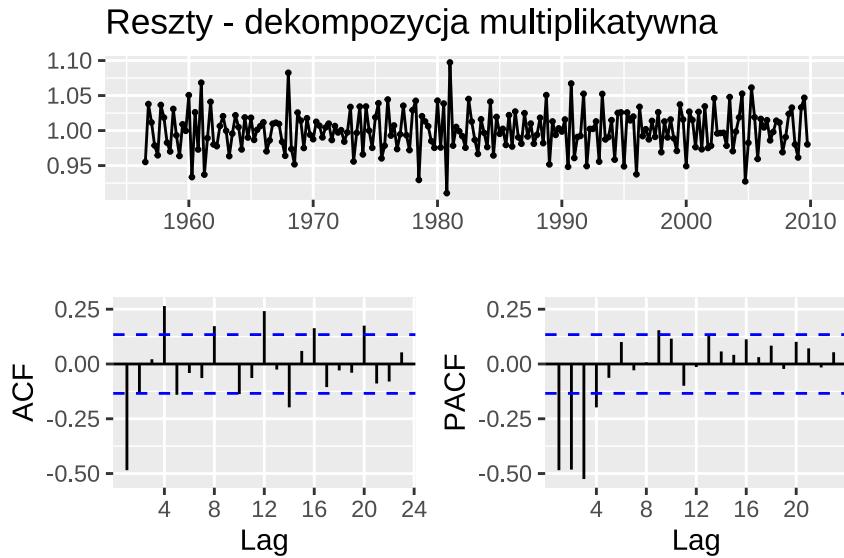
Możemy zobaczyć, czy model został dobrze dopasowany odtwarzając wykres naszych danych.



Rysunek 5: Sprawdzenie dopasowania modelu addytywnego do danych

Po rysunku 5 widzimy, że model addytywny nie dopasował się do danych. Na początku i na końcu zbioru danych wiele wartości odstaje od oryginalnych danych. W środku wykres jest trochę lepiej.

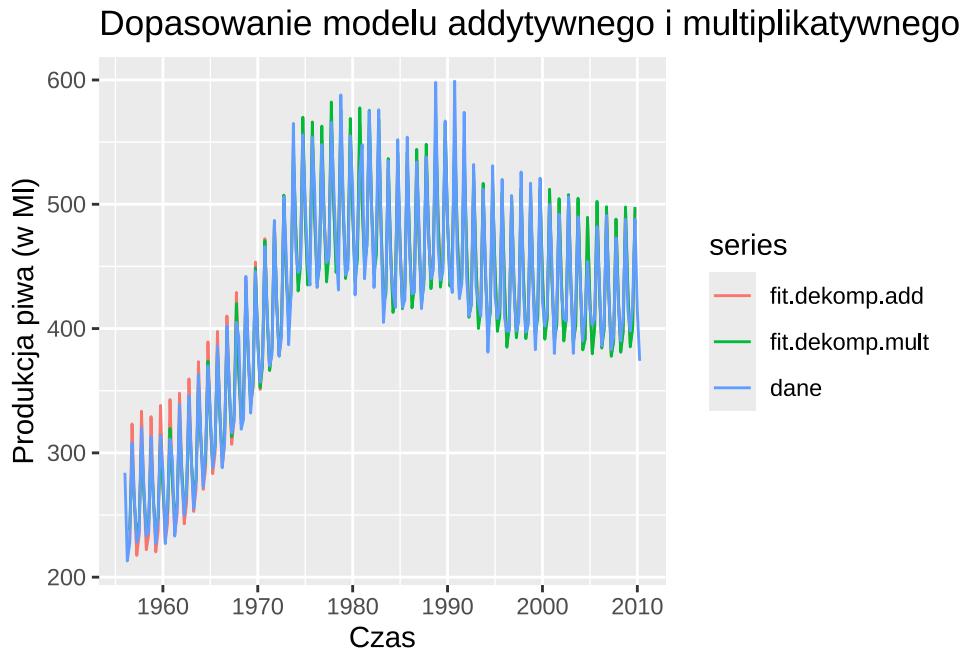
Zobaczmy, jak poradzi sobie model mnożnikowy. Wykres losowych reszt oraz ich ACF i PACF jest przedstawiony na wykresie 6.



Rysunek 6: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji mnożnikowej

Z wykresu 6 możemy wyciągnąć podobne wnioski, jak dla dekompozycji addytywnej (brak trendu, niewidoczna sezonowość, prawdopodobna stacjonarność).

Zobaczmy, czy model mnożnikowy dopasował się lepiej do danych niż model addytywny. Na wykresie 7 mamy porównanie obu modeli względem oryginalnych danych.



Rysunek 7: Sprawdzenie dopasowania modelu addytywnego i mnożnikowego do danych

Widzimy, że model mnożnikowy zdecydowanie lepiej dopasował się do danych od modelu addytywnego, jednak po roku 1970 widoczne jest odstanie od oryginalnych danych. Ogólnie, oba modele usunęły trend i

zniwelowały mocno sezonowość, ale nie dopasowały się do danych tak, jakbyśmy oczekiwali. Zobaczmy, czy metoda dekompozycji na podstawie modelu regresji będzie pod tym względem lepsza.

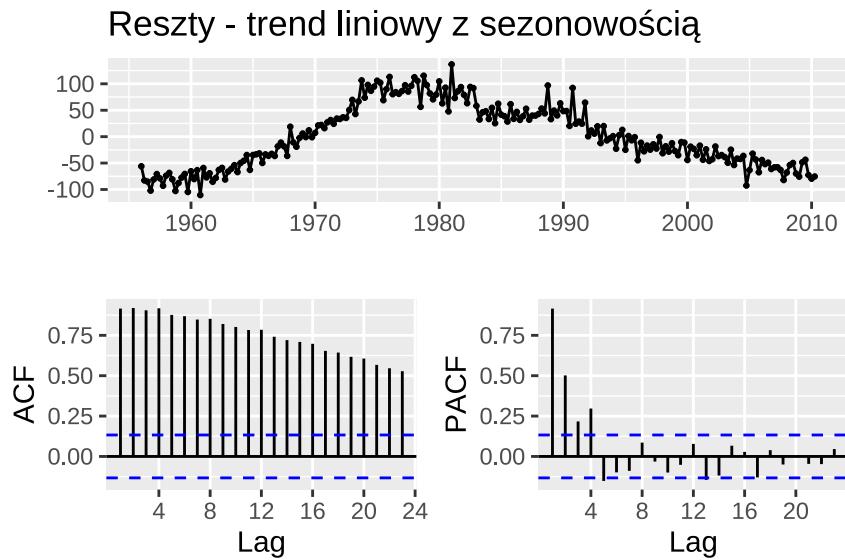
1.2.2 Dekompozycja na podstawie modelu regresji

Dla dekompozycji na podstawie modelu regresji przeprowadzimy podobną analizę. W tym przypadku na podstawie wykresu 1 możemy zdecydować się, jaki model zastosować dla trendu. Wcześniej, przy 1, stwierdziliśmy, że możemy mieć do czynienia z trendem kwadratowym i liniowym. Być może zatem odpowiedni będzie wielomian wyższego stopnia, który nam ujmie trend w latach 1956-1970 i po roku 1970. My zbadamy kilka modeli (ozn. trend - m_t , sezonowość - s_t):

- $m_t = a + b \cdot t$ z s_t ,
 - $m_t = a + b \cdot t + c \cdot t^2$ z s_t ,
 - m_t będzie wielomianem stopnia 5. z s_t .

Wpierw zajmiemy się modelem z trendem liniowym. Najpierw, sprawdźmy, co możemy wstępnie powiedzieć o tej metodzie z funkcji `summary`.

Możemy stąd podać współczynniki, które występują przy trendzie: $m_t = 339.3848485 + 0.7103976 \cdot t$. Dodatkowo, wszystkie współczynniki, łącznie z tymi przy sezonowości, są istotne przy poziomie istotności $\alpha = 0.05$. Współczynnik R^2 jest dosyć niski i wynosi 52.72195%. Podobnie, jak wcześniej zobaczymy wykresy reszt, aby zobaczyć, czy model pozbył się trendu i sezonowości.

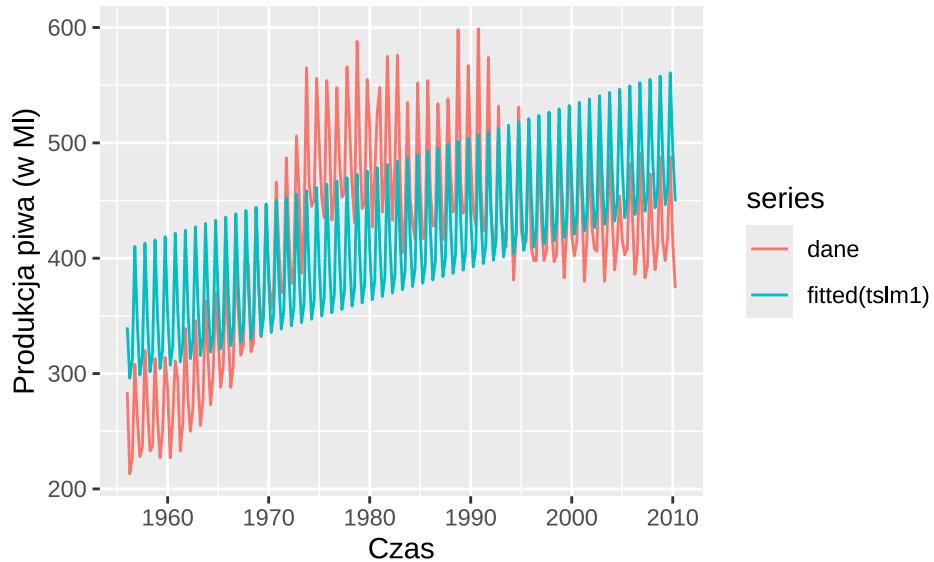


Rysunek 8: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji na podstawie modelu regresji z liniowym trendem i z sezonowością

Po przyjrzeniu się wykresowi 8 możemy wywnioskować, że trend nie został wyeliminowany, więc trend liniowy to za mało. Szczególnie widać to po zanikającej funkcji ACF. Sezonowość została jednak w większości wyeliminowana, jednak dalej widać delikatną okresowość ACF. Z powodu widocznego trendu na funkcji ACF można stwierdzić, że nie mamy do czynienia z szeregiem stacjonarnym, a tym bardziej z białym szumem.

