

Analiza białoszumowości i dekompozycja szeregów czasowych

Analiza szeregów czasowych

Ksawery Józefowski, 277513

2025-11-16

Spis treści

1	Analiza białego szumu w szeregach czasowych	2
1.1	Wstęp	2
1.2	Hipotezy	2
1.3	Test Graficzny	2
1.4	Testy Formalne Boxa-Pierce’a i Ljungu-Boxa	5
1.5	Symulacja	6
2	Dekompozycja Szeregów	9
2.1	Wstęp	9
2.2	Podstawowe własności szeregu	10
2.3	Dekompozycje	16
2.4	Porównanie wyników eliminacji trendu i sezonowości	26
2.5	Wnioski końcowe	29

1 Analiza białego szumu w szeregach czasowych

1.1 Wstęp

Badamy metody weryfikacji hipotezy, że obserwowany szereg czasowy jest białym szumem $\{X_t\} \sim WN(0, \sigma^2)$. Biały szum definiujemy jako ciąg niezależnych zmiennych losowych o identycznym rozkładzie z $E[X_t] = 0$, $Var(X_t) = \sigma^2$ i $Cov[X_t, X_{t+h}] = 0$, gdzie funkcja kowariancji nie zależy od t :

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2, & \text{gdy } h = 0, \\ 0, & \text{gdy } h \neq 0, \end{cases}$$

Analiza obejmuje zarówno test graficzny oparty na własnościach funkcji autokorelacji białego szumu i asymptotycznym rozkładzie autokorelacji próbkowej, jak i formalne testy statystyczne oparte na statystykach Boxa-Pierce’a oraz Ljunga-Boxa. W pierwszej części raportu przedstawiamy implementację obu podejść oraz ich ilustrację na podstawie realizacji szeregów czasowych. W drugiej części przeprowadzamy symulacje pozwalające ocenić skuteczność i zachowanie tych metod w różnych warunkach, uwzględniając liczbę powtórzeń, różne rozkłady, odmienne długości próby, wybór maksymalnego opóźnienia oraz przypadki procesów odbiegających od białego szumu.

1.2 Hipotezy

- **Hipoteza zerowa** H_0 : Szereg czasowy jest białym szumem

$$H_0 : \text{Wszystkie autokorelacje } \rho_x(h) = 0 \text{ dla } h > 0$$

- **Hipoteza alternatywna** H_1 : Szereg nie jest białym szumem

$$H_1 : \text{Istnieją istotne autokorelacje dla } h > 0$$

Do naszej hipotezy przyjmujemy poziom istotności $\alpha = 0.05$.

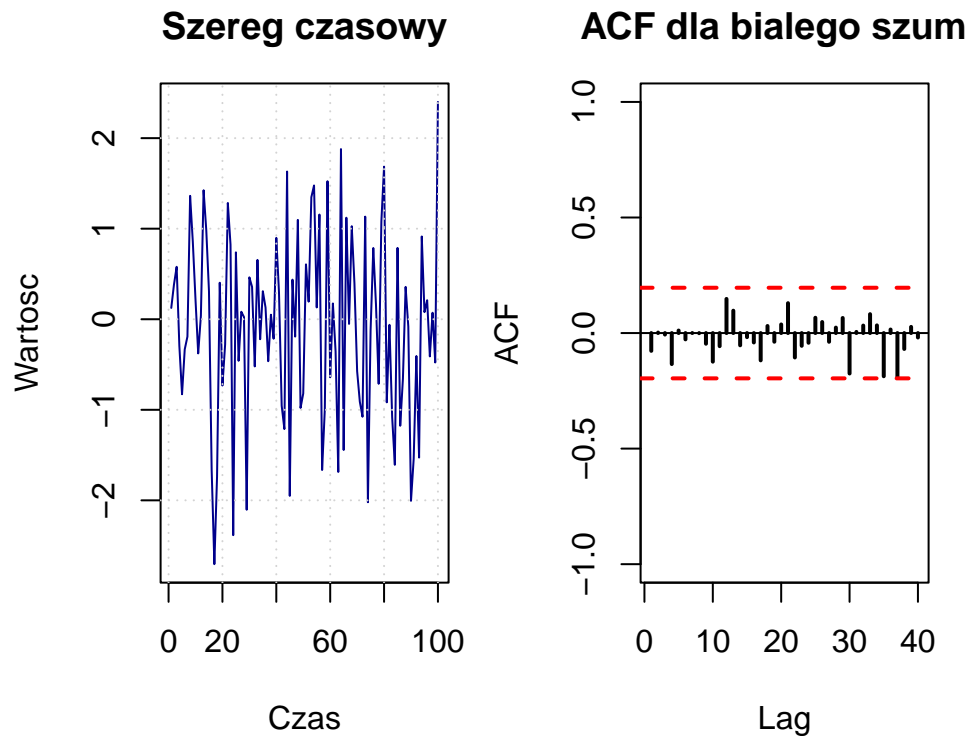
1.3 Test Graficzny

Do weryfikacji hipotezy H_0 można wykorzystać funkcję autokorelacji ACF. Autokorelacja dla opóźnienia h definiowana jest jako:

$$\rho_x(h) = \frac{\text{Cov}(X_t, X_{t+h})}{\text{Var}(X_t)}$$

Dla białego szumu oczekujemy $\rho_x(h) \approx 0$. Graficzny test polega na narysowaniu histogramu/wykresu słupkowego ACF i porównaniu wartości z granicami istotności.

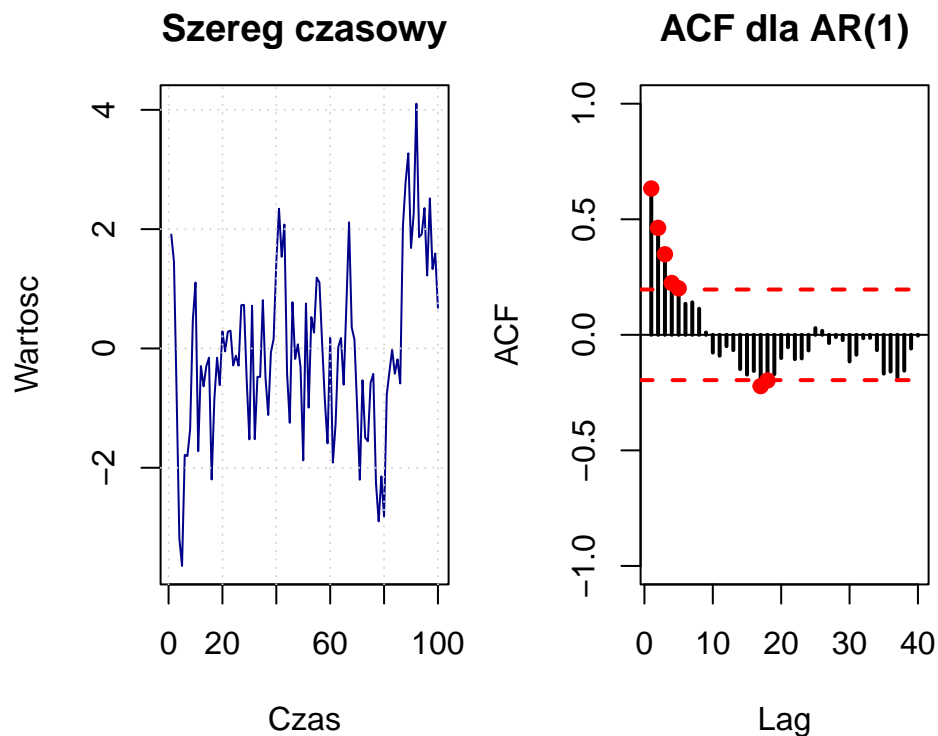
1.3.1 Przykład 1: Szereg białego szumu



Rysunek 1: Test graficzny dla białego szumu

Na Rysunek 1 przedstawiliśmy jak wygląda badany szereg białego szumu oraz jak zachowuje się funkcja autokorelacji. Wykres szeregu pokazuje losowe wahania bez wyraźnej struktury, więc jest to faktycznie biały szum. ACF mieści się w granicach istotności więc nasz test nie odrzuca H_0 potwierdzając biały szumowość tego szeregu.

1.3.2 Przykład 2: Szereg AR(1)



Rysunek 2: Test graficzny dla AR(1)

Na Rysunek 2 wizualizujemy jak wygląda szereg $AR(1)$ oraz jego funkcję ACF. Na wykresie szeregu można zauważyć zależność pomiędzy wartościami. ACF wychodzi poza granice istotności dla wielu opóźnień. Stąd nasz test dla tego szeregu odrzuca H_0 , co wskazuje na obecność korelacji.

1.3.3 Podsumowanie dla przykładów

Podsumowując, zaimplementowany test graficzny poprawnie weryfikuje hipotezę o białym szumie. W przypadku szeregu będącego białym szumem nie obserwuje się istotnych autokorelacji, a hipoteza zerowa nie jest odrzucana. Natomiast dla szeregu z pamięcią czasową $AR(1)$ test wykazuje autokorelacje i skutecznie odrzuca hipotezę zerową.

1.4 Testy Formalne Boxa-Pierce’a i Ljung-Boxa

Formalne testy statystyczne pozwalają zweryfikować hipotezę zerową H_0 , że obserwowany szereg czasowy jest białym szumem. W tym celu stosuje się m.in. test Boxa-Pierce’a oraz test Ljunga-Boxa.

Ich zasada działania jest bardzo prosta:

1. Najpierw obliczamy autokorelacje szeregu dla kolejnych opóźnień $h = 1, 2, \dots, h_{\max}$.
2. Następnie sumujemy kwadraty tych autokorelacji, tworząc statystykę testową, która rośnie wraz ze wzrostem zależności w szeregu.
3. Porównujemy uzyskaną wartość ze standardowym rozkładem χ^2 . Jeśli statystyka jest zbyt duża, oznacza to, że szereg nie zachowuje się jak biały szum, a hipoteza zerowa jest odrzucana.

1.4.1 Przykład 1: Szereg białego szumu

Tabela 1: Formalne testy białego szumu

Test	Statistic	p_value	Decision
Box-Pierce	33.542	0.96430	Nie odrzucono H_0
Ljung-Box	49.023	0.51255	Nie odrzucono H_0

Wyniki dla białego szumu przedstawiamy w tabeli 1. Zarówno test Boxa-Pierce’a, jak i test Ljunga-Boxa wykazały wysokie wartości p-value. Oznacza to, że nie odrzucamy hipotezy zerowej H_0 , a szereg można uznać za biały szum.

1.4.2 Przykład 2: Szereg AR(1)

Tabela 2: Formalne testy dla szeregu AR(1)

Test	Statistic	p_value	Decision
Box-Pierce	132.800	0	Odrzucono H_0
Ljung-Box	155.448	0	Odrzucono H_0

Statystyki dla szeregu AR(1) przedstawiamy w tabeli 2. W obu testach p-value wynosi 0, co oznacza, że statystyki testowe są znacznie większe od wartości krytycznych. Wnioskujemy więc, że odrzucamy hipotezę zerową H_0 .

1.4.3 Podsumowanie dla przykładów

Podobnie jak test graficzny, testy formalne skutecznie rozróżniają białe szumy od szeregów z zależnością czasową. Dla szeregu będącego białym szumem hipoteza zerowa nie została odrzucona, co potwierdza brak istotnych autokorelacji. Natomiast dla szeregu $AR(1)$, który nie jest białym szumem, testy formalne odrzuciły hipotezę zerową.

1.5 Symulacja

Poniższa symulacja porównuje skuteczność testu graficznego oraz testów formalnych w wykrywaniu biało-szumowości w różnych typach szeregów czasowych.

