

**WYDZIAŁ, PODSTAWOWYCH PROBLEMÓW TECHNIKI
POLITECHNIKI WROCŁAWSKIEJ**

SPRAWOZDANIE

JACEK MUCHA

Analiza Techniczna Dow Chemicals Company
Model AR dla średniej temperatury globalnej

WROCŁAW 2014

Spis treści

1	Wskaźniki analizy technicznej	2
1.1	Moving Average Convergence/Divergence (MACD)	2
1.2	Oscylator stochastyczny STS	2
1.3	Wstęga Bollingera	2
1.4	Average True Range (ATR)	3
1.5	Zmienność Chaikina	3
1.6	Indykator Aroon	3
1.7	Chaikin AD	3
1.8	Chaikin Money Flow (CMF)	4
1.9	Money Flow Index (MFI)	4
1.10	Commodity Channel Index (CCI)	4
1.11	Relative Strength Index(RSI)	4
2	Analiza techniczna Dow Chemicals Company	5
2.1	MACD	5
2.2	STS	7
2.3	Wstęga Bollingera	8
2.4	ATR	9
2.5	Chaikin's volatility	10
2.6	Aroon	11
2.7	CMF	12
2.8	MFI	13
2.9	CCI	14
2.10	RSI	14
2.11	Podsumowanie	15
3	Dopasowywanie modelu autoregresji - teoria	15
3.1	Estymatory Yule'a-Walkera	15
3.2	FPE (Final Prediction Error)	16
4	Dopasowanie modelu autoregresji - przykład	16
4.1	Metoda PACF	19
4.2	Metoda AIC	21
4.3	Metoda FPE	23
4.4	AR(3) i AR(4) ze współczynnikami wyestymowanymi innymi metodami	23
4.5	Wnioski	24

1 Wskaźniki analizy technicznej

1.1 Moving Average Convergence/Divergence (MACD)

Wskaźnik oparty na cenie; wykorzystuje różnicę dwóch eksponencjalnych średnich ruchomych - długookresowej i krótkookresowej, najczęściej - odpowiednio - 26- i 12- punktowa.

$$shortEMA = 0.15 \cdot price + 0.85 \cdot shortEMA_{[-1]}$$

$$longEMA = 0.075 \cdot price + 0.925 \cdot longEMA_{[-1]}$$

$$MACD = shortEMA - longEMA$$

Interpretacja:

- Gdy średnia przebija z dołu poziom 0, przyjmuje się, że trend będzie wzrostowy (kupować)
- Gdy średnia przebija z góry poziom 0, oczekuje się trendu zniżkowego (sprzedawać)
- Histogram ze słupkami nad poziomem 0 interpretowany jest jako trend o tendencji wzrostowej. Poniżej tego poziomu - trend spadkowy.

1.2 Oscylator stochastyczny STS

Służy do badania wykupienia i wyprzedania rynku. Wykorzystywany w krótkoterminowej analizie kursu. Przyjmuje wartości ze zbioru $[0, 100]$.

$$STS = 100 \frac{C - L_{14}}{H_{14} - L_{14}},$$

gdzie C - ostatnia cena zamknięcia, L_{14} - minimum z ostatnich 14 sesji, H_{14} - maksimum z ostatnich 14 sesji. Na wykresie umieszcza się także linię *signal* - średnią ruchomą z trzech dni.

Zastosowanie praktyczne:

- Kupować, gdy STS przecina z dołu linię *signal*
- Sprzedawać, gdy STS przecina z góry linię *signal*
- Wartość powyżej 80 oznacza wykupienie rynku (przebiecie od góry: sprzedawaj)
- Wartość poniżej 20 oznacza wyprzedanie rynku (przebiecie od dołu: kupuj)

1.3 Wstęga Bollingera

Wskaźnik zmienności trendu. Środkowa linia to średnia krocząca z n sesji. Górna wstęga: k -krotność n -okresowego odchylenia standardowego powyżej środkowej wstęgi; wstęga dolna - odpowiednio k -krotność n -okresowego odchylenia standardowego poniżej środkowej wstęgi. Do policzenia średniej ruchomej i zmienności bierze się cenę typową $p_t = \frac{H+L+C}{3}$, tj. średnią arytmetyczną z dziennego maksimum, minimum i ceny zamknięcia.