Teraz sprawdzimy, jak model dopasował się do danych na wykresie 9.

Dopasowanie trendu liniowego z sezonowością do danych

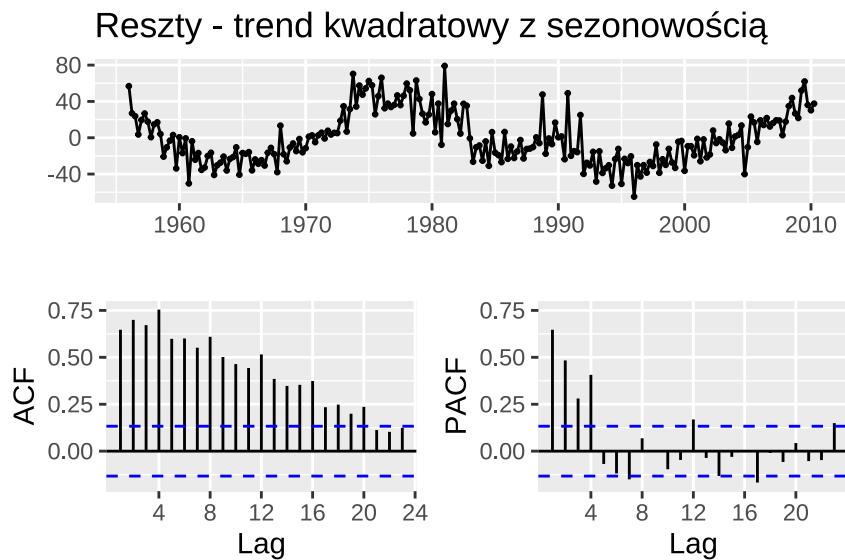


Rysunek 9: Sprawdzenie dopasowania modelu regresji dla trendu liniowego z sezonowością do danych

Model nie uchwycił wykrzywienia danych, jednak sezonowość jest na w miarę odpowiednim poziomie. Jest to jedyny pozytyw dopasowania modelu. W ogólności, model źle się dopasował do danych.

Kolejnym modelem regresji, który wykorzystamy dla trendu, będzie model, w którym przyjmujemy trend kwadratowy. Podobnie, jak wcześniej zobaczymy, co wywnioskujemy z funkcji **summary**.

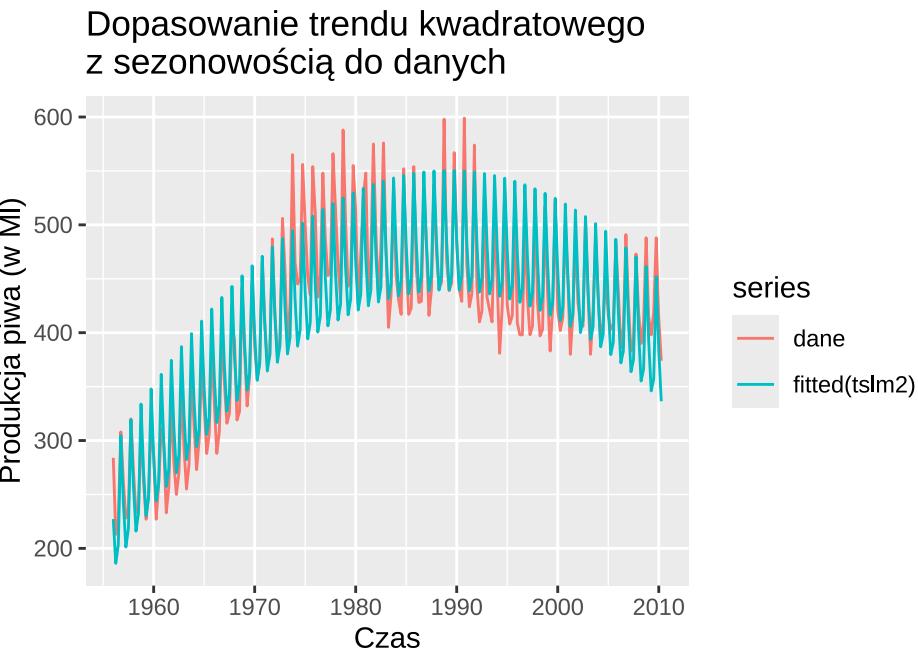
W tym przypadku trend ma postać $m_t = 418.223928 + 660.071536 \cdot t + (-763.0716051) \cdot t^2$. Wszystkie współczynniki zarówno przy trendzie, jak i przy sezonowości są istotne na poziomie istotności 0.05. Otrzymaliśmy też o wiele wyższy współczynnik R^2 , który wynosi 89.0880961%, co oznacza, że ten model jest już lepiej dopasowany do danych. Zobaczmy na wykresie reszt 12, czy udało się nam pozbyć trendu.



Rysunek 10: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji na podstawie modelu regresji z trendem kwadratowym i z sezonowością

Z wykresu 10 możemy wywnioskować, że w tym przypadku niestety też nie udało nam się do końca zniwelować trendu (zanikający wykres ACF). Tak samo, jak na 8, widać delikatne pozostałości po sezonowości. Z tych dwóch powodów możemy stwierdzić, że nie mamy do czynienia z szeregiem stacjonarnym, w tym także z białym szumem.

Sprawdźmy teraz dopasowanie tego modelu do danych.



Rysunek 11: Sprawdzenie dopasowania modelu regresji liniowej dla trendu kwadratowego z sezonowością do danych

Po wykresie 11 możemy stwierdzić, że model regresji z trendem kwadratowym zdecydowanie lepiej odwzorował

dane niż model z trendem liniowym, jednak dalej jest to słabe dopasowanie. Możemy jednak stwierdzić, że większy stopień wielomianu dał lepsze dopasowanie.

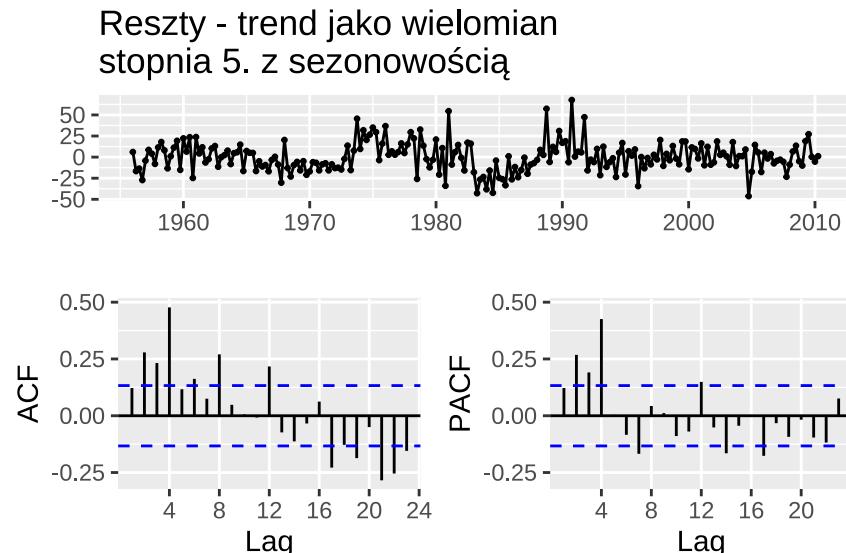
Ostatnim modelem regresji, jaki zbadamy będzie m_t jako wielomian stopnia 5 z sezonowością. Zrobimy taki duży krok, aby jak najlepiej dopasować dane i zobaczyć, czy uda nam się przy dosyć dużym stopniu wielomianu usunąć trend. Najpierw sprawdzimy, jakie informacje możemy wyciągnąć z funkcji `summary`.

W tym modelu

$$m_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5,$$

gdzie $a_0 = 417.7477184$, $a_1 = 660.0667094$, $a_2 = -762.9435811$, $a_3 = 164.0192627$, $a_4 = 226.8288624$, $a_5 = -170.7502984$. Współczynnik R^2 wzrósł nam do 95.8029503%, więc znowu nam znacznie wzrósł. Dodatkowo, wszystkie współczynniki są istotne (być może można było взять jeszcze wyższy stopień wielomianu).

Na wykresie 12 możemy przyjrzeć się resztom, jakie zostaną po usunięciu trendu i sezonowości.

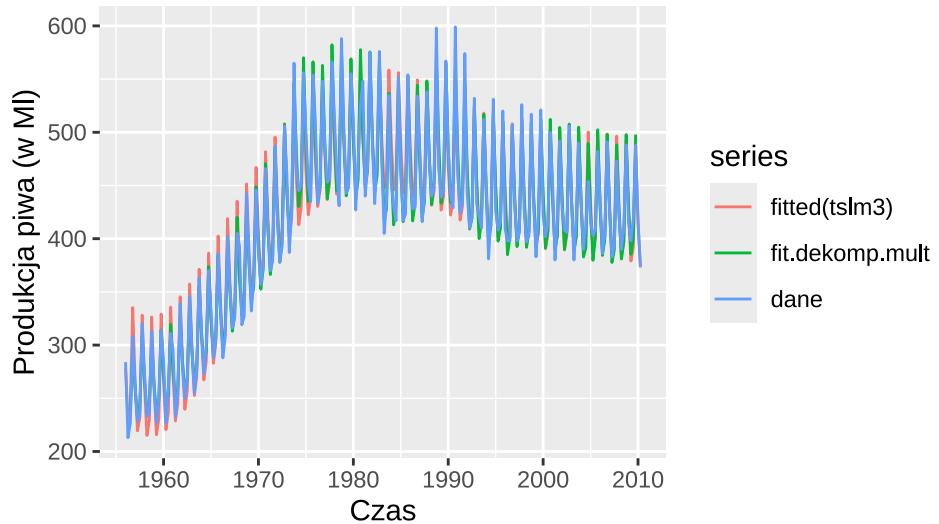


Rysunek 12: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji na podstawie modelu regresji z trendem jako wielomian stopnia 5. i z sezonowością

Na wykresie 12 możemy dostrzec jeszcze niewielkie pozostałości po trendzie, jednak ACF już nie zanika, co jest dobrym sygnałem. Problemem jest jednak brak jednorodnej wariancji. Sezonowość została w podobny sposób zniwelowana, co dla 10 i 12. Nie jest to raczej szereg stacjonarny, tym bardziej biały szum, ale być może jeszcze większy stopień wielomianu pozwoliłby nam ostatecznie dojść do stacjonarności.