1.5.1 Plan symulacji

- Liczba powtórzeń: $K = 100$ realizacji dla każdej konfiguracji
- Długość szeregów: $n = 20, 50, 100$
- Maksymalne opóźnienie: $h_{\max} = 5, 50$
- Rodzaje rozkładów:
 - Standardowy normalny $\mathcal{N}(0, 1)$
 - Jednostajny $U(-\sqrt{3}, \sqrt{3})$
 - t Studenta z 5 stopniami swobody
- Szeregi inne niż biały szum:
 - Błądzenie losowe
 - Szereg z trendem i sezonowością: $X_t = 0.05t + \sin(2\pi t/12)$
 - Proces $MA(1)$ z $\theta = 0.6$

Tabela 3: Odsetek odrzuceń H_0 dla różnych testów i różnych typów szeregów

Model	n	h_max	Grafik	BoxPierce	LjungBox
Normalny	20	5	0.13	0.03	0.06
Normalny	20	50	0.14	0.00	0.07
Normalny	50	5	0.30	0.01	0.03
Normalny	50	50	0.41	0.00	0.12
Normalny	100	5	0.62	0.05	0.06
Normalny	100	50	0.66	0.02	0.12
Jednostajny	20	5	0.19	0.04	0.12
Jednostajny	20	50	0.10	0.00	0.06
Jednostajny	50	5	0.38	0.02	0.04
Jednostajny	50	50	0.33	0.00	0.06
Jednostajny	100	5	0.67	0.03	0.04
Jednostajny	100	50	0.70	0.02	0.13
t5_Stud	20	5	0.12	0.02	0.06
t5_Stud	20	50	0.08	0.00	0.10
t5_Stud	50	5	0.41	0.01	0.02
t5_Stud	50	50	0.44	0.01	0.12
t5_Stud	100	5	0.61	0.00	0.00
t5_Stud	100	50	0.61	0.00	0.06
MA1	20	5	0.49	0.12	0.26
MA1	20	50	0.44	0.00	0.33
MA1	50	5	0.98	0.58	0.64
MA1	50	50	0.97	0.03	0.55
MA1	100	5	1.00	0.97	0.98
MA1	100	50	1.00	0.49	0.78
RW	20	5	0.91	0.77	0.84
RW	20	50	0.85	0.44	0.89
RW	50	5	1.00	1.00	1.00
RW	50	50	1.00	0.98	1.00
RW	100	5	1.00	1.00	1.00
RW	100	50	1.00	1.00	1.00
TrendSezon	20	5	0.28	0.05	0.12
TrendSezon	20	50	0.30	0.00	0.16
TrendSezon	50	5	0.93	0.81	0.83
TrendSezon	50	50	0.99	0.21	0.92
TrendSezon	100	5	1.00	1.00	1.00
TrendSezon	100	50	1.00	1.00	1.00

Analiza wyników z tabeli 3 pozwala wyciągnąć następujące wnioski. Dla szeregów będących białym szumem, obserwujemy, że test graficzny radzi sobie coraz gorzej wraz ze wzrostem długości szeregu. Im dłuższy szereg, tym wyższy jest odsetek odrzuceń hipotezy H_0 , co wskazuje na tendencję tego testu do błędnego klasyfikowania białego szumu. Parametr h_{max} nie wpływa przy tym w sposób znaczący na wyniki. W przeciwieństwie do tego, testy formalne radzą sobie z tym zadaniem znacznie lepiej. Test Boxa-Pierce’a wykazuje niemal idealne wyniki, mieszczące się w granicach błędu, przy czym im większa wartość h_{max} , tym lepiej rozpoznaje on biały szum, a długość szeregu nie ma istotnego wpływu na jego skuteczność. Test Ljung-Boxa radzi sobie nieco gorzej niż test Boxa-Pierce’a, co wynika z jego nieco wyższej mocy, która w tym przypadku prowadzi do nieco częstszego błędnego odrzucania H_0 . Dla testu Ljung-Boxa obserwuje się, że im większa wartość h_{max} , tym gorzej radzi on sobie z poprawną identyfikacją białego szumu, podczas gdy długość szeregu ponownie nie wpływa znacząco na rezultaty.

W przypadku szeregów, które nie są białym szumem, test graficzny wykazuje bardzo dobrą skuteczność. Im dłuższy jest badany szereg, tym lepiej radzi sobie z poprawnym odrzuceniem H_0 , przy czym h_{max} nie odgrywa tu znaczącej roli. Dla testu Boxa-Pierce’a również im dłuższy szereg, tym lepiej wykrywa on niebiałoszumowość. Jednakże w jego przypadku obserwuje się wyraźny negatywny wpływ parametru h_{max} . Im jest on większy, tym gorzej test ten radzi sobie z wykrywaniem zależności. Doskonale widać to na przykładzie szeregu z trendem i sezonowością, gdzie dla $n = 20$ i $h_{max} = 50$ odsetek odrzuceń jest równy 0, co oznacza, że test uznaje ten proces w pełni za biały szum. Test Ljung-Boxa również nie radzi sobie w tych warunkach idealnie, lecz znacznie lepiej niż Box-Pierce. Ogólnie im dłuższy szereg, tym lepiej odrzuca H_0 . W przeciwieństwie do testu Boxa-Pierce’a, im większy parametr h_{max} tym lepsze wyniki testu.

2 Dekompozycja Szeregów

2.1 Wstęp

Celem zadania jest analiza wybranego szeregu czasowego. W naszym przypadku będzie to **bezrobocie{TSAFBook}**, obejmująca identyfikację jego podstawowych własności oraz zastosowanie i porównanie wybranych metod dekompozycji. Dekompozycja szeregów czasowych, umożliwia rozbięcie obserwowanego szeregu $X_t = f(s_t, m_t, Z_t)$ na trzy główne komponenty: trend m_t , sezonowość s_t oraz składnik losowy Z_t .

W zależności od charakteru danych rozważamy modele addytywne:

$$X_t = s_t + m_t + Z_t$$

oraz modele multiplikatywne:

$$X_t = s_t \cdot m_t \cdot Z_t$$

W pierwszej części raportu przeprowadzimy wstępną analizę graficzną szeregu bezrobocia. Pozwoli to ocenić obecność trendu, struktury sezonowej, potencjalnych zmian wariancji oraz obserwacji odstających.

Następnie zastosujemy trzy metody dekompozycji:

- dekompozycję opartą na ruchomej średniej `decompose()`,
- dekompozycję regresyjną opartą na modelu `tslm()`.
- dekompozycję STL bazującą na lokalnym wygładzaniu `stl()`.

W analizie uwzględnimy również wpływ parametrów poszczególnych metod: wybór modelu addytywnego lub multiplikatywnego w `decompose()`, wykorzystanie różnych modeli trendu w `tslm()`, oraz wpływ parametrów wygładzających w metodzie STL. Sprawdzimy także, czy zastosowanie transformacji Boxa–Coxa może poprawić jednorodność wariancji i jakość otrzymanej dekompozycji.

Ostatnim etapem raportu będzie porównanie wyników dekompozycji z rezultatami uzyskanymi poprzez odpowiednie różnicowanie szeregu. Analiza szeregu reszt pozwoli ocenić, czy pozostałości po usunięciu trendu i sezonowości można uznać za realizację szeregu stacjonarnego.

2.2 Podstawowe własności szeregu

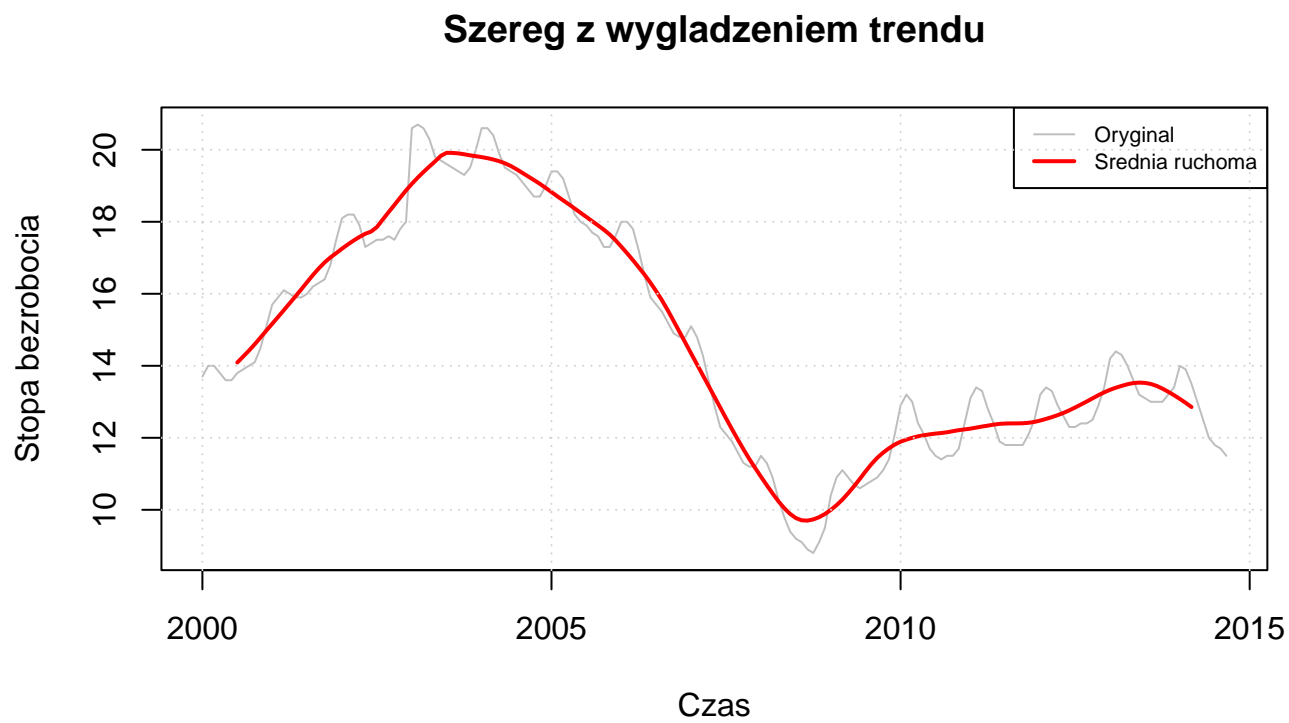
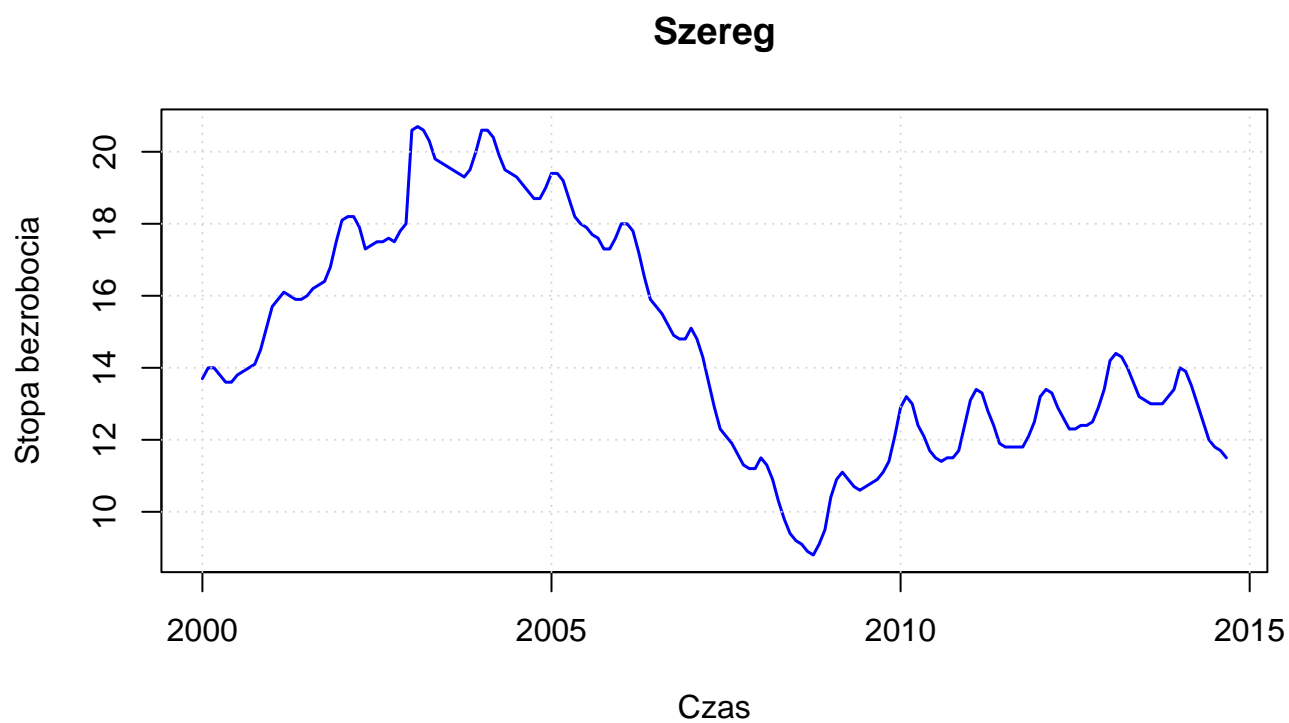
W tej części raportu analizujemy podstawowe własności szeregu czasowego `bezrobocie`{TSAFBook}. W szczególności badamy obecność trendu, sezonowości, jednorodność wariancji oraz ewentualne obserwacje odstające.