Rozszerzanie się wstęgi oznacza rosnące ryzyko gwałtownych ruchów, sugeruje zmniejszenie pozycji. Zwężanie - przeciwnie. Zwężanie stanowi ostrzeżenie przed silnym ruchem w przyszłości. Przy szerokiej wstędze sugeruje się kupować przy dolnej wstędze i sprzedawać przy górnej.

1.4 Average True Range (ATR)

ATR jest średnią z rzeczywistego zakresu zmian TR .

$$TR = \max(H_{[0]}, C_{[-1]}) - \min((L_{[0]}, C_{[-1]}))$$

$$ATR = \frac{(n-1)TR_{[-1]} + TR}{n}$$

Wysoka wartość wskaźnika podczas gwałtownego wzrostu lub spadku może zapowiadać dłuższą zmianę trendu; niskie wartości są prognozą dla trendu horyzontalnego.

1.5 Zmienność Chaikina

Algorytm polega na badaniu różnic $H - L$ na przestrzeni n dni (najczęściej $n = 10$). Z tych różnic wyznacza się średnią kroczącą EMAHL. Wówczas uważa się, że duże wartości wskaźnika oznaczają przełamanie trendu horyzontalnego, zaś niskie wartości oznaczają ustabilizowanie się emocji inwestorów, a w związku z tym sugerują mniejszą zmienność w przyszłości. Inna interpretacja mówi, że wzrost różnic między kursem maksymalnym i minimalnym w ciągu krótkiego okresu czasu może zwiastować nadchodzącą gwałtowną bessę, zaś spadek tychże różnic - nadchodzący szczyt.

$$ChaikinVolatility = \frac{EMAHL - EMAHL_{[-n]}}{100 \cdot EMAHL_{[-n]}}$$

1.6 Indykator Aroon

Są to dwie linie, które pokazują, jak dużo czasu (unormowanego) upłynęło od ostatniej zanotowanej najwyższej i najniższej wartości. Każdy okres długości T , w którym nie pojawiło się notowanie z nową ekstremalną wartością, powoduje obniżenie wskaźnika o $\frac{100}{T}$.

$$AroonUp = 100 \cdot \frac{n - PeriodsSinceHighestHigh}{n}$$

$$AroonDown = 100 \cdot \frac{n - PeriodsSinceLowestLow}{n}$$

Przecinanie się linii sugeruje odwrócenie trendu; zwłaszcza przecięcie powyżej 50 - początek trendu rosnącego; przecięcie poniżej 50 - malejącego. Trend rosnący występuje, jeśli składowa Up przyjmuje wartość powyżej 70, zaś malejący - $Down > 70$. Silny trend występuje wówczas, gdy jedna składowa jest duża, podczas gdy druga - mała.

1.7 Chaikin AD

Wskaźnik oparty na wolumenie wyznaczany jako różnice dwóch średnich wykładniczych; określa przepływ kapitału na rynku. Zazwyczaj używa się go jako pomoc przy interpretacji innych wskaźników; w niniejszej analizie posłuży jako wprowadzenie do CMF.

$$CLV = \frac{(C - L) - (H - C)}{H - L}$$

$$ChaikinAD = EMA(CLV \cdot V; 3) - EMA(CLV \cdot V; 10),$$

gdzie V - wolumen.

1.8 Chaikin Money Flow (CMF)

Jest to suma CAD z ostatnich n okresów (najczęściej $n = 21$ podzielona przez sumę wolumenów z tego okresu.

$$CMF_i = \frac{\sum_{i-n}^i (CLV \cdot V)}{\sum_{i-n}^i V}$$

Interpretacja:

- $CMF > 0.25 \implies$ kupować
- $CMF < -0.25 \implies$ sprzedawać
- Wybicie kursu ponad opór przez dodatnie wartości wskaźnika - zapowiedź wzrostu kursu
- Przebicie poziomu wsparcia przez wykres kursu za pomocą ujemnych wartości wskaźnika - zapowiedź trendu malejącego.