Widzimy zatem, że pod względem usunięcia sezonowości i trendu ten model poradził sobie najlepiej ze wszystkich dotychczasowych modeli regresji. Nie wydaje się jednak, żeby poradził sobie lepiej niż modele oparte na metodzie dekompozycji na podstawie ruchomej średniej. W celu porównania tych obu metod dekompozycji sprawdzimy na wykresie 13, czy lepiej do danych dopasował się model dekompozycji na podstawie modelu regresji z trendem jako wielomian stopnia 5. i z sezonowością, czy model mnożnikowy dekompozycji na podstawie ruchomej średniej.

Porównanie modelu regresji z trendem jako wielomian stopnia 5. z modelem multiplikatywnym



Rysunek 13: Porównanie dopasowania regresja z trendem jako wielomian stopnia 5. vs. model multiplikatywny dekompozycji opartej na podstawie ruchomej średniej do danych

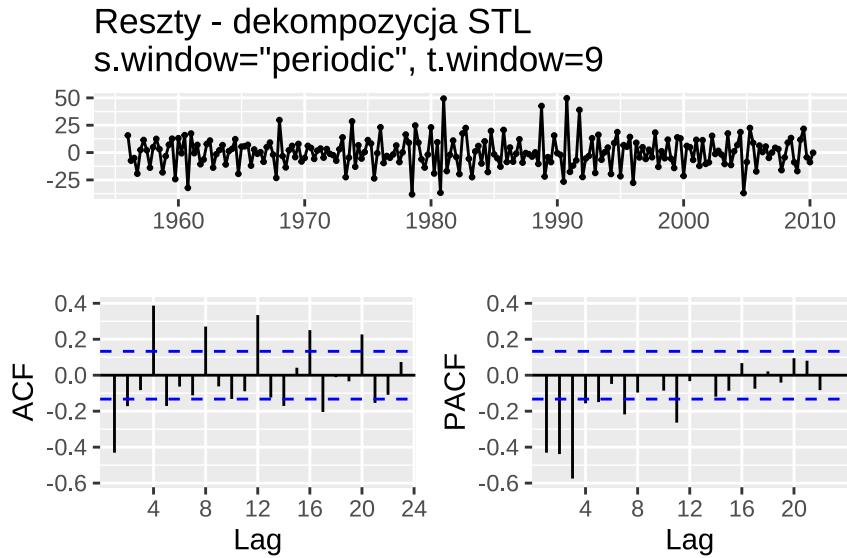
Widzimy, że lepiej do danych dopasował się model multiplikatywny. Stąd możemy stwierdzić, że lepiej będzie użyć metody dekompozycji na podstawie ruchomej średniej.

1.2.3 Dekompozycja STL oparta na metodzie loess

Ostatnią metodą dekompozycji, jaką zbadamy będzie dekompozycja STL oparta na metodzie loess. Do wykonania dekompozycji za pomocą tej metody posłuży nam funkcja `stl`. W tym przypadku zbadamy 4 modele, gdzie w każdym ustawimy 2 parametry - `s.window` - parametr mówiący o okresie dla okna loess, w którym będzie wyciągana sezonowość, `t.window` - parametr mówiący o okresie dla okna loess, w którym będzie wyciągany trend. Dokładniej rozważymy następujące zestawy parametrów:

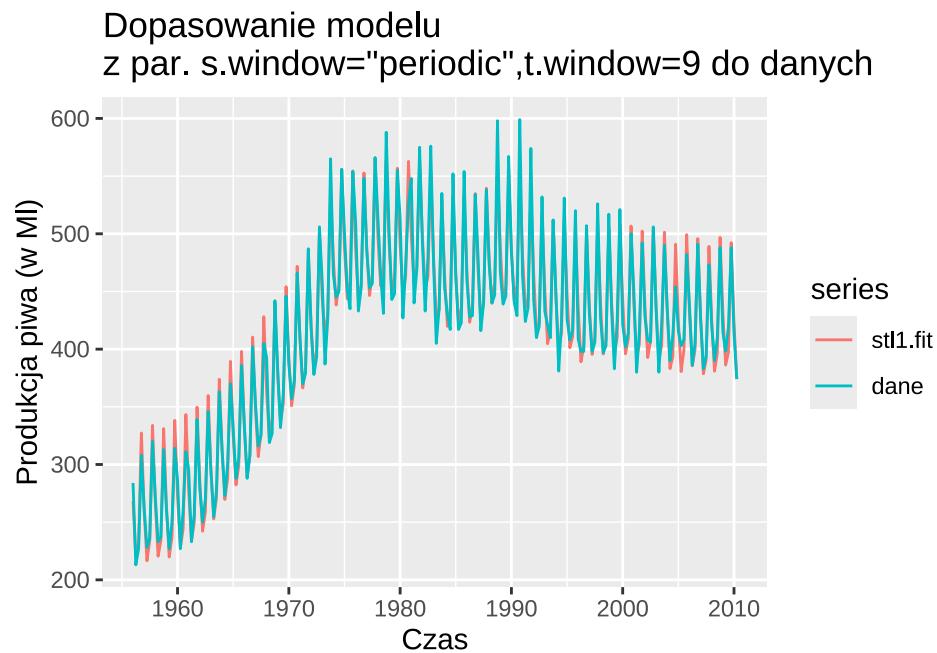
1. `s.window = 'periodic', t.window = 9,`
2. `s.window = 'periodic', t.window = 5,`
3. `s.window = 13, t.window = 9,`
4. `s.window = 13, t.window = 5.`

Najpierw rozważymy 1. Na wykresie 14 są pokazane reszty i funkcje ACF oraz PACF.



Rysunek 14: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji STL z par. `s.window = 'periodic'`, `t.window = 9`

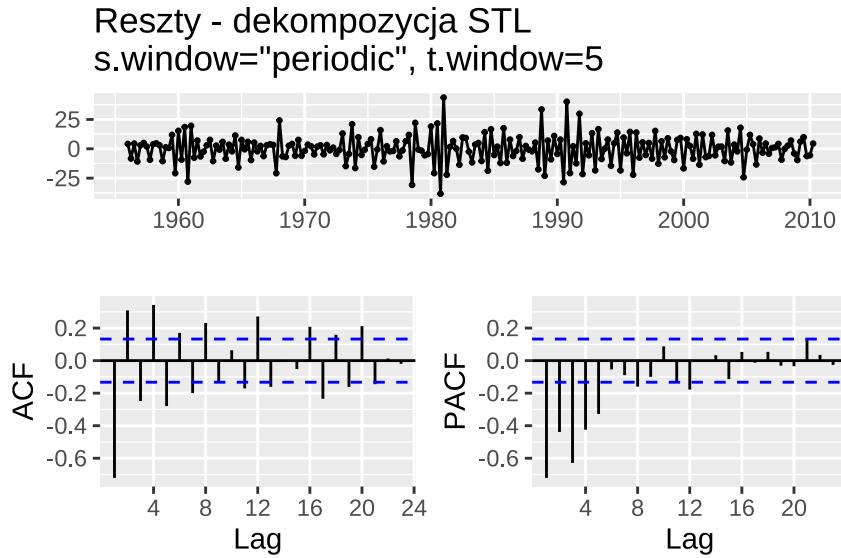
Wykres reszt wygląda znacznie lepiej niż w przypadku dekompozycji na podstawie modelu regresji, jeśli chodzi o trend, który na wykresie 14 jest niezauważalny. Większy problem jest z sezonowością, gdyż wydaje się, że funkcja ACF jest okresowa. Tak samo występują odchylenia w wartościach, co może świadczyć o braku jednorodności wariancji. W celu zbadania stacjonarności tego najlepiej byłoby przeprowadzić odpowiedni test, gdyż z rysunku 14 ciężko jest jednoznacznie wywnioskować, czy szereg jest stacjonarny, czy nie. Można za to jednoznacznie powiedzieć, że nie jest to biały szum (za dużo obserwacji odstaje poza przedział ufności). Zobaczmy wykres dopasowania modelu do danych.



Rysunek 15: Sprawdzenie dopasowania dekompozycji STL z par. `s.window = 'periodic'`, `t.window = 9` do danych

Widzimy, że model ma problem z dopasowaniem się do danych, szczególnie na początku i na końcu wykresu 15. Trend został odpowiednio uchwycony przez model, jednak w przypadku sezonowości są wady w dopasowaniu modelu. Zobaczmy, czy inny parametr `t.window` lub `s.window` coś zmieni.

Najpierw, zmienimy parametr `t.window` na 5 (2). W tym przypadku szereg reszt prezentuje się następująco:

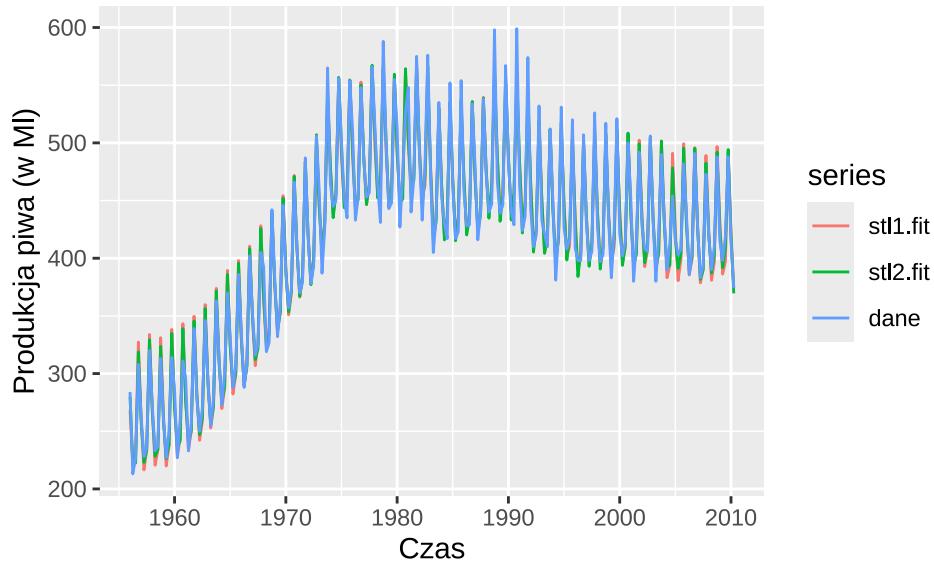


Rysunek 16: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji STL z par. `s.window = 'periodic'`, `t.window = 5`

W przypadku rysunku 16 na wykresie ACF nie da się zauważyc sezonowości (brak wyraźnej okresowości). Problemem tu jest jednak większe “ściśnięcie” wszystkich wartości, z wyjątkiem wartości odstających, przez co wartości odstające są jeszcze bardziej zauważalne niż w przypadku 14. Poza tym, wydaje się, że sezonowość i trend zostały usunięte. W związku z tym, bardziej prawdopodobne jest, że ten szereg jest stacjonarny niż szereg reszt dla 1. Nie jest to na pewno biały szum, ponieważ za dużo obserwacji ACF odstaje poza przedział ufności.