Tabela 4: Podstawowe informacje o szeregu `bezrobocie`

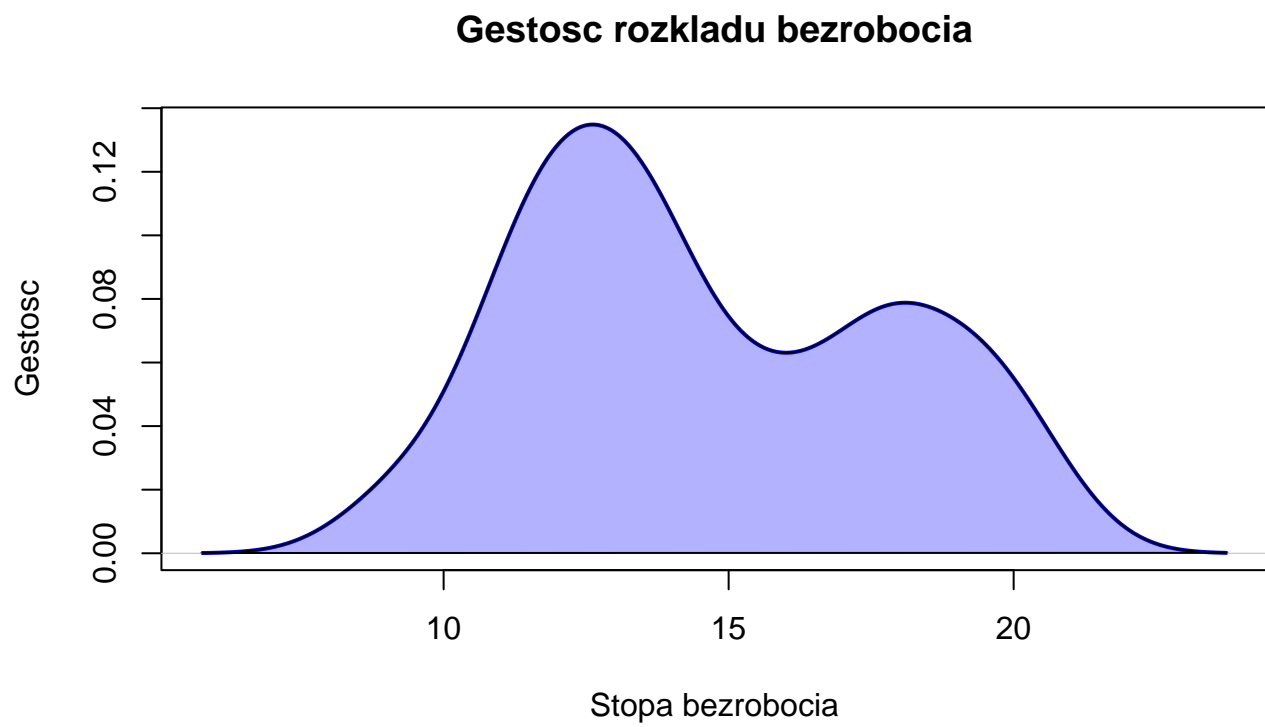
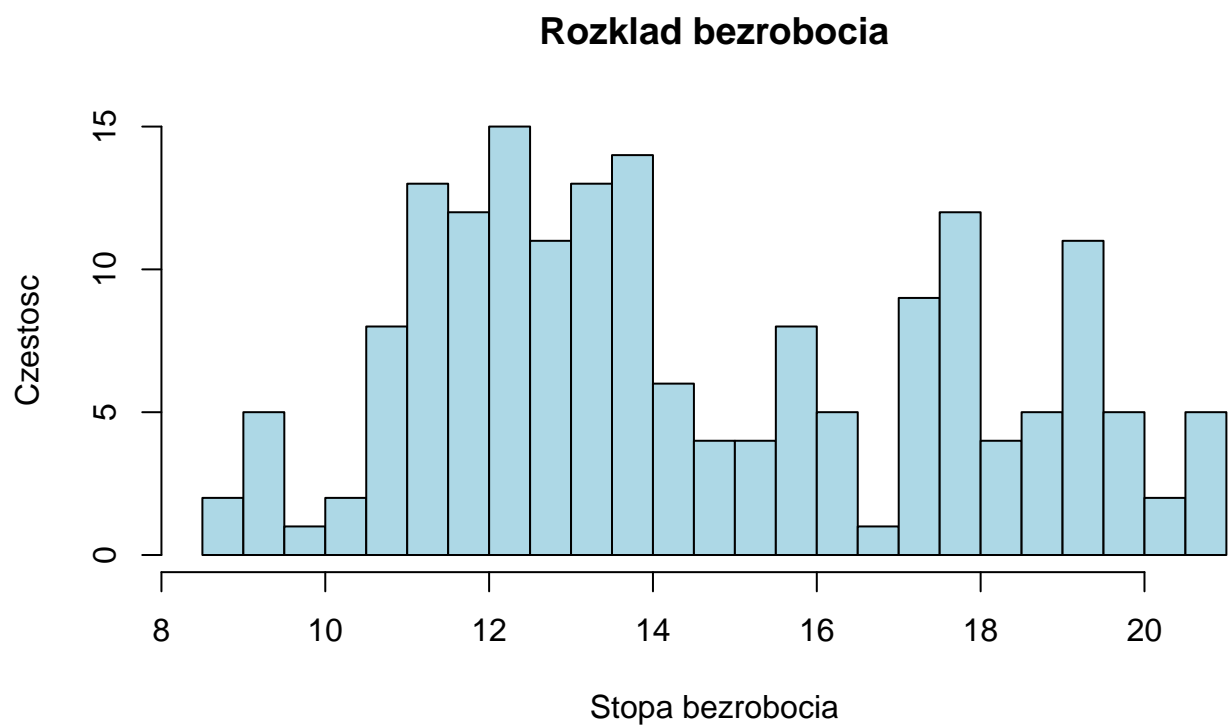
Parametr	Wartość
Początek szeregu	2000 1
Koniec szeregu	2014 9
Częstotliwość	12
Długość szeregu	177