1.9 Money Flow Index (MFI)

Indeks przepływów pieniężnych.

$$MoneyFlow = p_t \cdot V$$

Jeśli $p_t > p_{t[-1]}$, to $PositiveMoneyFlow = PositiveMoneyFlow_{[-1]} + MoneyFlow$. W przeciwnym wypadku: $NegativeMoneyFlow = NegativeMoneyFlow_{[-1]} + MoneyFlow$. Teraz $MoneyRatio = \frac{PositiveMoneyFlow}{NegativeMoneyFlow}$

$$MFI = 100 - \left(\frac{100}{1 + MoneyRatio} \right)$$

Interpretacja:

- $MFI > 80$ - przewartościowanie rynku; sprzedawać
- $MFI < 20$ - niedowartościowanie rynku; kupować

1.10 Commodity Channel Index (CCI)

Bazuje na kursie typowym.

$$CCI = \frac{p_t - SMA(p_t)}{0.015 \cdot \sigma(p_t)}$$

Interpretacja: $CCI > 100$ - zapowiada trend zniżkowy. $CCI < -100$ - prognoza dla trendu wyżującego.

1.11 Relative Strength Index(RSI)

RS - stosunek średniej wartości wzrostu cen zamknięcia do średniej wartości ich spadku z n okresów.

$$RSI = 100 - \frac{100}{1 + RS}$$

Interpretacja:

- RSI przebija od dołu poziom wyprzedania (20-30) - kupować
- RSI przebija od góry poziom wykupienia (70-80)- sprzedawać

2 Analiza techniczna Dow Chemicals Company

Analizie zostaną poddane notowania akcji korporacji Dow Chemicals. Na przestrzeni kilku ostatnich lat ceny akcji i wolumen kształtowały się następująco:



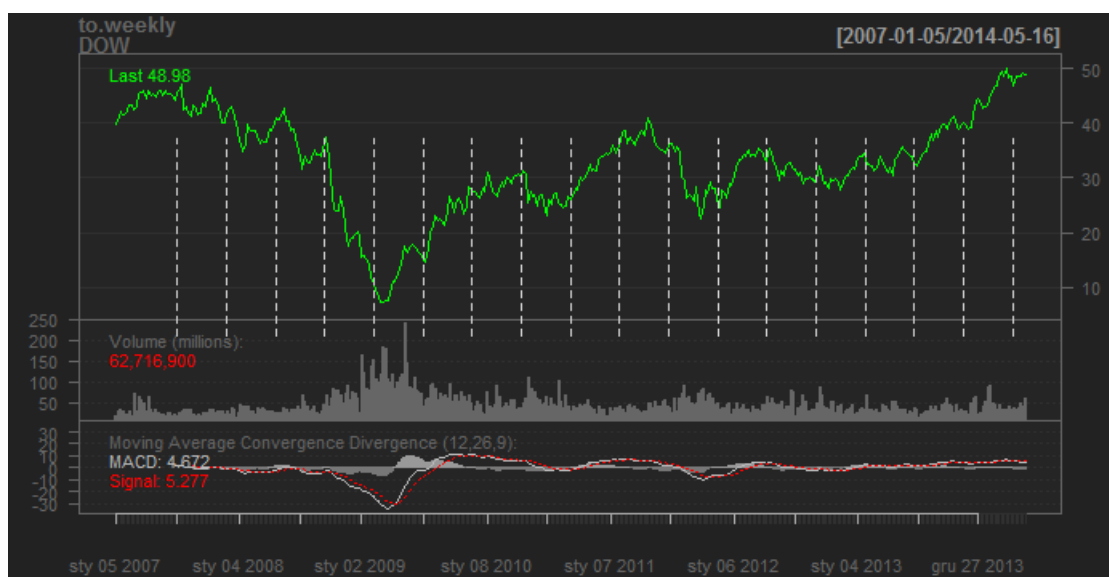
Wykres 1. Notowania korporacji Dow Chemicals Company.

Można wyróżnić kilka wyraźnych okresów kształtowania się kursu: (i) styczeń 2007-wrzesień 2008 - umiarkowany spadek ceny, wolumen na stałym poziomie; (ii) wrzesień 2008-marzec 2009 - mocny spadek cen, istotny wzrost wolumenu; (iii) marzec 2009-styczeń 2010 - szybki wzrost cen i powoli spadający wolumen; (iv) od stycznia 2010, po początkowych zawirowaniach, w miarę stabilny wzrost cen.