Sprawdzimy teraz, jak model dopasował się do danych.

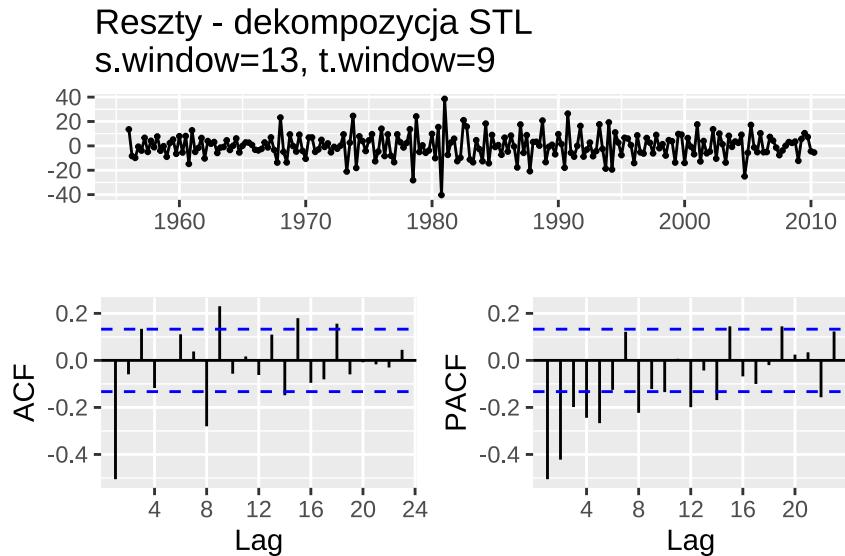
Dopasowanie modelu z par. s.window="periodic", t.window=9 do danych



Rysunek 17: Sprawdzenie dopasowania dekompozycji STL z par. `s.window = 'periodic'`, `t.window = 5` i `s.window = 'periodic'`, `t.window = 9` do danych

Widzimy na 17, że model z parametrami 2 trochę lepiej dopasował się do danych niż model z par. 1, jednak jest to niewielka różnica. Oba modele mają ten sam problem - gorsze dopasowanie na początku i na końcu danych.

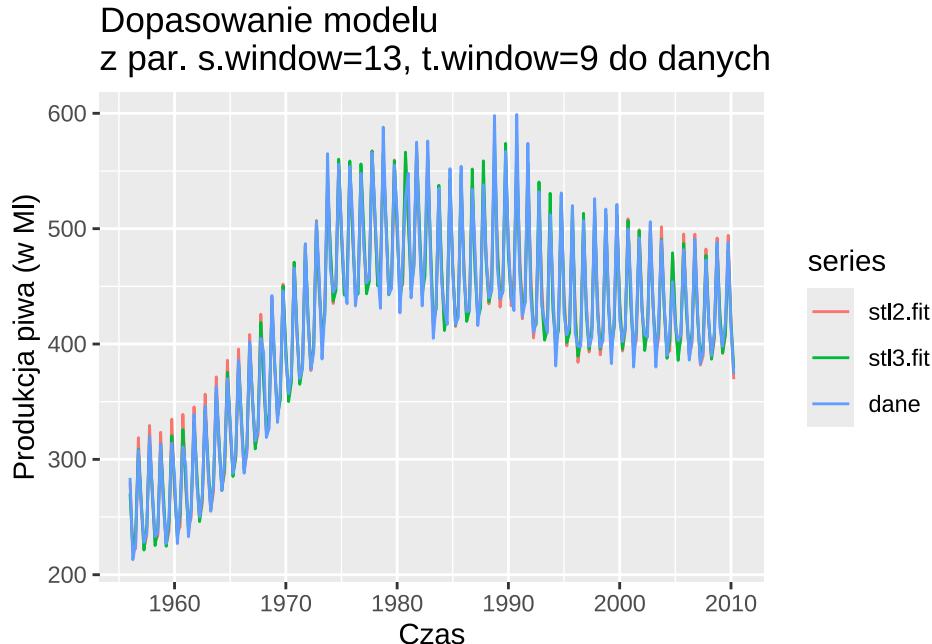
Ten problem może rozwiązać model z par. z 3. Najpierw jednak sprawdzmy, czy reszty, jakie pozostaną po usunięciu trendu i sezonowości w tym modelu będą szeregiem stacjonarnym.



Rysunek 18: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji STL z par. `s.window = 13`, `t.window = 9`

Z wykresu 18 możemy wywnioskować, że podobnie, jak dla 1 i 2 niewidoczny jest trend. W porównaniu z tamtymi modelami jednak, tutaj ACF nie jest na pewno okresowa, a wariancja wydaje się stabilniejsza. Prawdopodobnie, szereg reszt z rysunku 18 jest zatem szeregiem stacjonarnym. Nie jest to jednak biały szum z tego samego powodu, co 14 i 16.

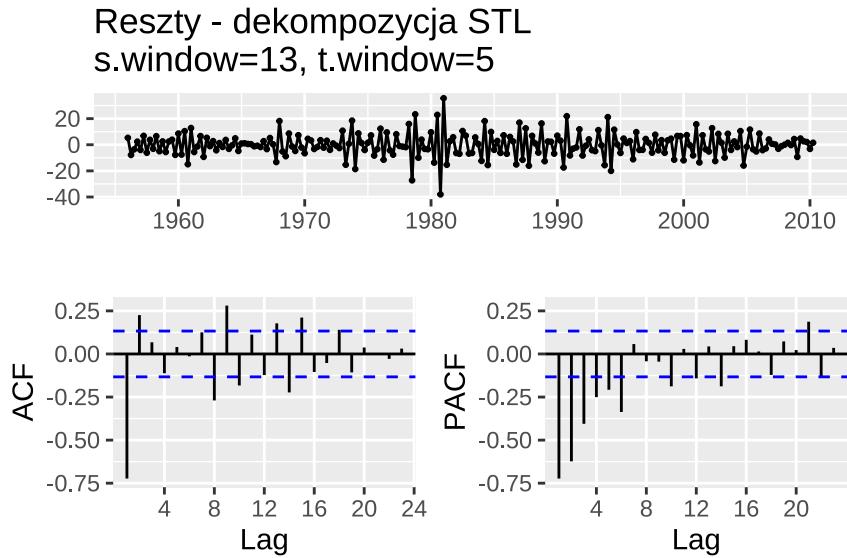
Sprawdźmy jakość dopasowania modelu z par. 3 na tle modelu z par. 2.



Rysunek 19: Sprawdzenie dopasowania dekompozycji STL z par. `s.window = 'periodic'`, `t.window = 9` i `s.window = 13`, `t.window = 9` do danych

Wykresy na 19 bardzo ciężko ze sobą porównać - to wskazuje, że nie ma dużych różnic w dopasowaniu obu modeli do danych. Po bliższym przyjrzeniu się jednak możemy zauważyć, że model z parametrami `s.window = 13`, `t.window = 9` jest lepiej dopasowany do danych (zdecydowanie bliższe dopasowanie do danych na początku wykresu, w większości punktów jest bliżej oryginalnych danych).

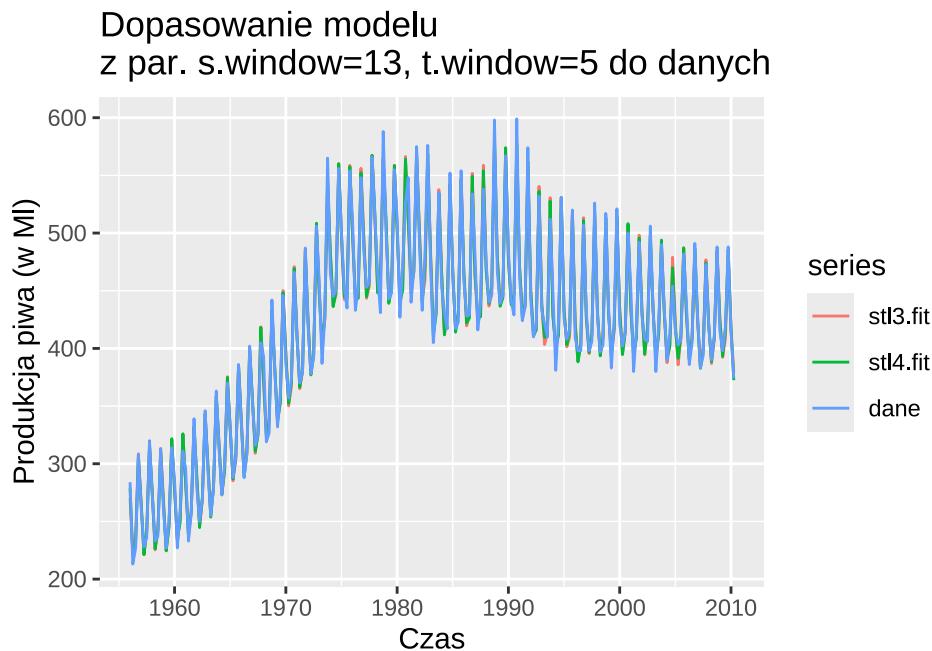
Pozostaje sprawdzić, czy zmiana parametru `t.window` na 5 poprawi jeszcze bardziej dopasowanie lub zbliży szereg reszt do bycia szeregiem stacjonarnym. Najpierw, zbadamy szeregi reszt dla modelu 4.