Tabela 5: Podstawowe statystyki opisowe szeregu `bezrobocie`

Min	1st Q	Mediana	Średnia	3rd Q	Max
8.8	12.1	13.8	14.616	17.5	20.7

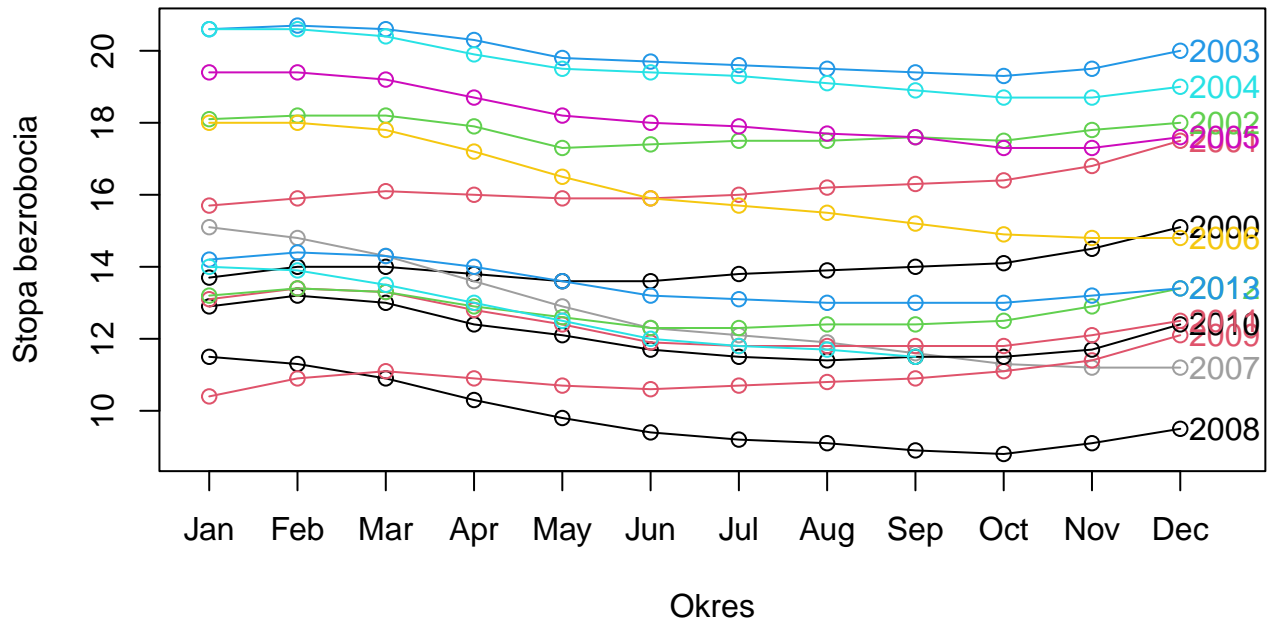


Rysunek 3: Wykresy Szeregu i wygładzenia trendu

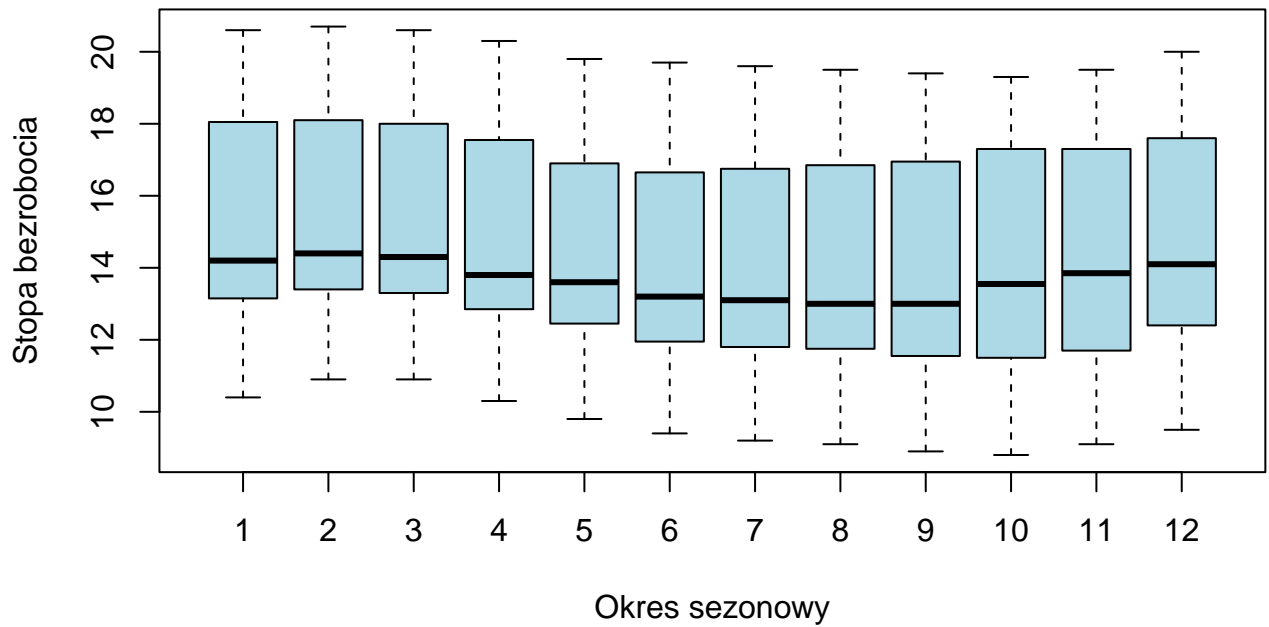


Rysunek 4: Wykresy rozkładu i gęstości szeregu bezrobocia

Wykres sezonowy

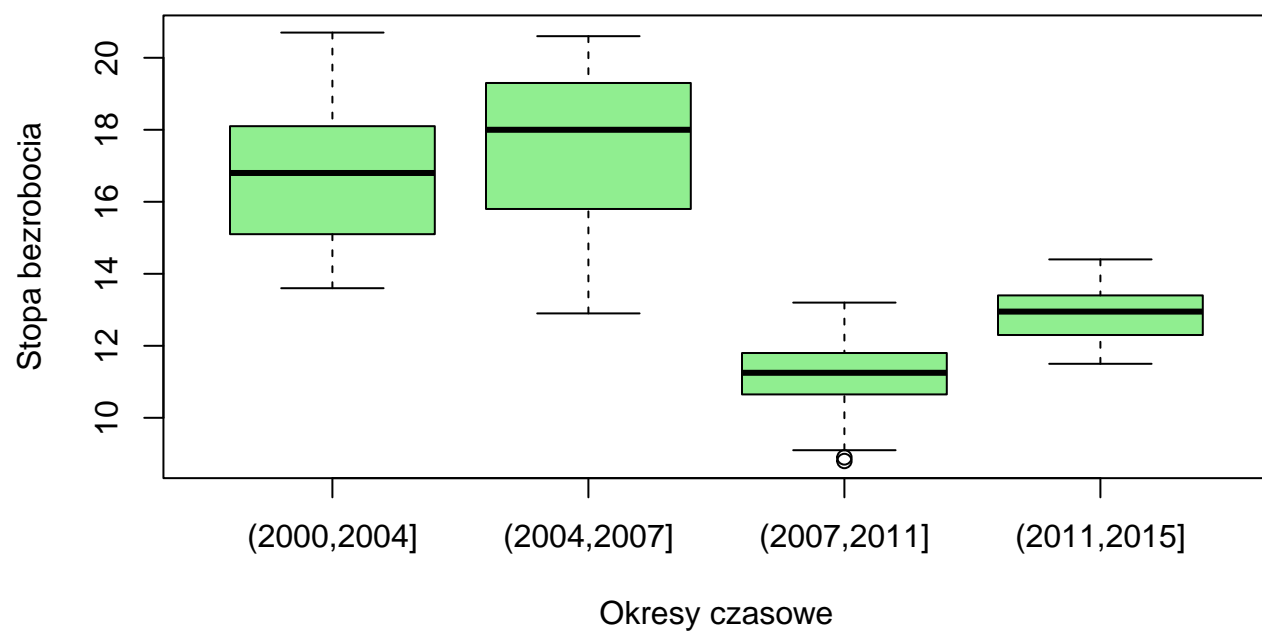


Rozkład sezonowy

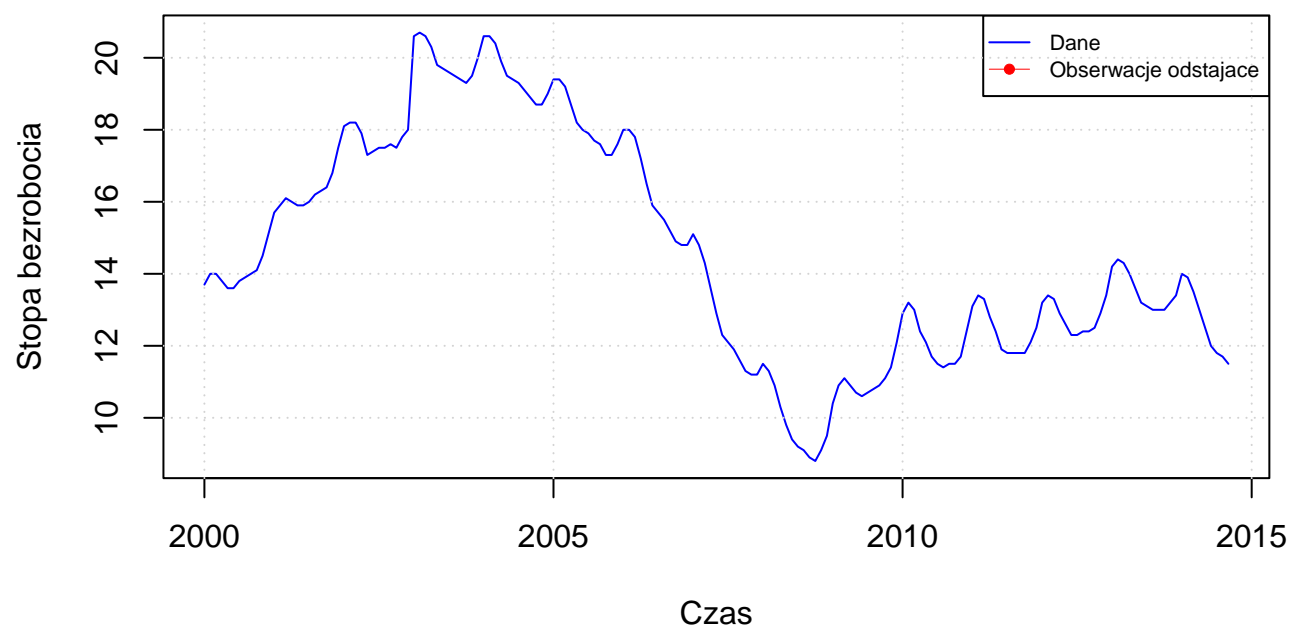


Rysunek 5: Wykresy sezonowości szeregu

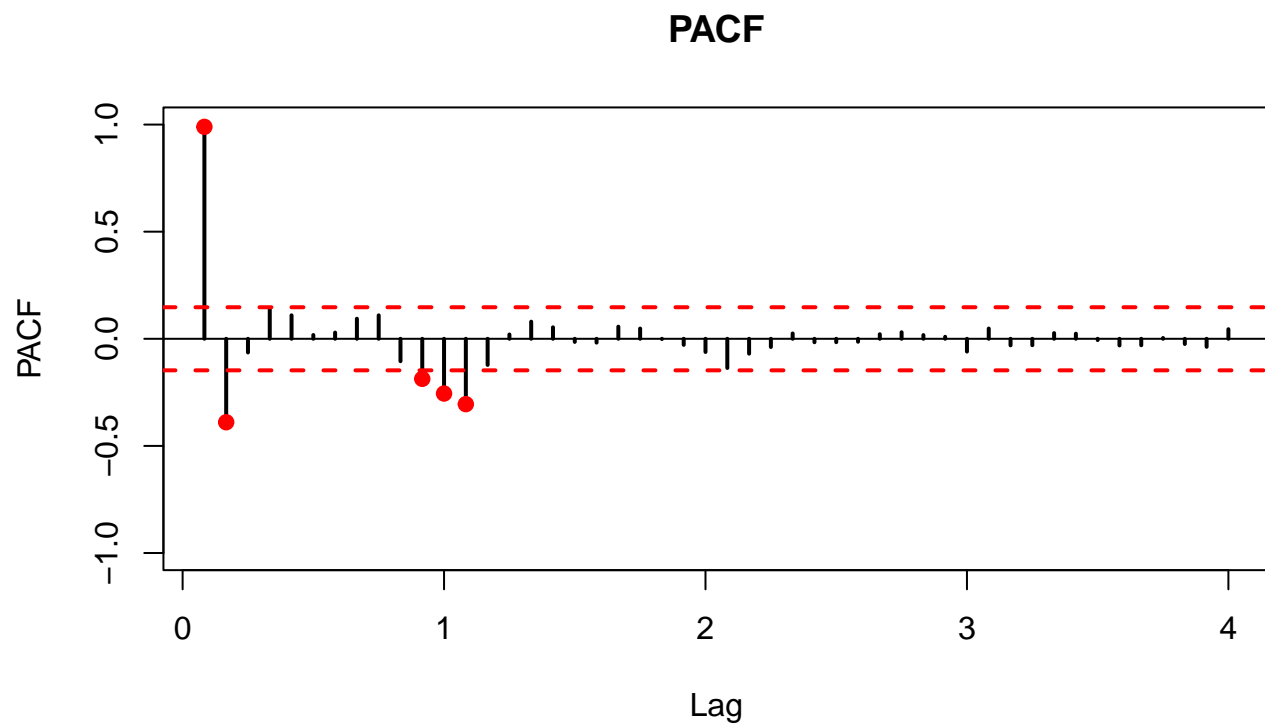
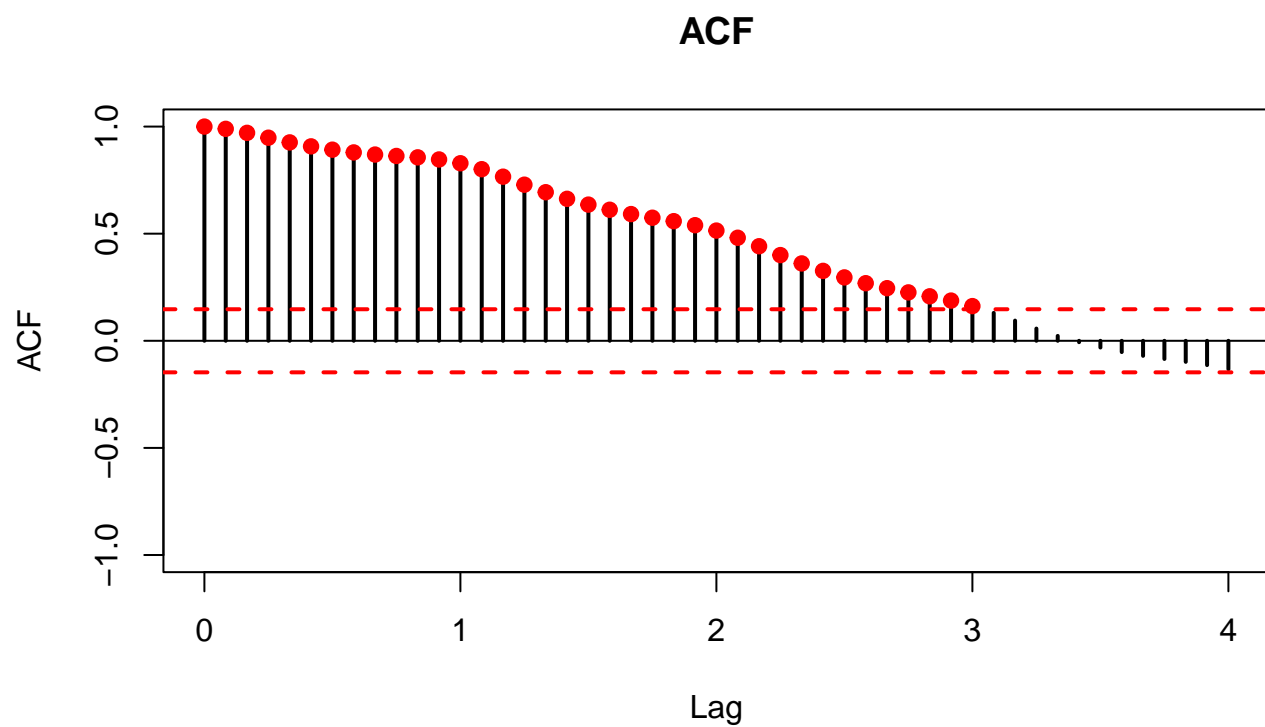
Wariancja w czasie



Obserwacje odstające szeregu



Rysunek 6: Wykresy wariancji i obserwacji odstających



Rysunek 7: Wykresy ACF i PACF

2.2.1 Analiza wyników podstawowych własności szeregu

W oparciu o przedstawione wykresy, statystyki opisowe oraz analizę sezonowości i autokorelacji szeregu `bezrobocie` {TSAFBook}, możemy sformułować następujące wnioski:

1. Trend i ogólne zachowanie szeregu: Na rysunku 3 z wykresu szeregu oraz wykresu z nałożoną średnią ruchomą wynika, że stopa bezrobocia wykazuje wyraźne trendy długookresowe. W pewnym okresie obserwujemy spadki, w innych wzrosty, przy czym wygładzenie sugeruje wolnozmiennie trendy. Nie jest to szereg stacjonarny.
2. Sezonowość: Wykres sezonowy oraz rozkład sezonowy z rysunku 5 wskazują na powtarzalną sezonowość. W poszczególnych okresach w roku wartości bezrobocia układają się podobnie, co oznacza, że:

$$X_t = s_t + m_t + Z_t$$

, gdzie s_t jest istotnym składnikiem.