Można wspomnieć o ciekawostce z zakresu analizy fundamentalnej: we wrześniu 2008 upadł bank Lehman Brothers, co pociągnęło za sobą gwałtowny spadek indeksu Dow Jones. Stąd bierze się najciekawszy okres (ii) w poniższej analizie. Wskaźniki analizy technicznej skonfrontowałem w niniejszym sprawozdaniu z tą sytuacją nadzwyczajną, bo jakkolwiek jest ona pozamodelowa, to jednak inwestor wolałby raczej wskazówki dotyczące rzeczywistych problemów, a nie załamujących się podczas kryzysów matematycznych *toy models*...

2.1 MACD

Spróbujmy zastosować wskaźniki analizy technicznej do rozstrzygnięcia dylematów inwestycyjnych. Zaczniemy od MACD.



Wykres 2. MACD dla short=12, long=26, signal=9, agregacja tygodniowa

Na podstawie MACD (agregacja tygodniowa na wykresie, short=12, long=26, signal=9) można wyróżnić parę sygnałów zachęcających do przyjęcia krótkiej lub długiej pozycji przy agregacji tygodniowej. Pierwszym takim sygnałem było przecięcie przez wykres MACD linii *signal* od dołu około kwietnia 2009, co wskazywało, że należy kupować akcje. Mniej więcej w listopadzie 2009 pojawił się sygnał do sprzedaży. Postępując zgodnie z tą strategią można było kupić akcje po cenie 9, a sprzedać po cenie 25 (według skali na wykresie). Kolejny moment do zakupu miał miejsce w czerwcu 2010; akcje należało trzymać do grudnia 2010. Poza tym nie było wyraźniejszych sygnałów; może poza wrześniem 2008, kiedy to należało wyprzedać akcje w celu ochrony przed znacznymi stratami.

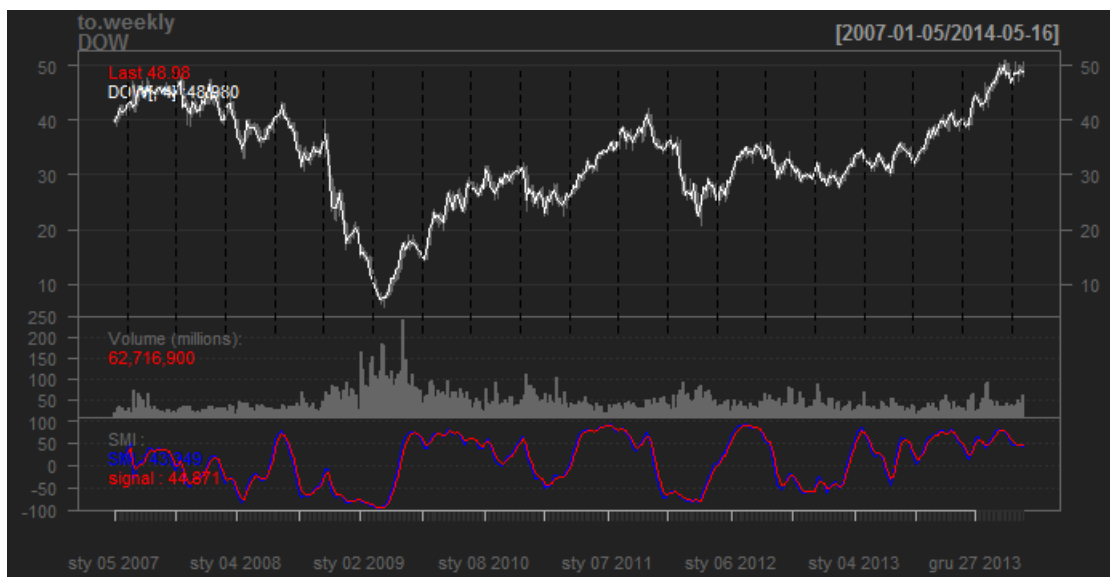


Wykres 3. MACD dla short=2, long=4, signal=4, agregacja miesięczna.

Na Wykresie 3. dobrze widać utrzymujące się trendy spadkowe i wzrostowe. Strategie inwestycyjne pozostają podobne, zysk również niewiele się zmienia. Sygnał alarmowy do wyprzedawania akcji pojawia się w czerwcu 2008 (a więc wcześniej!), następnie w marcu 2009 (wcześniej!) sy-

gnał do zakupu i we wrześniu 2009 - sygnał sprzedaży. Parametry dla MACD dobrane do Wykresu 3. (z agregacją miesięczną, short=2, long=3, signal=4) wydają się zatem zapewniać strategię nieco ostrożniejszą niż prezentowana wcześniej.