Rysunek 20: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji STL z par. `s.window = 13, t.window = 5`

Na wykresie 20 widać, że model z parametrami `s.window = 13, t.window = 5` ma bardziej spłaszczony wykres reszt dla obserwacji, które mniej odstają niż model z parametrami `s.window = 13, t.window = 9`, przez co wartości odstające są bardziej widoczne. Poza tym, wnioski dla tego szeregu są takie same. W takim razie to też prawdopodobnie jest szereg stacjonarny, ale nie jest to biały szum (ten sam powód, co dla 1, 2 i 3).

Porównajmy dopasowanie modeli z parametrami `s.window = 13, t.window = 5` i `s.window = 13, t.window = 9` do oryginalnych danych (21).



Rysunek 21: Sprawdzenie dopasowania dekompozycji STL z par. `s.window = 13, t.window = 5` i `s.window = 13, t.window = 9` do danych

Cieźko stwierdzić, który model lepiej się dopasowuje do danych, co tylko świadczy o podobieństwie obu modeli. Możemy jednak zauważyc przy uważnym przypatrzeniu się, że model z par. `s.window = 13`, `t.window = 5` trochę lepiej przybliża oryginalne dane. Stąd wynika, że najlepszą analizowaną przez nas dekompozycją STL jest dekompozycja z parametrami `s.window = 13`, `t.window = 5`.

1.3 Dodanie transformacji Boxa-Coxa

Kolejnym krokiem, jaki podejmiemy, będzie zastosowanie transformacji Boxa-Coxa na kilku z przedstawionych przez nas modeli. Przypomnijmy, że transformacja Boxa-Coxa jest dana wzorem:

$$f_\lambda(x) = \begin{cases} \frac{x^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \ln x, & \lambda = 0 \end{cases}$$

Transformację Boxa-Coxa zastosujemy na następujących modelach:

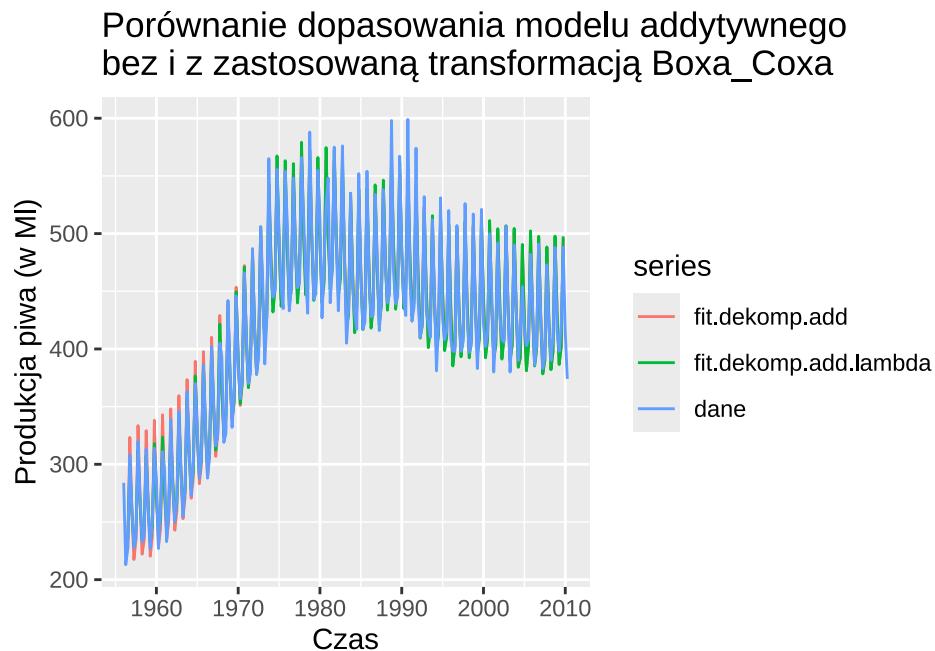
- model addytywny (`decompose`),
- model multiplikatywny (`decompose`),
- model regresji z trendem jako wielomian stopnia 5. i z sezonowością (`tslm`),
- model dekompozycji STL z parametrami `s.window = 'periodic'`, `t.window = 9` (`stl`),
- model dekompozycji STL z parametrami `s.window = 13`, `t.window = 5` (`stl`).

Sprawdzać będziemy tylko jakość dopasowania modeli do danych, więc pominiemy analizę szeregu reszt. Do dopasowania odpowiedniego parametru λ wykorzystamy funkcję `BoxCox.lambda`. Wynosi ona 0.2046987.

Uwaga! 1 Pokazane wykresy ze względu na małą widoczność różnic zostały wcześniej przeanalizowane przez autora w większym przybliżeniu, aby uchwycić wszelkie różnice pomiędzy wykresami.

1.3.1 Model addytywny

Najpierw porównamy dopasowanie modelu addytywnego. Na wykresie 22 pokazane jest dopasowanie tego modelu bez transformacji Boxa-Coxa i z transformacją Boxa-Coxa.

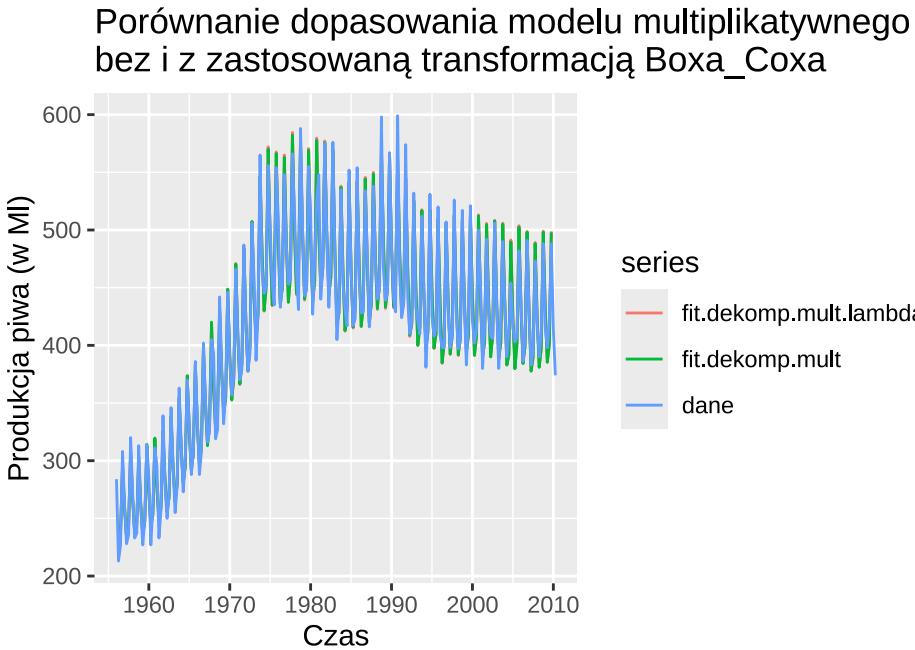


Rysunek 22: Dopasowanie modelu addytywnego: brak Boxa-Coxa vs. Box-Cox

Transformacja Boxa-Coxa poprawiła dopasowanie na początku danych, ale w środku i na końcu jest gorzej. Ogółem, transformacja Boxa-Coxa nie poprawiła znacznie dopasowania do danych.

1.3.2 Model multiplikatywny

Teraz sprawdzimy drugi model uzyskany za pomocą funkcji `decompose` - multiplikatywny. Na wykresie 23 widzimy różnicę dopasowania, gdy zastosujemy wcześniej na danych transformację Boxa-Coxa.



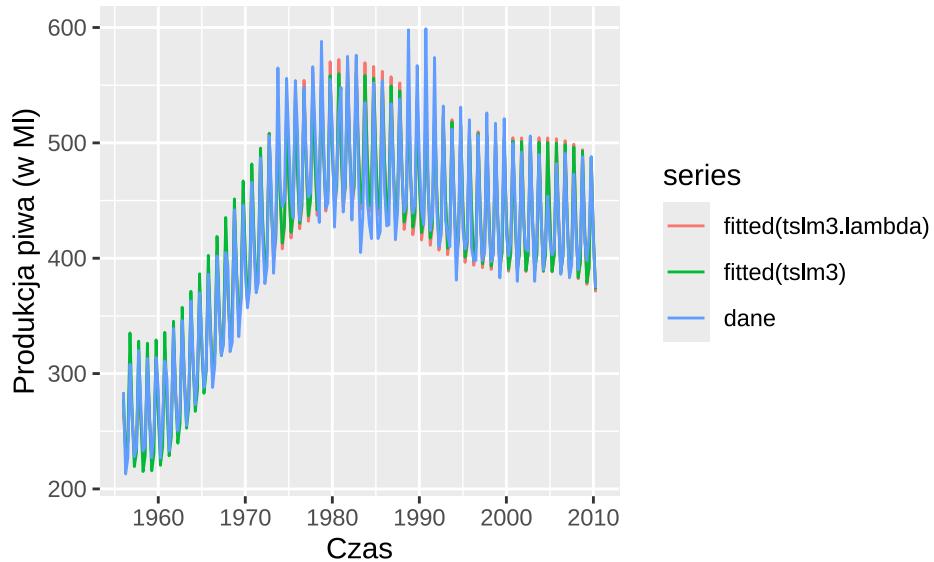
Rysunek 23: Dopasowanie modelu multiplikatywnego: brak Boxa-Coxa vs. Box-Cox

Po wykresie 23 możemy stwierdzić, że zastosowanie transformacji Boxa-Coxa nie poprawiło, ani nie pogorszyło dopasowania modelu multiplikatywnego w wyraźny sposób.

1.3.3 Model regresji z trendem jako wielomian stopnia 5. i z sezonowością

Kolejnym punktem będzie sprawdzenie jakości dopasowania modelu regresji z transformacją Boxa-Coxa i bez niej. Na wykresie 24 widzimy porównanie dopasowania obu modeli.