3. Wariancja w czasie i Obserwacje odstające: Analizując rysunek 6 zauważamy, że wariancja w różnych segmentach czasowych nie wskazuje na duże zmiany rozproszenia wartości. Obserwacje odstające nie pojawiają się, co oznacza, że szereg nie zawiera poważnych anomalii.
4. Rozkład: Histogram i estymowana gęstość z rysunku 4 wskazują, że rozkład stopy bezrobocia różni się od rozkładu normalnego, co jest typowe dla danych ekonomicznych. Nie można więc przyjąć $X_t \sim \mathcal{N}(\mu, \sigma^2)$.
5. Autokorelacja i częściowa autokorelacja: Na rysunk 7 wykresy ACF i PACF pokazują liczne istotne autokorelacje dla małych i sezonowych opóźnień, jak i cykliczne wzorce zgodne z częstotliwością sezonową. Oznacza to, że szereg wykazuje zależność czasową, co potwierdza niestacjonarność i obecność sezonowości $\rho(k) \not\approx 0$ dla wielu k .

2.3 Dekompozycje

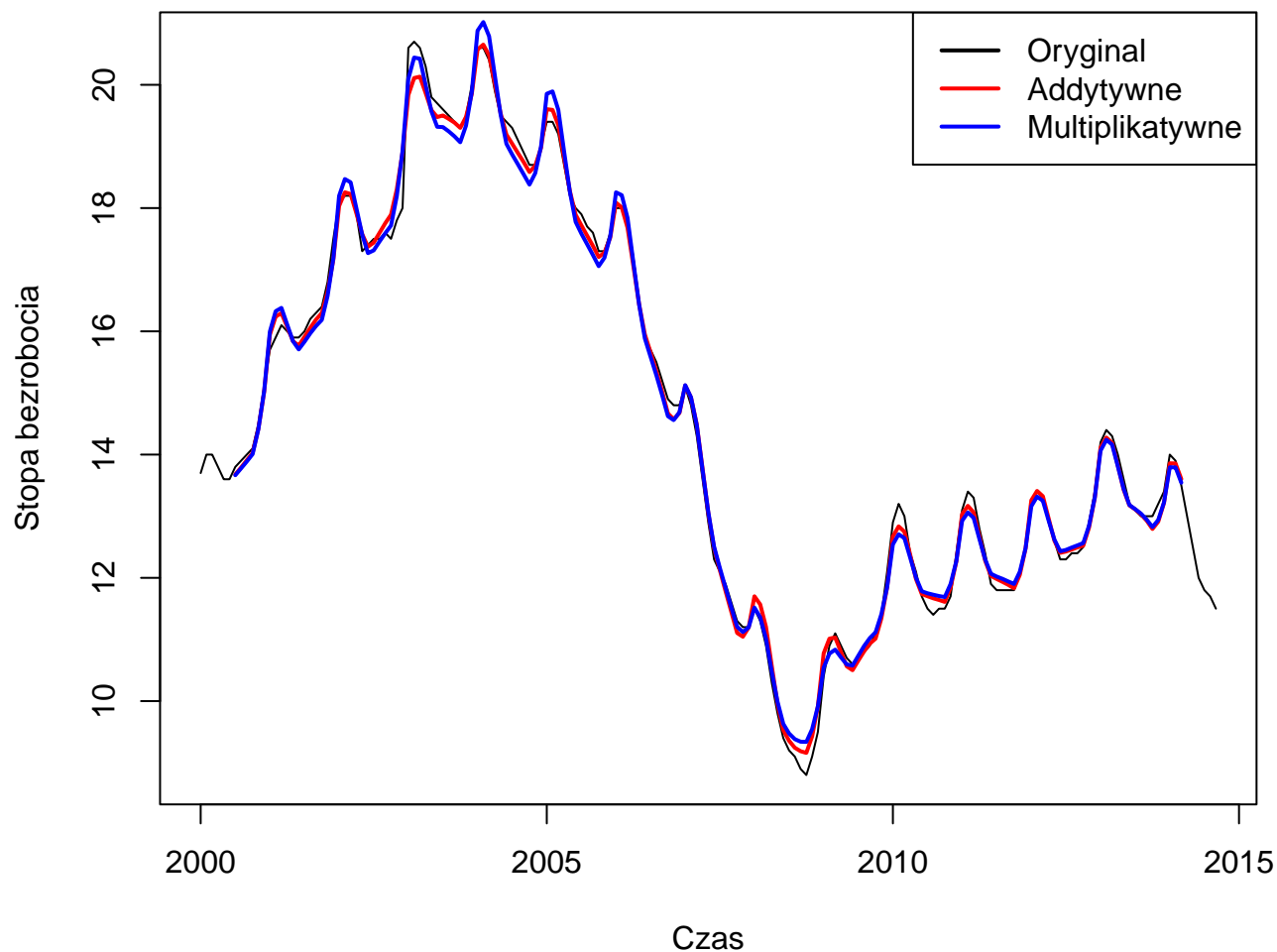
W tej części analizujemy trzy podejścia do dekompozycji szeregu `bezrobocie`:

- klasyczną dekompozycję opartą na średnich ruchomych (`decompose()`),
- dekompozycję modelową (`tslm()`),
- dekompozycję opartą na wygładzaniu lokalnym (`stl()`).

Celem jest ocena jakości dopasowania oraz wpływu parametrów poszczególnych metod na uzyskane wyniki.

2.3.1 Dekompozycja na podstawie ruchomej średniej

Zaczynamy najpierw od dekompozycji na podstawie ruchomej średniej. Rozpatrzmy model addytywny i model multiplikatywny, które dogłębniej opisałem we wstępie.



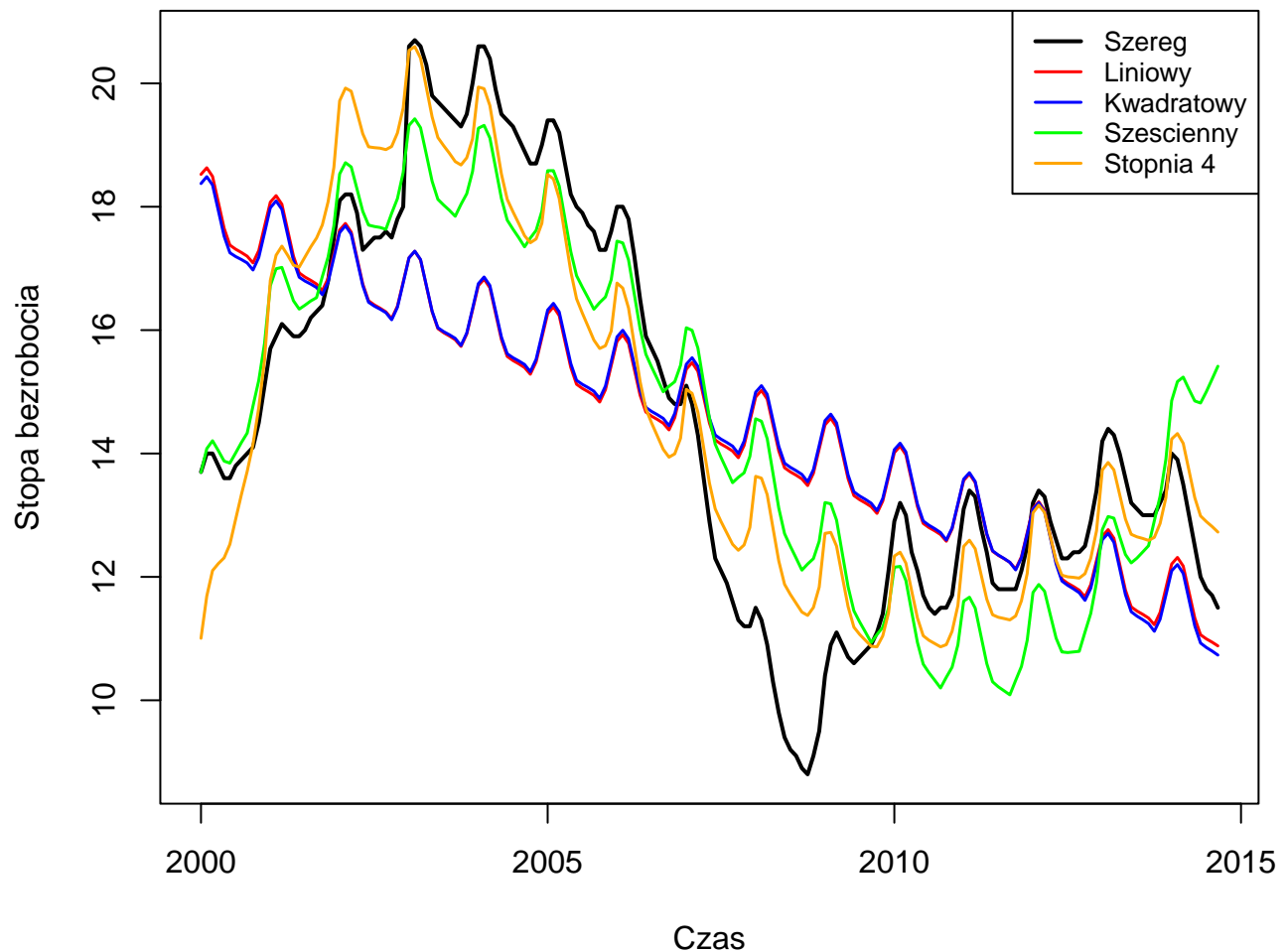
Rysunek 8: Dopasowanie dekompozycji na podstawie ruchomej średniej, addytywnie i multiplikatywnie

W przypadku szeregu **bezrobocie** amplituda sezonowości jest względnie stała w czasie, a poziom szeregu ulega długookresowym zmianom. Z tego powodu zauważamy na wykresie 8, że model addytywny daje bardziej stabilne dopasowanie i nie przeszacowuje wahań sezonowych na przełomie okresów o niskiej wartości szeregu. Model multiplikatywny dopasowuje trend poprawnie, ale zawyża amplitudę sezonowości w okresach wysokiego bezrobocia i zaniża ją w okresach niskiego, co wskazuje na gorsze dopasowanie.

2.3.2 Dekompozycja na podstawie modelu

Następną dekompozycją jaką rozpatrzymy jest dekompozycja na podstawie modelu. W tej sekcji zakładamy, że trend m_t można przybliżyć poprzez wielomian trendu, pewnego stopnia. Natomiast sezonowość s_t opiszemy zero-jedynkowo. Modele, które będziemy rozważać to:

- liniowy $m_t = a + bt$
- kwadratowy $m_t = a + bt + ct^2$
- sześcienny $m_t = a + bt + ct^2 + dt^3$
- wielomian czwartego stopnia $m_t = a + bt + ct^2 + dt^3 + et^4$



Rysunek 9: Dopasowanie dekompozycji na podstawie różnych modeli

Z rysunku 9 widzimy wyraźnie, że model liniowy jest najprostszą konstrukcją i nie jest w stanie uchwycić nieliniowych zmian poziomu bezrobocia. W efekcie trend jest znacznie niedopasowany, choć sama sezonowość jest estymowana poprawnie, ponieważ zmienne sezonowe w modelu działają niezależnie od postaci trendu. Podobna sytuacja występuje w modelu kwadratowym. Z jednej strony dodanie składnika t^2 pozwala na opisanie prostych krzywizn trendu, z drugiej jednak trend bezrobocia zmienia się w sposób bardziej złożony niż parabola. Model kwadratowy nie potrafi więc dobrze uchwycić zmienności długookresowej, mimo że składnik sezonowy pozostaje poprawnie estymowany.