2.2 STS



Wykres 4. STS, agregacja tygodniowa, SMI (Smart Money Index, inna nazwa dla STS) w pakiecie R wywołana z parametrami $n = 14$, $nFast = 2$, $nSlow = 4$, $nSig = 3$; linia *signal* jest czerwona. Wykres STS jest przeskalowany z $[0, 100]$ do $[-100, 100]$.

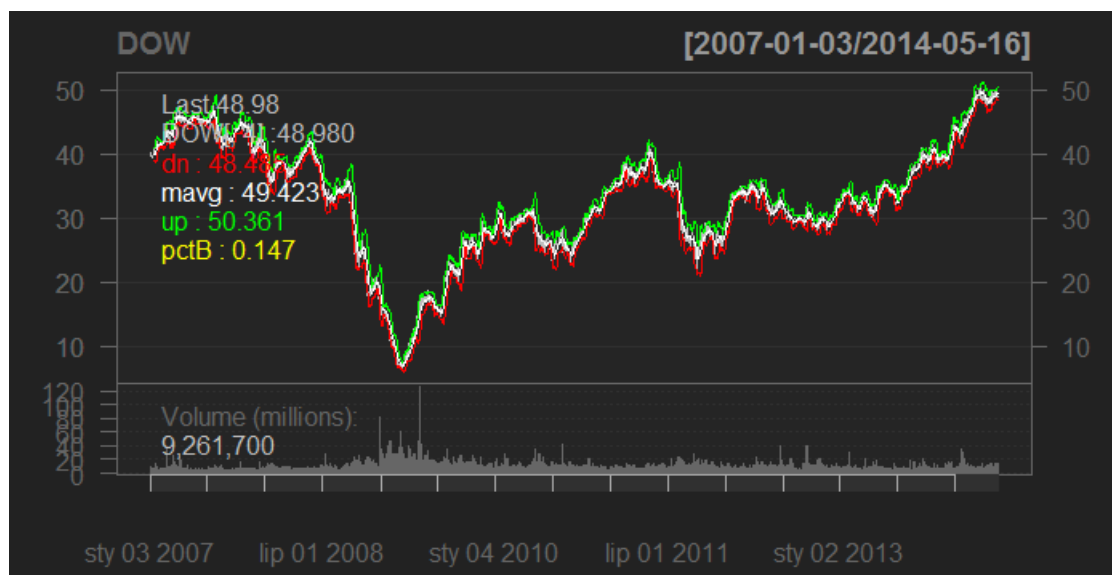
STS daje pierwszy wyraźny sygnał do sprzedaży mniej więcej w czerwcu 2008, kiedy, po pierwsze, mamy informację o wykupieniu rynku (przecięcie od góry poziomu 60), a po drugie, linia *signal* zostaje przecięta od góry. Mniej więcej w maju 2009 pojawia się sygnał do zakupu i we wrześniu 2009 - sygnał sprzedaży. Dobrane parametry uczyniły STS bardzo wrażliwym na zmiany trendu. Kolejne sygnały: maj 2011 (sprzedawaj), listopad 2011 (kupuj), lipiec 2012 (kupuj), marzec 2013 (kupuj).



Wykres 5. STS, agregacja miesięczna, SMI (Smart Money Index, inna nazwa dla STS) w pakiecie R wywołana z parametrami $n = 14$, $nFast = 2$, $nSlow = 4$, $nSig = 3$; linia *signal* jest czerwona. Wykres STS jest przeskalowany z $[0, 100]$ do $[-100, 100]$.