Porównanie dopasowania modelu tslm bez i z transformacją Boxa-Coxa



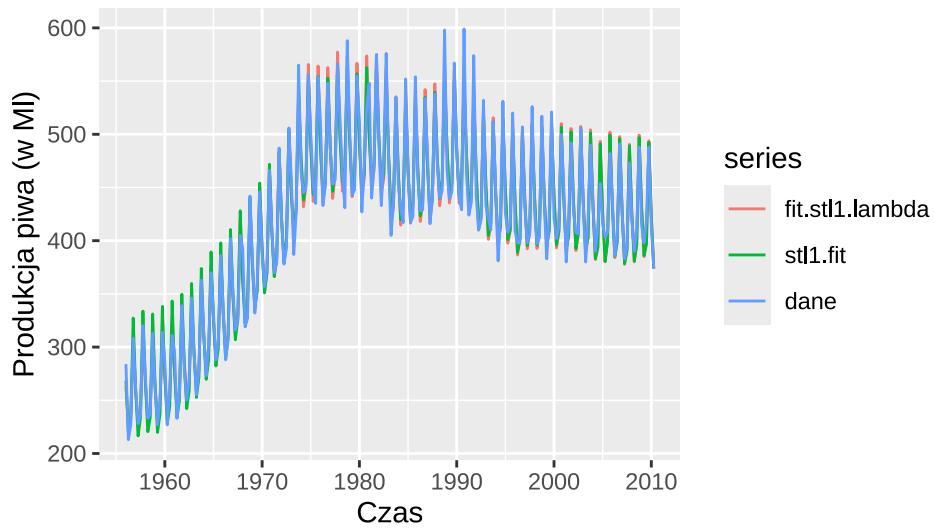
Rysunek 24: Dopasowanie modelu regresji: brak Boxa-Coxa vs. Box-Cox

Zastosowanie transformacji Boxa-Coxa w tym przypadku polepszyło dopasowanie na początku danych, natomiast w środku i przy końcu danych trochę pogorszyło jakość dopasowania. Ogółem, można uznać, że transformacja Boxa-Coxa doprowadziła do minimalnie lepszego dopasowania.

1.3.4 Dekompozycja STL z parametrami `s.window = 'periodic'`, `t.window = 9`

Kolejnym rodzajem dekompozycji, przy którym dokonamy transformacji Boxa-Coxa będzie dekompozycja STL oparta na metodzie loess. Najpierw zastosujemy przekształcenie dla modelu z parametrami `s.window = 'periodic'`, `t.window = 9`. Na wykresie 25 widać porównanie modelu bez transformacji Boxa-Coxa i z tą transformacją.

Porównanie dopasowania modelu STL z par. s.window="periodic", t.window=9 bez i z transformacją Boxa-Coxa



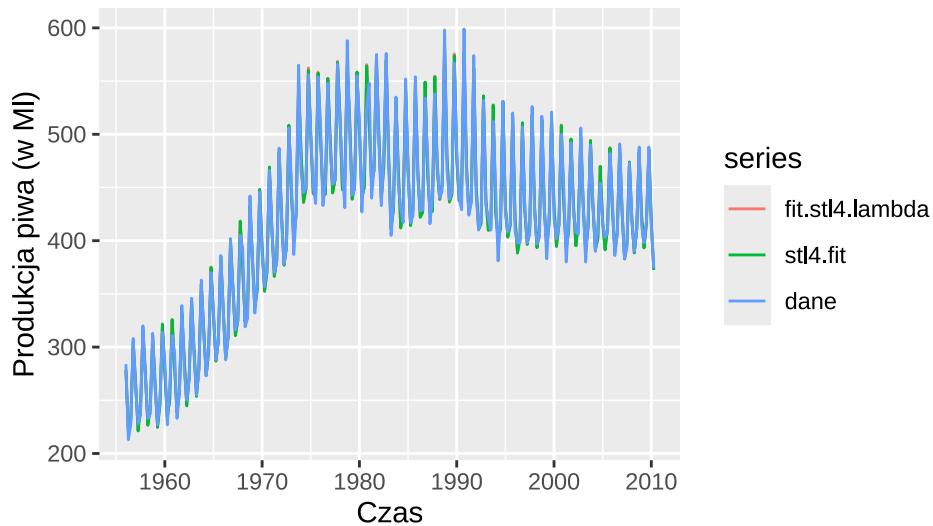
Rysunek 25: Dopasowanie modelu STL z par. `s.window = 'periodic'`, `t.window = 9`: brak Boxa-Coxa vs. Box-Cox

Transformacja Boxa-Coxa w tym przypadku poprawiła jakość dopasowania na początku i na końcu zbioru danych i nieznacznie pogorszyła ją w środku zbioru danych. Ogółem, można stwierdzić, że przekształcenie Boxa-Coxa wpłynęło pozytywnie na jakość dopasowania danych.

1.3.5 Dekompozycja STL z parametrami `s.window = 13`, `t.window = 5`

Ostatnim modelem, na którym sprawdzimy zastosowanie transformacji Boxa-Coxa będzie model dekompozycji STL z par. `s.window = 13`, `t.window = 5`. Na wykresie 26 zaprezentowano jakość dopasowania modelu do danych z i bez transformacji Boxa-Coxa.

Porównanie dopasowania modelu STL z par. s.window=13, t.window=5 bez i z transformacją Boxa-Coxa

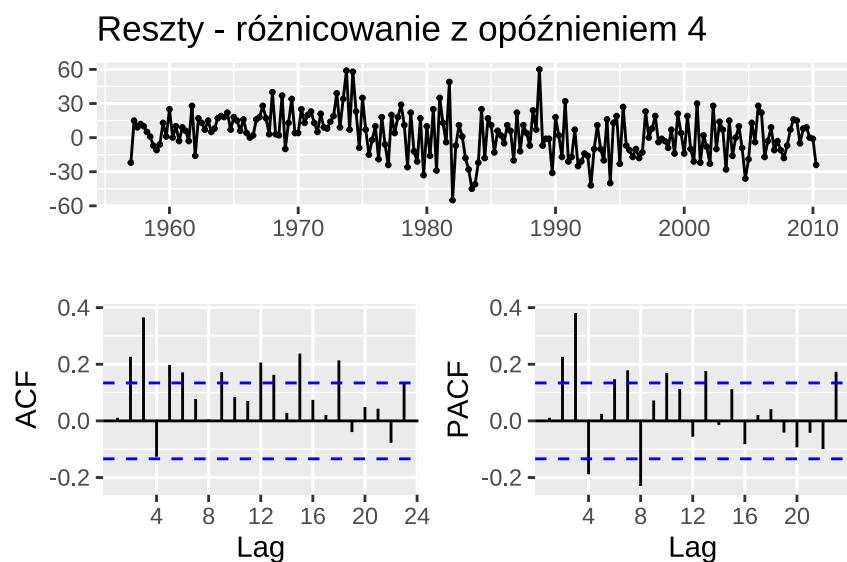


Rysunek 26: Dopasowanie modelu STL z par. `s.window = 13, t.window = 5`: brak Boxa-Coxa vs. Box-Cox

Zastosowanie transformacji Boxa-Coxa w tym przypadku nie spowodowało praktycznie żadnych różnic w dopasowaniu modelu do danych. Można więc uznać, że transformacja Boxa-Coxa dla tego modelu była niepotrzebna.

1.4 Porównanie z różnicowaniem danych

Ostatnim krokiem będzie porównanie wyników eliminacji trendu i sezonowości na podstawie rozważanych metod dekompozycji z wynikami uzyskanymi poprzez odpowiednie różnicowanie danych. W tym przypadku będziemy porównywać szeregi reszt. Do różnicowania danych zastosujemy funkcję `diff` z parametrem `lag=4`, gdyż dane mają okresowość 4. Wykres 27 przedstawia reszty po różnicowaniu danych.

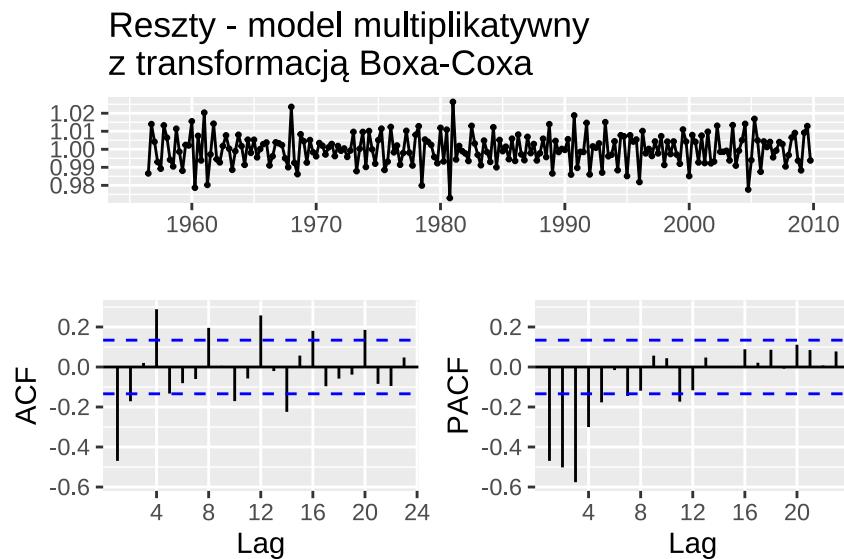


Rysunek 27: Wykres losowych reszt oraz ich ACF i PACF dla różnicowania danych z opóźnieniem 4

Z wykresu 27 możemy odczytać, że:

- nie ma pozostałości po sezonowości,
- są delikatne pozostałości po trendzie,
- dużo wartości w ACF odstaje poza wymagany przedział ufności, więc nie mamy do czynienia z białym szumem,
- nie ma wartości odstających,
- wariancja nie jest stabilna,
- do rozstrzygnięcia stacjonarności potrzebne są formalne testy.