Znaczącą poprawę dopasowania obserwujemy w modelu sześciennym, który dzięki dodatkowej elastyczności znacznie lepiej odwzorowuje zmienność trendu, a tym samym wyraźnie redukuje autokorelację reszt. Krzywizna trzeciego stopnia okazuje się wystarczająca, by uchwycić główne nieliniowości w przebiegu bezrobocia. Model czwartego stopnia dopasowuje trend nieco lepiej niż model sześcienny, jednak równocześnie zaczynają pojawiać się oznaki overfittingu.

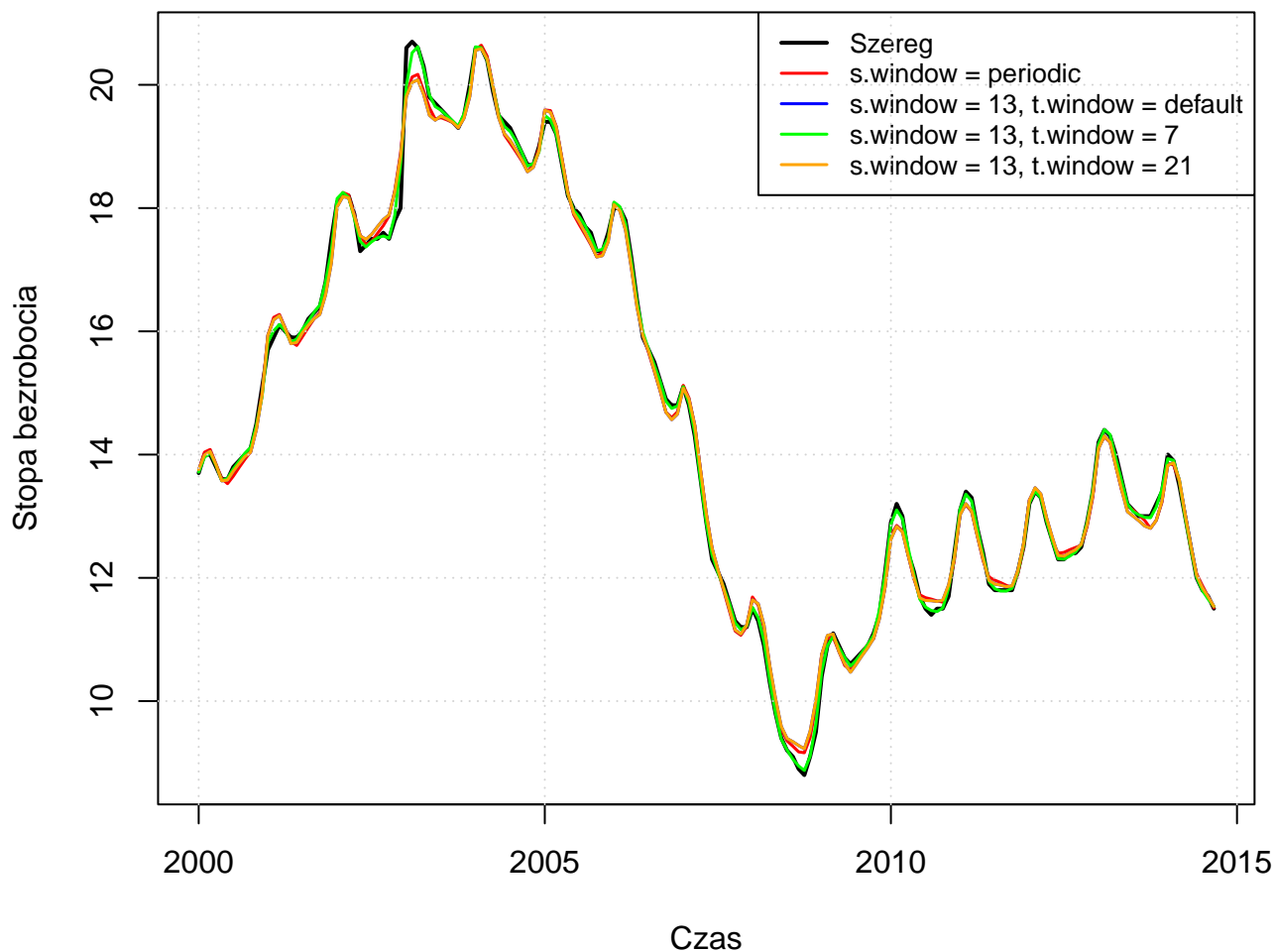
Pod względem jakości odwzorowania trendu i reszt, modele można uporządkować następująco:

$$\textit{Liniowy} < \textit{Kwadratowy} < \textit{Sześcienny} < \textit{Stopnia4}$$

2.3.3 Dekompozycja STL oparta na metodzie loess

W tej sekcji rozważymy dekompozycję STL opartą na metodzie loess, która umożliwia elastyczne i niezależne modelowanie komponentów trendu, sezonowości i reszt. STL pozwala regulować stopień wygładzenia zarówno trendu, jak i sezonowości, co daje większą kontrolę nad jakością dopasowania. Kluczową rolę odegrają tutaj parametry wygładzenia. Będziemy analizować parametry:

- $s.window = \textit{periodic}$ — sezonowość stała, bez wygładzania,
- $s.window = 13$
- $t.window = 7, 21$

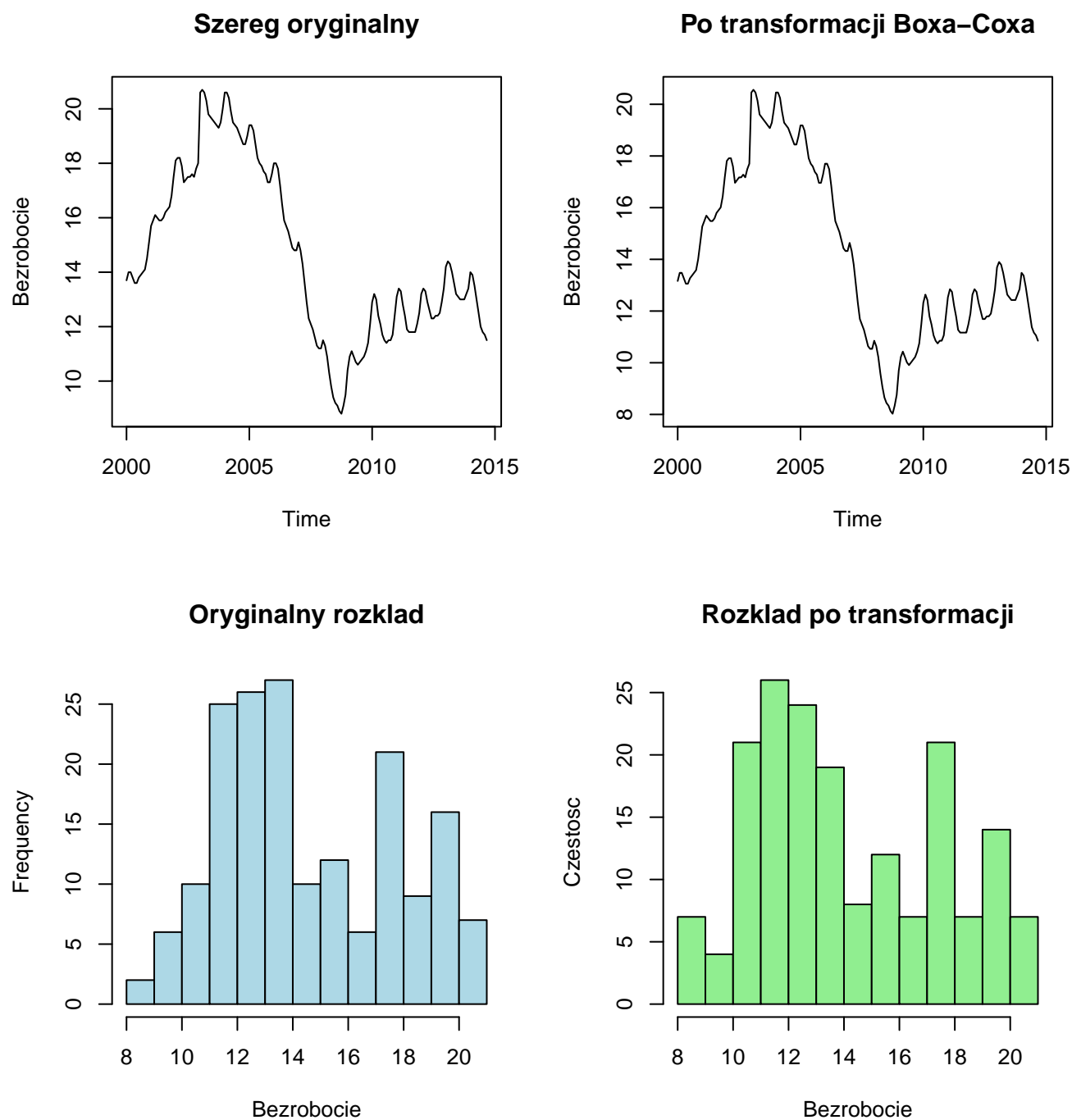


Rysunek 10: Porównanie dopasowania dekompozycji STL z różnymi parametrami

Z rysunku 10 odczytujemy, że Ustawienie $s.window = periodic$ wymusza identyczny przebieg sezonowości w każdym roku. Model dopasowuje się stabilnie, ale ignoruje subtelne zmiany sezonowe w czasie. Wariant z $s.window = 13$ pozwala na dostosowanie sezonowości lokalnie, dzięki czemu uchwycone zostają drobne zmiany struktury sezonowej. Parametr $t.window$ kontroluje wygładzanie trendu, $t.window = 7$ daje trend bardziej “poszarpany”, reagujący na krótkookresowe zmiany, zato $t.window = 21$ daje trend gładki, wygładzony, miejscami zbyt mocno. STL daje największą elastyczność spośród wszystkich metod, ponieważ umożliwia niezależne kontrolowanie wygładzania trendu i sezonowości. Najlepsze dopasowanie uzyskaliśmy dla ustawień $s.window = 13$, $t.window = 7$.