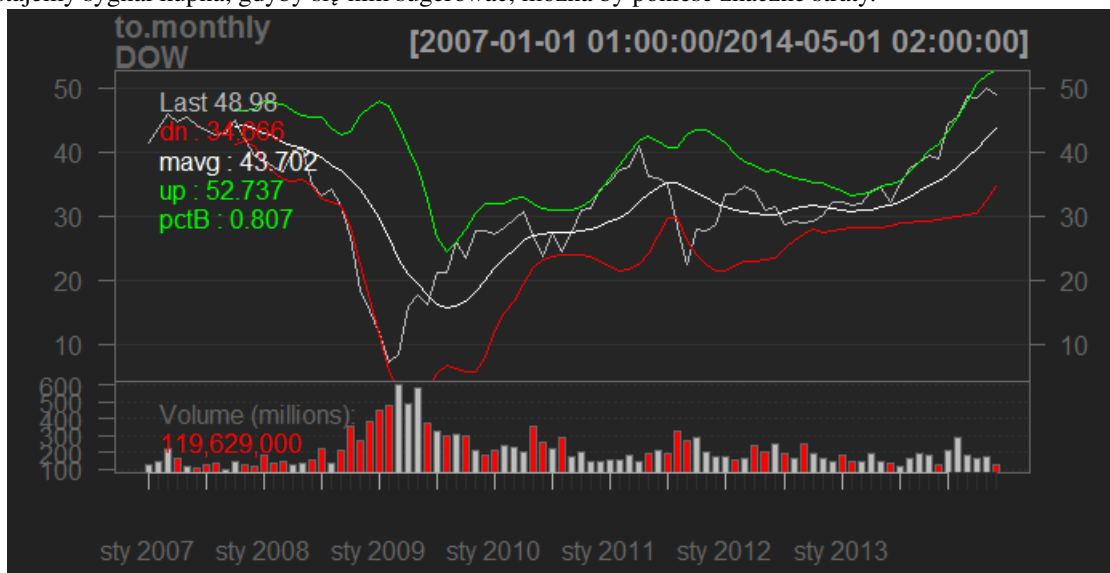
Wykres 5. z agregacją miesięczną wygładza linie SMI oraz *signal*. Zakładając, że kupujemy, jeśli zarówno SMI przekroczy od dołu poziom -60, jak i przecnie od dołu linię *signal*, otrzymujemy znacznie mniej sygnałów do kupna (K) i sprzedaży (S), mianowicie: kwiecień 2009 (K), maj 2010 (S), maj 2011 (S). Gdyby posługiwać się tylko kryterium przecięcia linii *signal*, można by, po pierwsze, uchronić się przed bessą od kwietnia 2008 do stycznia 2009, a po drugie, wykorzystać hossę trwającą od marca 2013. W takim razie poziom wykupienia rynku przy agregacji miesięcznej nie wydaje się zbyt rozsądnym kryterium do podejmowania strategicznych decyzji.

2.3 Wstęga Bollingera



Wykres 6. Wstęga Bollingera, $n = 10$, $k = 2$, brak agregacji.

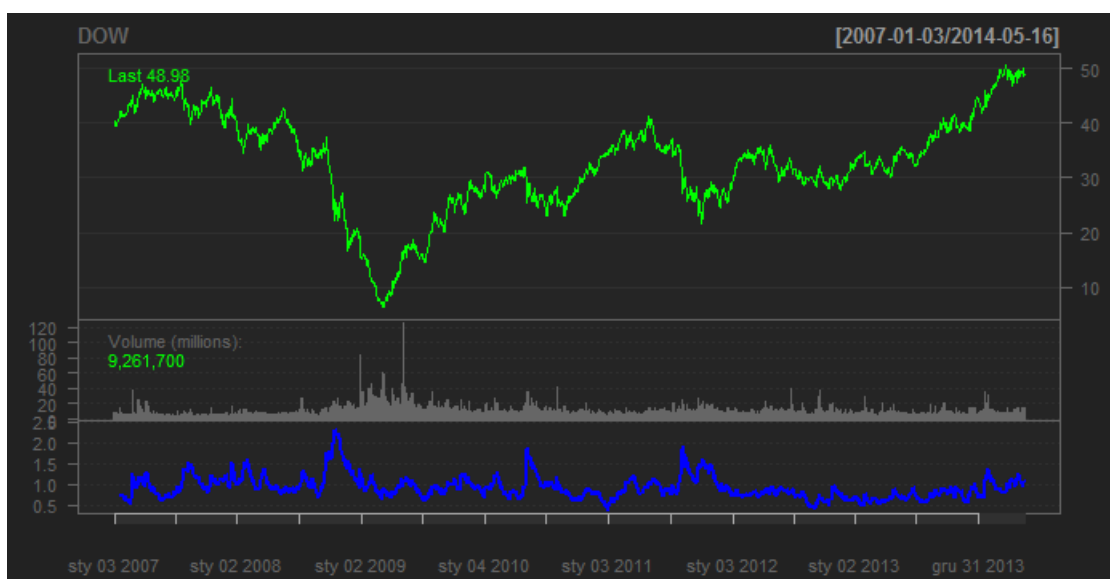
Wykres 6. jest dynamiczny, z uwagi na brak agregacji. Kupno: szara linia dotyka lub przecina linię czerwoną. Sprzedaż: szara linia przecina linię zieloną. Niestety w okresie (ii) wcześnie dostajemy sygnał kupna; gdyby się nim sugerować, można by ponieść znaczne straty.