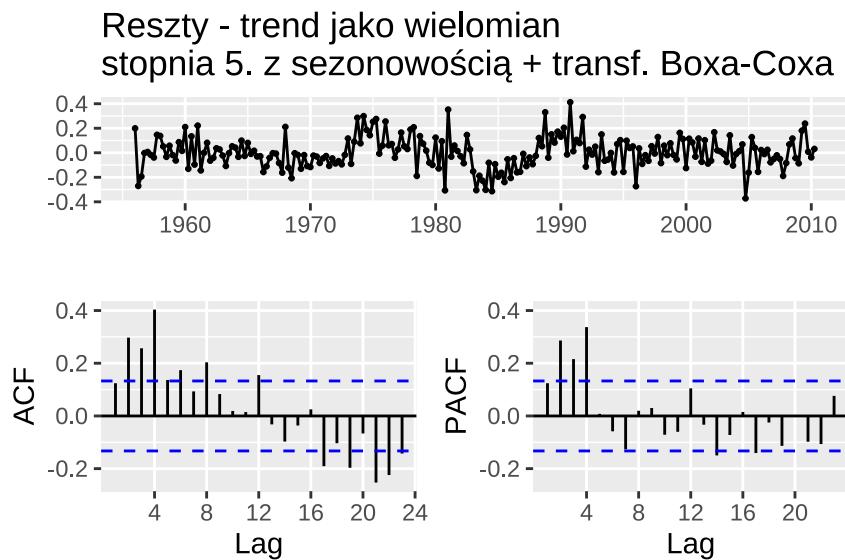
Widzimy, że dzięki różnicowaniu udało nam się otrzymać szereg reszt, który może mieć charakter stacjonarny. Porównamy to teraz z 3 modelami, na których zastosowana została transformacja Boxa-Coxa. Zaczniemy od modelu multiplikatywnego.



Rysunek 28: Wykres losowych reszt oraz ich ACF i PACF dla modelu multiplikatywnego z przekształceniem Boxa-Coxa

Reszty na wykresie 28 mają bardzo podobny charakter do reszt w modelu multiplikatywnym bez zastosowania transformacji Boxa-Coxa (6). Wnioski będą zatem podobne do tych z 6 (niewidoczna sezonowość, eliminacja trendu, możliwa stacjonarność). W porównaniu do szeregu 27 możemy dostrzec mniejsze pozostałości po trendzie oraz mniej "rozrzucone" obserwacje.

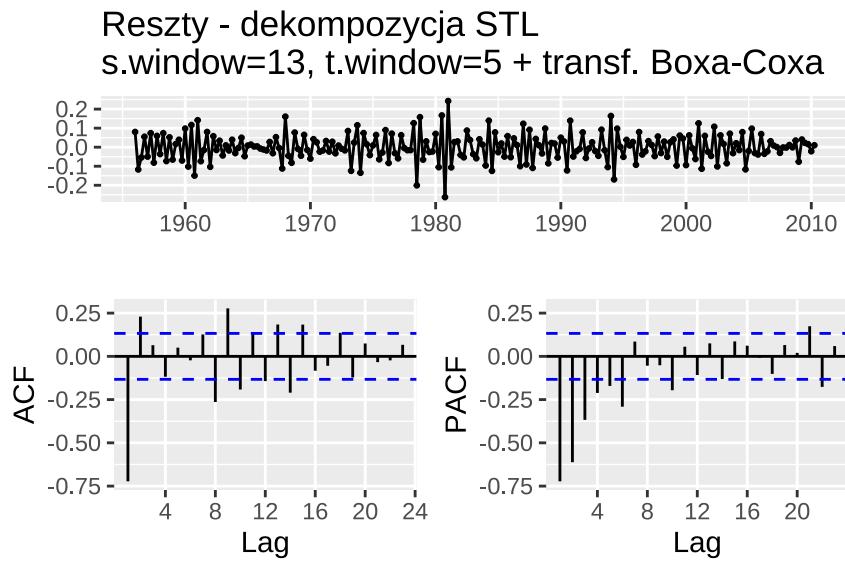
Kolejnym modelem, z jakim porównujemy różnicowanie będzie model regresji z trendem jako wielomian 5. stopnia oraz z zastosowaną transformacją Boxa-Coxa. Wykres reszt jest przedstawiony na rysunku 29.



Rysunek 29: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji na podstawie modelu regresji z trendem jako wielomian stopnia 5. i z sezonowością po zastosowaniu transformacji Boxa-Coxa

Na wykresie reszt 29 widzimy pozostałości po trendzie i wahańia szeregu, przez które nie możemy mówić o stacjonarności szeregu. Udało się jednak prawdopodobnie wyeliminować sezonowość. Stąd szereg danych po zróżnicowaniu ma lepsze własności teoretyczne.

Ostatnim szeregiem, z jakim porównamy szereg danych po zróżnicowaniu będzie szereg reszt dla modelu STL z parametrami `s.window = 13, t.window = 5` z zastosowaną transformacją Boxa-Coxa. W tym przypadku szereg reszt prezentuje się następująco:



Rysunek 30: Wykres losowych reszt oraz ich ACF i PACF dla dekompozycji STL z par. `s.window = 13, t.window = 5` po zastosowaniu transformacji Boxa-Coxa

Wnioski dla tego szeregu będą podobne, jak w przypadku 20, z tą różnicą, że wykres reszt 30 wydaje się szerszy, przez co jeszcze częściej zobaczyć wartości odstające. Ogółem też jest to raczej szereg stacjonarny, być może ma nawet lepsze własności niż szereg przedstawiony w 27.

W podsumowaniu:

- do rozstrzygnięcia stacjonarności szeregu danych po zróżnicowaniu potrzebny jest formalny test,
- szereg reszt 28 jest szybciej szeregiem stacjonarnym niż 27,
- szereg 29 ma gorsze własności teoretyczne niż szereg 27,
- szereg 30 możemy uznać za stacjonarny - ma zatem lepsze własności niż szereg danych po zróżnicowaniu.

2 Zastosowanie wskaźników analizy technicznej w analizie szeregów czasowych

W tej sekcji przyjrzymy się wybranym wskaźnikom analizy technicznej, które mogą być użyteczne w analizie szeregów czasowych.

2.1 Wybrane wskaźniki analizy technicznej

- **SMA (Simple Moving Average) - Prosta Średnia Krocząca, SMA(n):**
 - $\text{SMA}(n) = \frac{1}{n} \sum_{i=1}^n P_i$
 - (P_i) : cena zamknięcia z (i) -tego okresu
 - (n) : liczba okresów, dla których obliczamy średnią
 - bazowo: $(n = 50), (n = 200)$
 - sygnał kupna: SMA krótkookresowa przecina SMA długookresową od dołu
 - sygnał sprzedaży: SMA krótkookresowa przecina SMA długookresową od góry
 - działa najlepiej w trendach, mniej skuteczna w konsolidacji
- **MACD (Moving Average Convergence Divergence) - Zbieżność/Rozbieżność Średnich Krocących, MACD(n,m,k):**
 - $\text{MACD} = \text{EMA}_n - \text{EMA}_m$
 - linia MACD: różnica między krótkookresową EMA (n) a długookresową EMA (m) cen zamknięcia
 - linia sygnału: (k) -okresowa EMA linii MACD
 - bazowo: $(n = 12), (m = 26), (k = 9)$
 - sygnał kupna: MACD przecina linię sygnału od dołu
 - sygnał sprzedaży: MACD przecina linię sygnału od góry
 - najlepiej działa w trendach, ale może dawać opóźnione sygnały w konsolidacji
- **STS (Stochastic Oscillator) - Oszylator Stochastyczny, STS(n):**
 - $\text{STS} = 100 \times \frac{C - L_n}{H_n - L_n}$
 - (C) : cena zamknięcia, (L_n) : najniższa cena, (H_n) : najwyższa cena z ostatnich (n) okresów
 - bazowo: $(n = 14)$, poziom wykupienia = 80, poziom wyprzedania = 20
 - wykupienie: $\text{STS} > 80$, wyprzedaż: $\text{STS} < 20$
 - sygnał kupna: STS przecina 20 od dołu, sygnał sprzedaży: STS przecina 80 od góry
 - skuteczny w trendzie bocznym, mniej w silnych trendach
- **SMI (Stochastic Momentum Index) - Indeks Momentum Stochastycznego, SMI(n):**

- wygładzona wersja STS, uwzględniająca średnie kroczące
- bazowo: ($n = 14$), poziom wykupienia = 80, poziom wyprzedania = 20
- redukuje fałszywe sygnały w porównaniu do STS
- interpretacja podobna do STS, ale bardziej stabilna

- **Bollinger Bands - Wstęgi Bollingera, BB(n,k):**

Składają się z trzech linii:

- środkowa: (n)-okresowa SMA
- górna: $\text{SMA} + k \times \sigma$
- dolna: $(\text{SMA} - k \times \sigma)$
- (σ): odchylenie standardowe cen z (n) okresów
- bazowo: ($n = 20$), ($k = 2$)
- wykupienie: cena blisko górnej wstęgi, wyprzedanie: cena blisko dolnej wstęgi
- zwężezenie wstępów: niska zmienność, rozszerzenie wstępów: wysoka zmienność
- działa dobrze w ocenie zmienności i punktów odwrócenia

- **RSI (Relative Strength Index) - Indeks Siły Względnej, RSI(n):**

- $\text{RSI} = 100 - \frac{100}{1 + \frac{\text{średni zysk}}{\text{średnia strata}}}$
- średni zysk i strata obliczane z (n) okresów
- bazowo: ($n = 14$), poziom wyprzedania = 30, poziom wykupienia = 70
- wykupienie: $\text{RSI} > 70$, wyprzedanie: $\text{RSI} < 30$
- sygnał kupna: RSI przecina 30 od dołu, sygnał sprzedaży: RSI przecina 70 od góry
- przydatny do identyfikacji odwróceń trendu, szczególnie w połączeniu z innymi wskaźnikami

2.2 Wizualizacja i interpretacja wybranych wskaźników

Poniższa analiza techniczna dotyczy spółki Apple Inc. (AAPL) w okresie od 1 stycznia 2020 do 31 grudnia 2022 i wskaźników: SMA, MACD, Wstęp Bollingera i RSI.