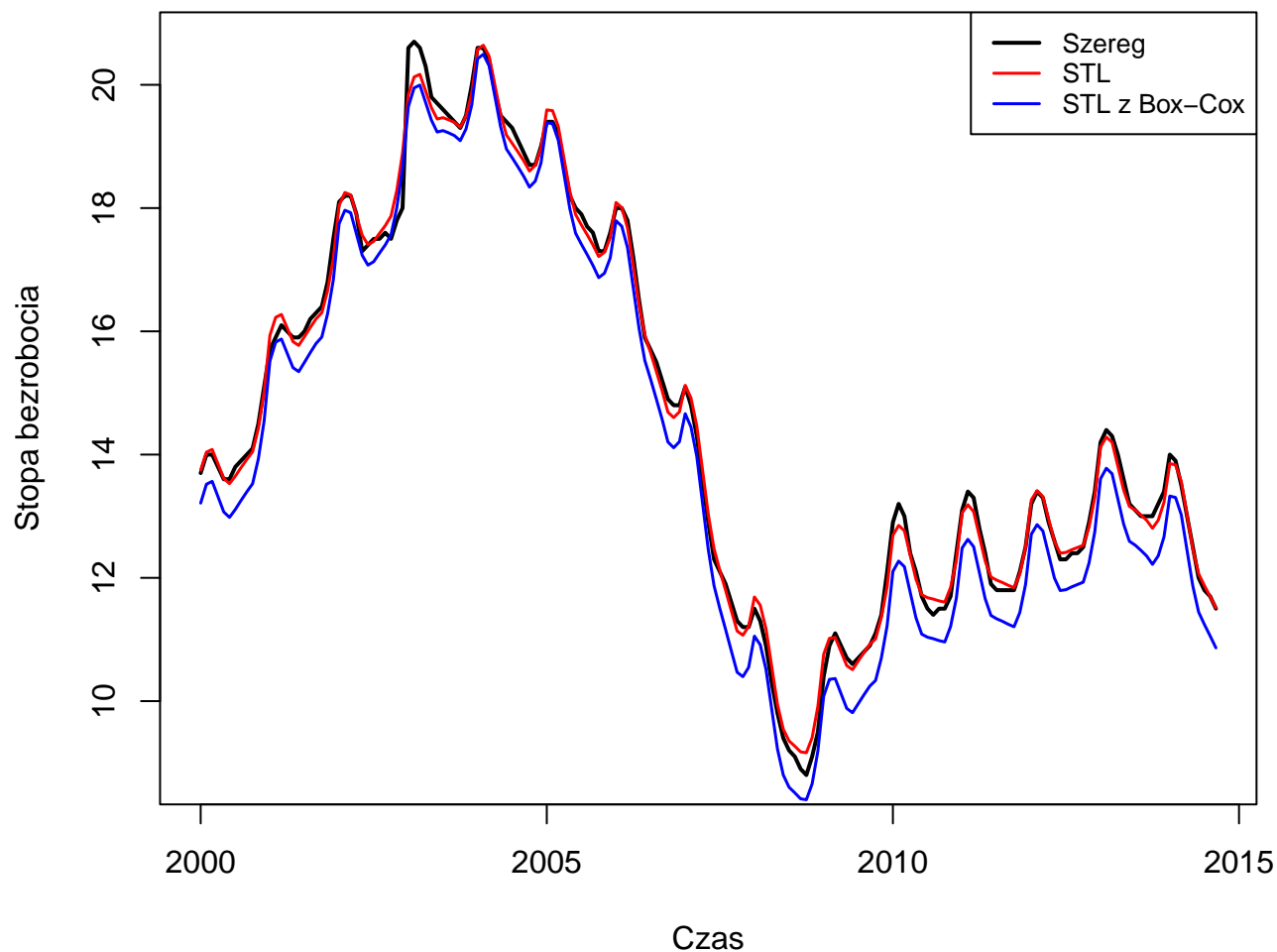
2.3.4 Wpływ transformacji Box’a-Cox’a

Transformacja Boxa–Coxa jest klasycznym narzędziem służącym do stabilizacji wariancji i poprawy zgodności szeregu z założeniami modeli liniowych oraz metod dekompozycji. Jej zastosowanie może zmniejszyć skośność rozkładu, ograniczyć wpływ wartości odstających oraz sprawić, że błędy modelu będą bardziej jednorodne. Aby ocenić, czy transformacja poprawia jakość dopasowania, porównamy dekompozycję szeregu w wersji oryginalnej i przetransformowanej.



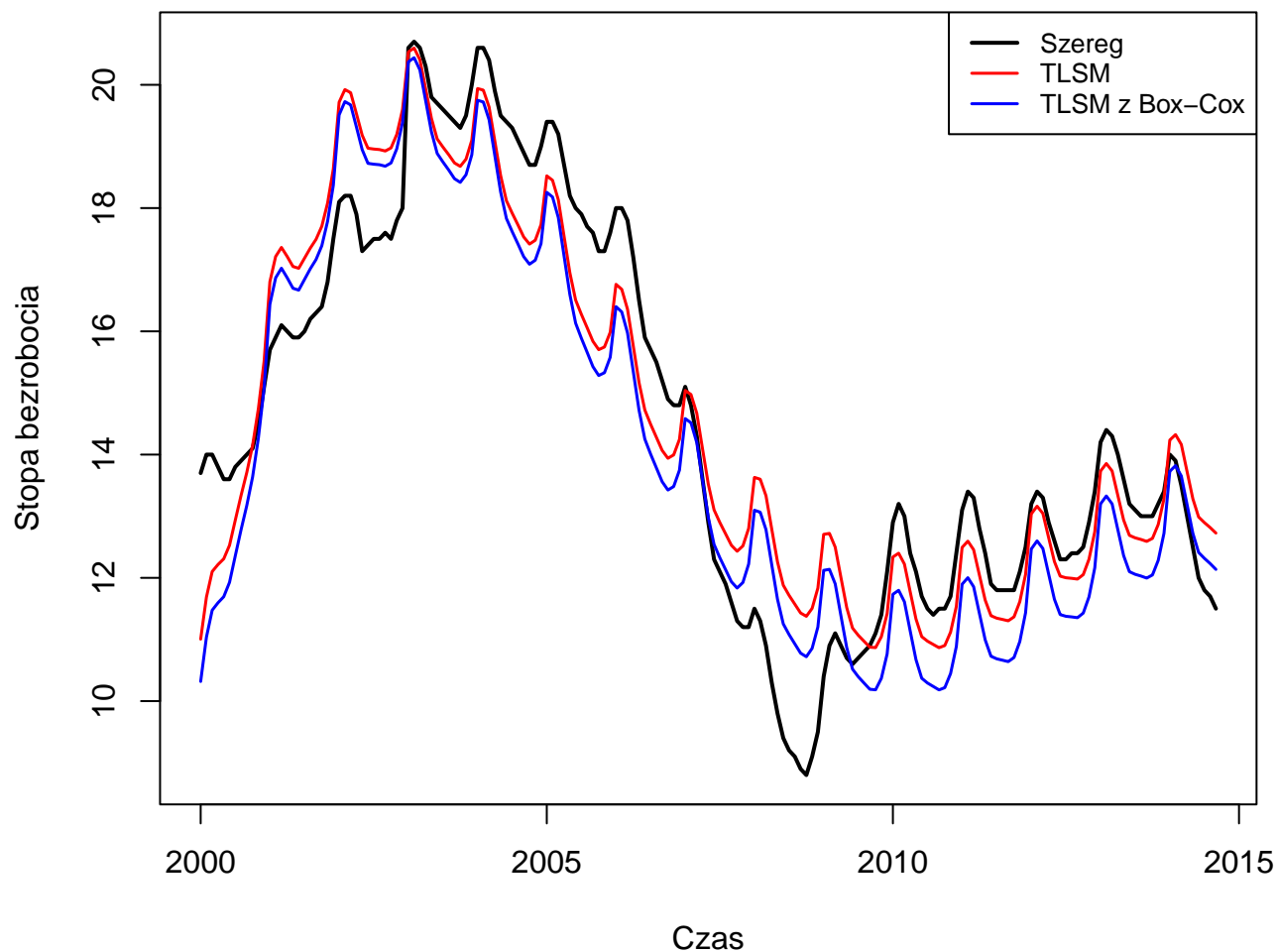
Rysunek 11: Porównanie oryginalnego szeregu i szeregu po transformacji Boxa-Coxa

Na wykresach 11 widać, że transformacja Boxa-Coxa prowadzi do zmniejszenia zmienności. Amplituda krótkoterminowych wahań jest mniejsza.



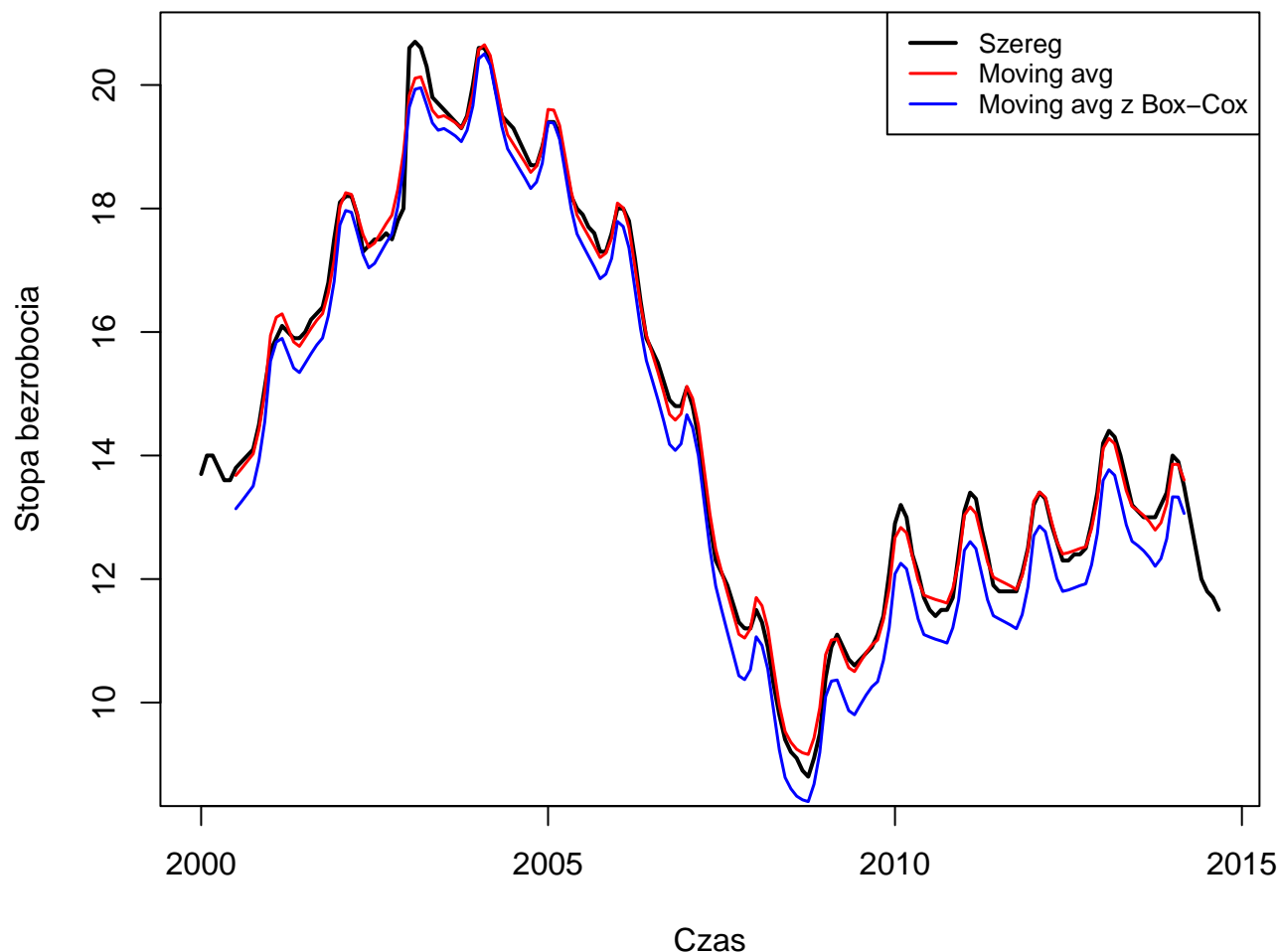
Rysunek 12: Porównanie dekompozycji STL bez i z Transformacją Boxa-Coxa

Dla dekompozycji STL przedstawionej na rysunku 12 obserwujemy, że zastosowanie transformacji Boxa-Coxa nie poprawia jakości dopasowania, a wręcz prowadzi do jego pogorszenia. Po transformacji trend staje się zbyt wygładzony i zaczyna odbiegać od rzeczywistego kształtu szeregu. Najbardziej widoczne jest to w latach 2009–2015, gdzie oryginalny STL dobrze odzwierciedla zarówno rosnący, jak i później malejący trend. Tymczasem wersja po transformacji Boxa-Coxa wygładza te zmiany zbyt mocno, przez co trend reaguje z opóźnieniem.



Rysunek 13: Porównanie dekompozycji TLSM z wielomianem 4 stopnia bez i z Transformacją Boxa-Coxa

W przypadku modeli parametrycznych opartych na regresji z rysunku 13 transformacja Boxa-Coxa może redukować autokorelację reszt i poprawiać dopasowanie trendu szczególnie przy wielomianach wyższych rzędów. Mimo to w naszym przypadku wielomianu 4 stopnia, znaczna większość trendu jest niedoszacowana lub przeszacowana. Może to też wynikać z faktu, że model bez transformacji również słabo dopasowuje się do szeregu.