SMA(50) i SMA(200)

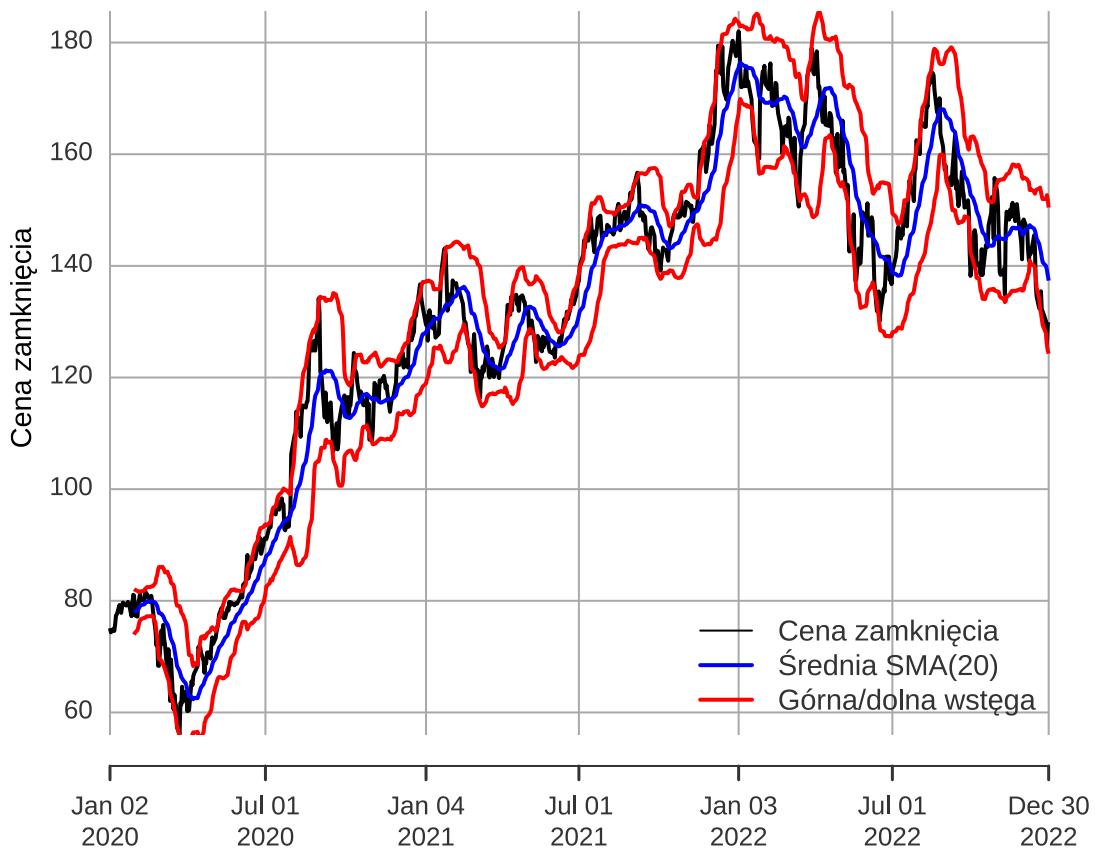


Rysunek 31: SMA dla cen zamknięcia AAPL w okresie 1.01.2020-31.12.2022

Interpretacja SMA (Rysunek 31):

- koniec 2022: SMA(50) przecinająca SMA(200) od góry - sygnał sprzedaży
- w latach 2020-2021 SMA(50) nad SMA(200) potwierdzała silny trend wzrostowy
- W 2022 SMA(50) pod SMA(200) wskazywała na dominację niedźwiedzi

Wstęgi Bollingera (20,2)



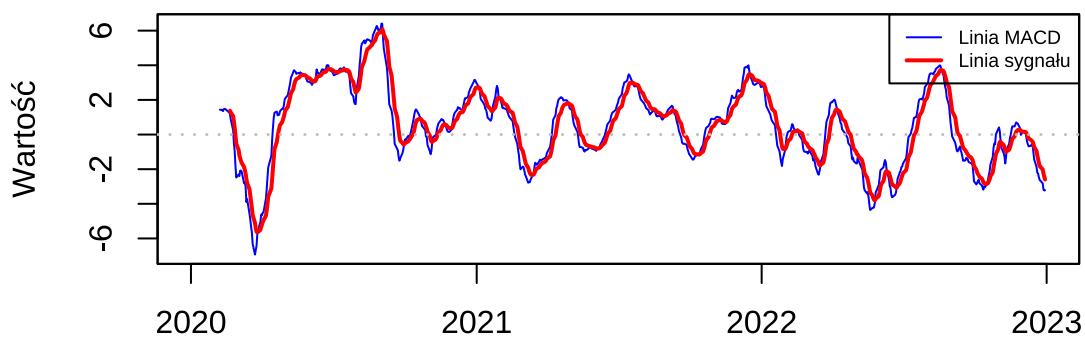
Rysunek 32: Wstęgi Bollingera dla cen zamknięcia AAPL w okresie 1.01.2020-31.12.2022

Interpretacja Wstęg Bollingera (Rysunek 32):

- marzec 2020: odbicia od dolnej wstęgi - sygnał kupna
- dotknięcie górnej wstęgi w 2020-2021 często poprzedzały korekty
- zwiększona szerokość wstęp w marcu 2020 i w 2022 wskazywała na podwyższoną zmienność



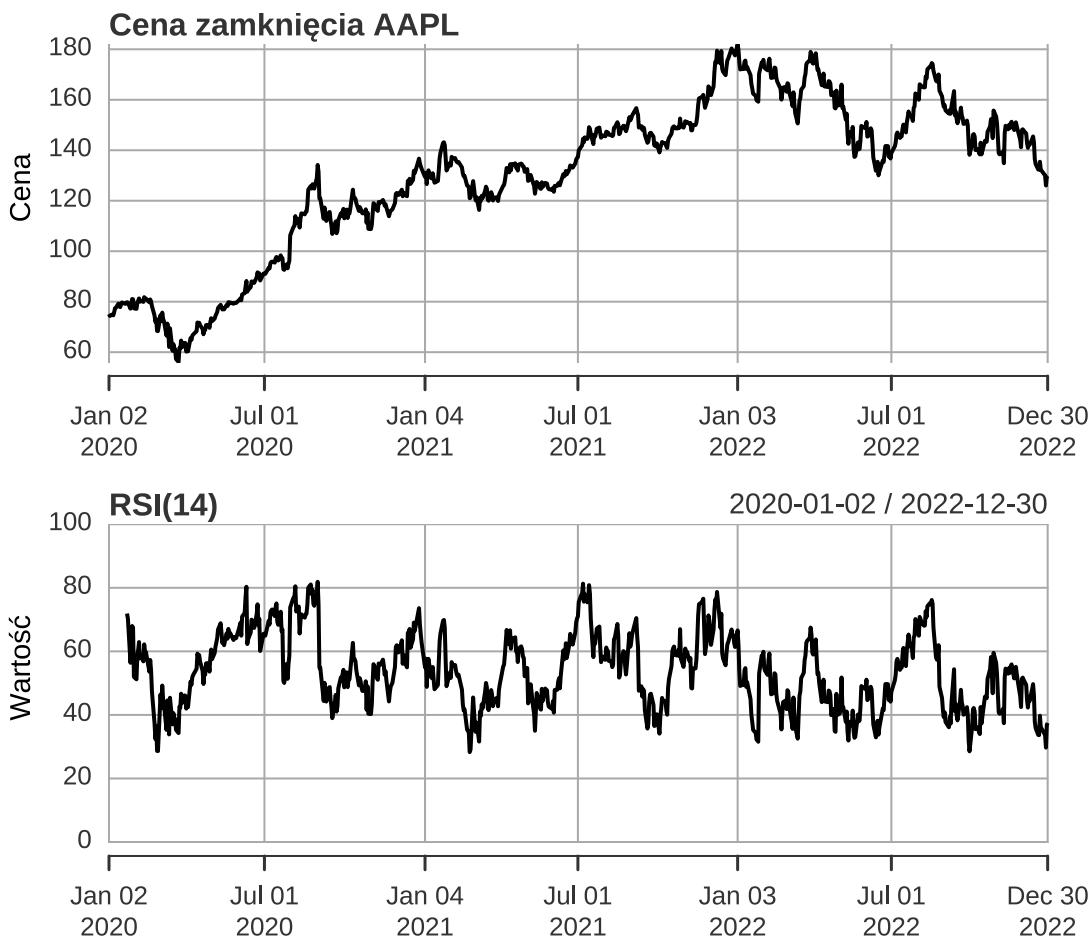
MACD(12,26,9)



Rysunek 33: MACD dla cen zamknięcia AAPL w okresie 1.01.2020-31.12.2022

Interpretacja MACD (Rysunek 33):

- marzec 2020: linia MACD poniżej linii sygnału - sygnał sprzedaży
- kwiecień 2020: przecięcie linii sygnału od dołu - sygnał kupna
- 2021: wielokrotne przecięcia linii sygnału - seria krótkoterminowych sygnałów
- 2022: MACD częściej poniżej zera - osłabienie trendu wzrostowego



Rysunek 34: RSI dla cen zamknięcia AAPL w okresie 1.01.2020-31.12.2022

Interpretacja RSI(14):

- Marzec 2020: spadek do 30 - wyprzedanie
- Połowa 2020: wzrost >70 - silne wykupienie
- 2020-2022: kilka okresów przekroczenia 70 - lokalne szczyty ceny
- 2022: RSI rzadko powyżej 60 - osłabienie trendu wzrostowego

Wpływ zmiany parametrów we wskaźnikach:

- **SMA:**
 - krótsze okresy (SMA(20) zamiast SMA(50)) generują szybsze sygnały, ale więcej fałszywych alarmów
 - przecięcia przy SMA(20/100) pojawiają się wcześniej niż przy SMA(50/200)
 - w 2022 SMA(20/100) wskazałyby wcześniejsze sygnały spadkowe niż SMA(50/200)
- **MACD:**
 - Parametry (8,17,9) zwiększą czułość wskaźnika – więcej sygnałów w trendach bocznych
 - Parametry (5,35,5) zwiększą opóźnienie, ale redukują fałszywe sygnały z 2022 roku
- **Wstęgi Bollingera:**
 - węższe wstęgi (1.5σ zamiast 2σ) dają więcej sygnałów, zwłaszcza w spokojnych okresach 2021
 - krótszy okres (10 zamiast 20) zwiększa czułość na gwałtowne ruchy z marca 2020

- szersze wstępki (2.5σ zamiast 2σ) lepiej identyfikują wyprzedanie z marca 2020
- **RSI:**
 - RSI(7) szybciej reaguje na zmiany – więcej sygnałów wyprzedania w 2022
 - RSI(21) daje mniej, ale bardziej wiarygodne sygnały ekstremów rynkowych