Rysunek 14: Porównanie dekompozycji ruchomej średniej bez i z Transformacją Boxa-Coxa

W dekompozycji opartej na ruchomej średniej z rysunku 14 transformacja Boxa-Coxa również nie poprawia jakości dopasowania. Wręcz przeciwnie, dopasowanie trendu po transformacji staje się wyraźnie słabsze niż w przypadku analizy wykonanej na szeregu oryginalnym. Trend uzyskany po transformacji jest nadmiernie wygładzony i gorzej odwzorowuje rzeczywiste zmiany poziomu bezrobocia.

Podsumowując, w przypadku badanego szeregu transformacja Boxa-Coxa nie prowadzi do poprawy jakości dekompozycji, a w wielu sytuacjach wręcz osłabia dopasowanie modeli. Choć transformacja skutecznie stabilizuje wariancję i krótkookresowe wahania szeregu, to jednocześnie nadmiernie wygładza strukturę danych, co negatywnie wpływa na rekonstrukcję trendu.

2.4 Porównanie wyników eliminacji trendu i sezonowości

Teraz porównamy reszty dla wszystkich 3 typów dekompozycji. Na początek przeprowadzimy test białoszumowości.

Wynik dla dekompozycji opartej na ruchomej średniej:

```
##
## Box-Ljung test
##
## data:  res_add
## X-squared = 63.275, df = 1, p-value = 1.776e-15
```

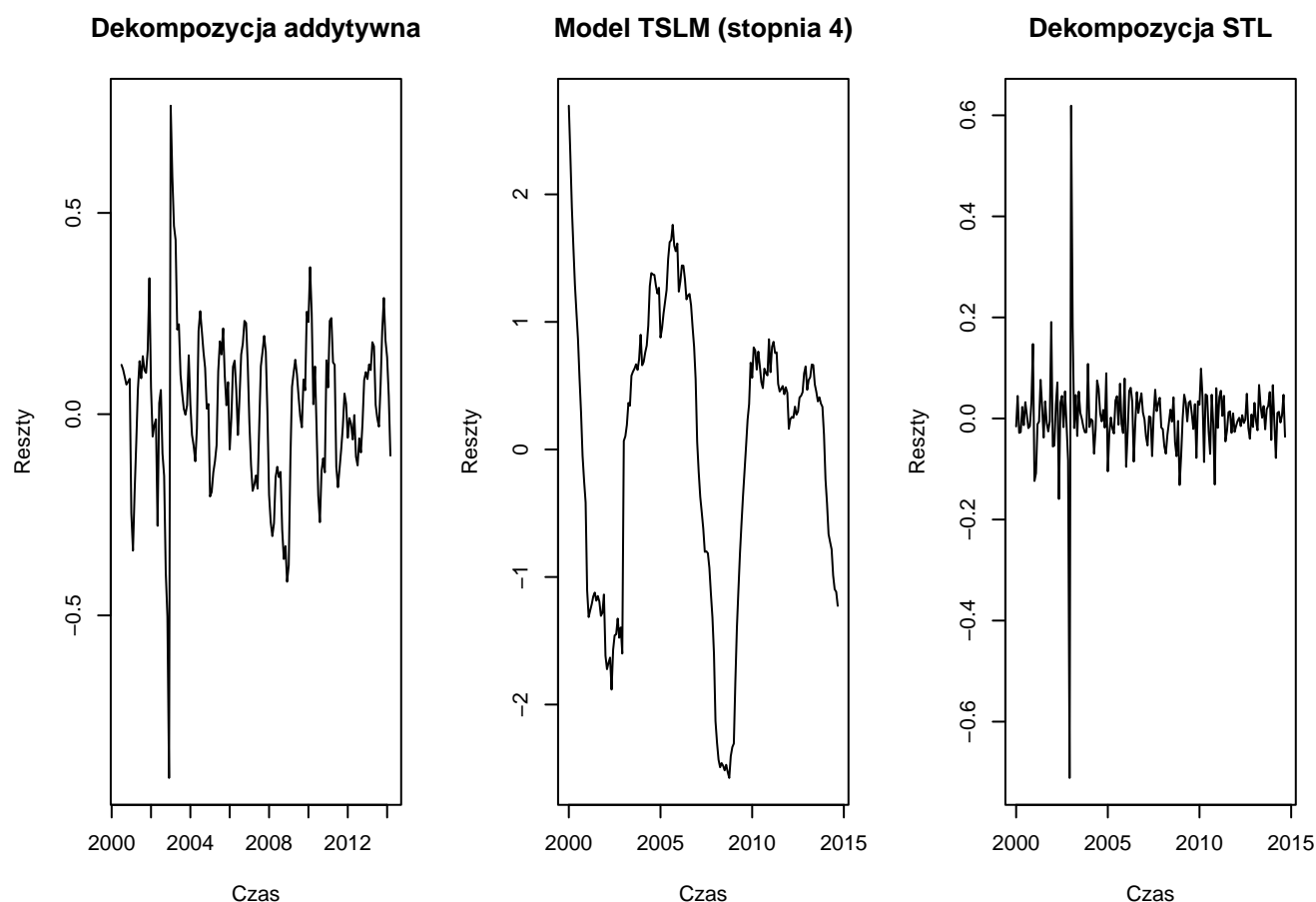
Wynik dla dekompozycji o model wielomianu 4 stopnia:

```
##
## Box-Ljung test
##
## data:  res_tslm
## X-squared = 167.4, df = 1, p-value < 2.2e-16
```

Wynik dla dekompozycji STL:

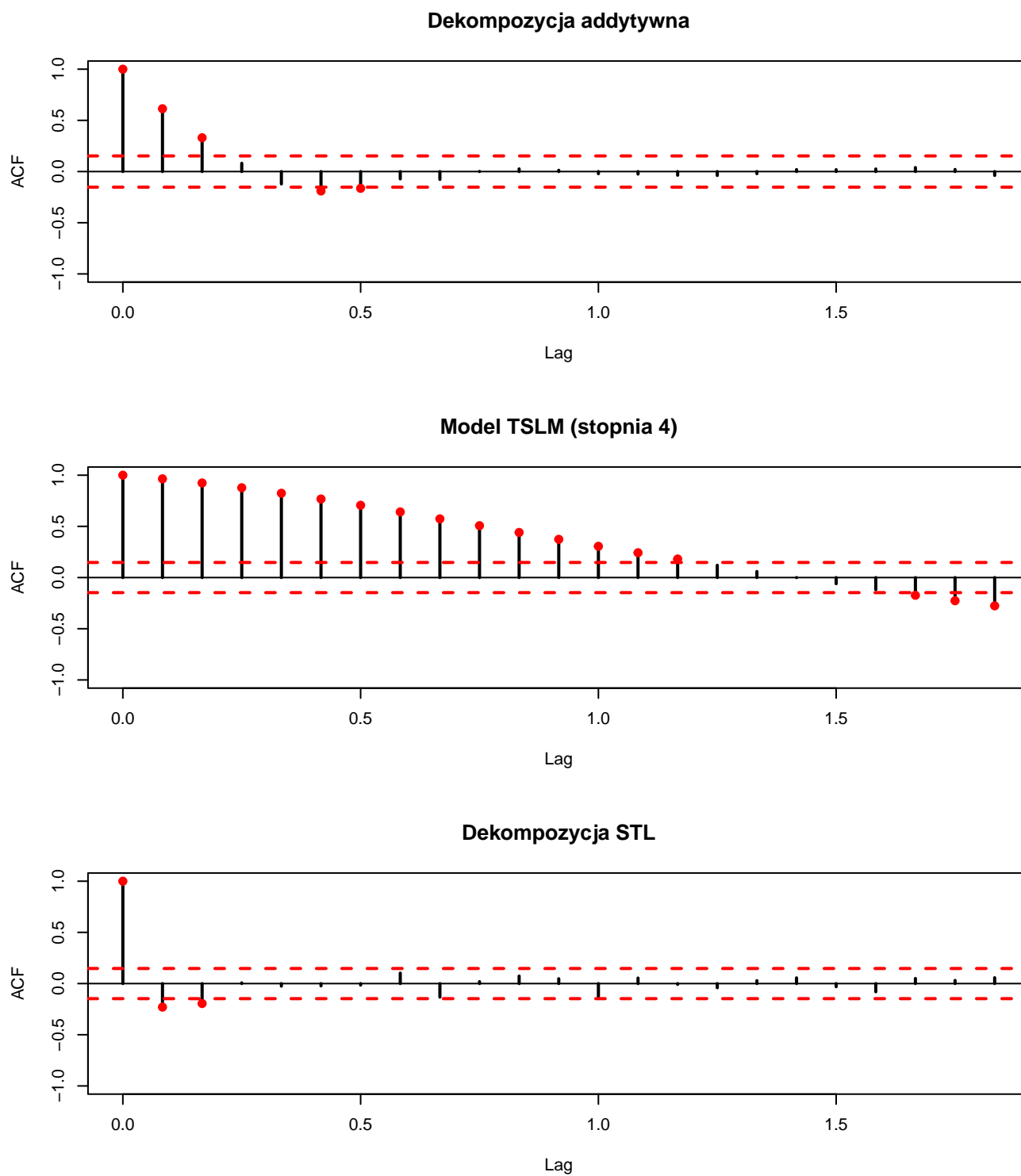
```
##
## Box-Ljung test
##
## data:  res_stl
## X-squared = 9.5252, df = 1, p-value = 0.002027
```

Aby nie polegać tylko na testach Ljunga-Box'a, zwizualizujemy reszty na wykresach.



Rysunek 15: Wykresy reszt dla każdej dekompozycji

Analizując wykresy przedstawione na rysunku 15, można stwierdzić, że reszty pochodzące z dekompozycji opartej na modelu nie zachowują się jak biały szum. Wyraźnie widoczny jest w nich trend. W przypadku dekompozycji STL reszty wydają się całkowicie losowe, co sugeruje, że można je uznać za biały szum. Dla dekompozycji addytywnej wykres reszt wygląda w przybliżeniu jak biały szum, jednak formalny test Ljunga wskazuje na odrzucenie hipotezy zerowej, co oznacza, że reszty nie są w pełni losowe. Oznacza to, że metoda STL najlepiej radzi sobie z usunięciem struktury z szeregu czasowego, pozostawiając reszty losowe. Natomiast w dekompozycji addytywnej i modelu klasycznym pewne wzorce pozostają niewyjaśnione.



Rysunek 16: Wykresy ACF dla reszt każdej dekompozycji

Wykresy ACF dla reszt przedstawione na rysunku 16 potwierdzają wcześniejsze wnioski z testu Ljunga-Boxa i analizy wykresów reszt. Analiza funkcji autokorelacji wyraźnie wskazuje na obecność istotnych zależności w resztach z modelu TSLM oraz dekompozycji addytywnej, gdzie wartości ACF wielokrotnie przekraczają granice istotności. W przeciwieństwie do nich, wykres ACF dla dekompozycji STL nie wykazuje żadnych istotnych autokorelacji, co potwierdza jej skuteczność w eliminacji struktury z danych. Tym samym metoda STL okazuje się najbardziej efektywna w usuwaniu trendu i sezonowości, pozostawiając reszty najbliższe białemu szumowi.

2.5 Wnioski końcowe

Przeprowadzona analiza szeregu **bezrobocia** pozwala wyciągnąć kluczowe wnioski dotyczące efektywności różnych metod dekompozycji. Spośród porównywanych podejść, dekompozycja STL okazała się najskuteczniejsza w eliminacji trendu i sezonowości, pozostawiając reszty najbardziej zbliżone do białego szumu. Metody oparte na ruchomej średniej i modelach TSLM, choć poprawne teoretycznie, nie radziły sobie równie dobrze z niestacjonarnością danych. Transformacja Boxa-Coxa nie przyniosła oczekiwanej poprawy jakości dekompozycji, a wręcz pogorszyła dopasowanie trendu w większości przypadków. Analiza wykresów ACF jednoznacznie potwierdziła, że jedynie reszty z dekompozycji STL można uznać za wystarczająco zbliżone do białego szumu dla celów praktycznych.