Wykres 7. Wstęga Bollingera, $n = 10$, $k = 2$, agregacja miesięczna.

Czy straty, jakie ponieśliśmy, sugerując się wykresem 6., są spowodowane brakiem agregacji? Na podstawie wykresu 7. należałoby raczej odrzucić tę hipotezę. Mimo agregacji miesięcznej otrzymujemy fałszywe sygnały. Stąd wskaźnik Bollingera w tym przypadku nie działał dobrze. Myślę, że było to spowodowane silnymi trendami krótkookresowymi.

2.4 ATR



Wykres 8. ATR bez agregacji.

Wskaźnik ATR dobrze zaprognozował spadki w okresie (ii); ponadto notowania w okresie (iv) zgadzają się z przewidywaniami. Można podejrzewać, że zmienność trendu będzie rosła w roku 2014.



Wykres 9. ATR z agregacją miesięczną.

Agregacja miesięczna pozwala lepiej rozeźnić zawirowania między lipcem 2011, a styczniem 2012. Wyraźnie widać stabilizację trendu i wariancji, kiedy poziom wskaźnika zbliża się do zera.

2.5 Chaikin's volatility

Aby nie zaciemniać analizy, pomijam wykresy bez agregacji i z agregacją miesięczną, z których w przypadku tego wskaźnika niewiele wynika. Przyjrzyjmy się agregacji tygodniowej:

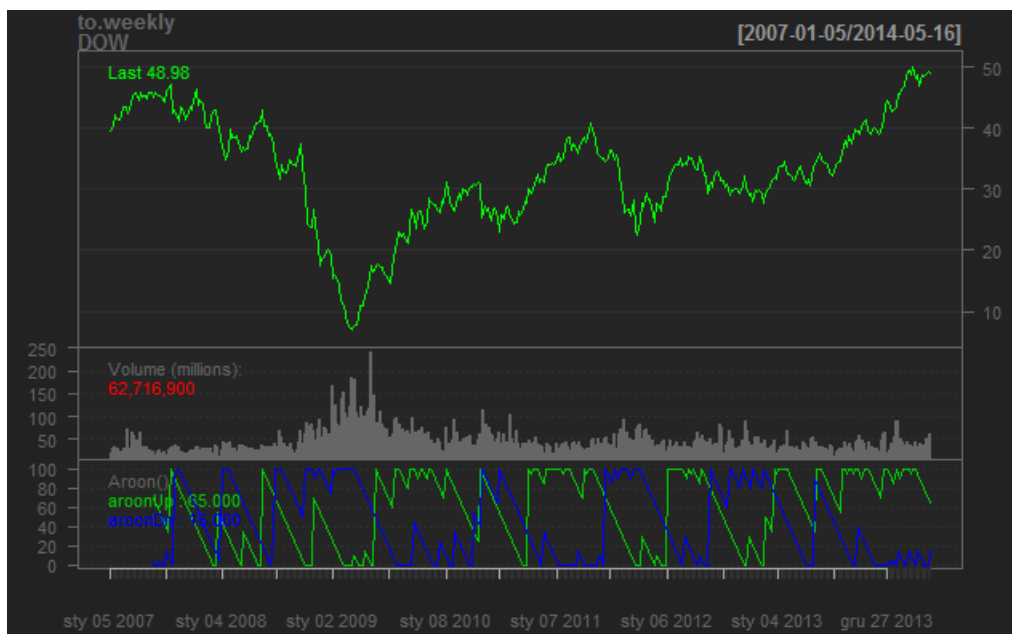


Wykres 10. Chaikin's volatility z agregacją tygodniową.

Wskaźnik dobrze określił szczyt tuż przed znacznymi spadkami okresu (ii), jak również dobrze wskazał na koniec spadków (z pewnym opóźnieniem). We wrześniu 2009 dał sygnał fałszywy - na

podstawie wcześniejszych obserwacji uznalibyśmy, że należy sprzedać akcje, a jednak przyszłość tego nie potwierdziła.

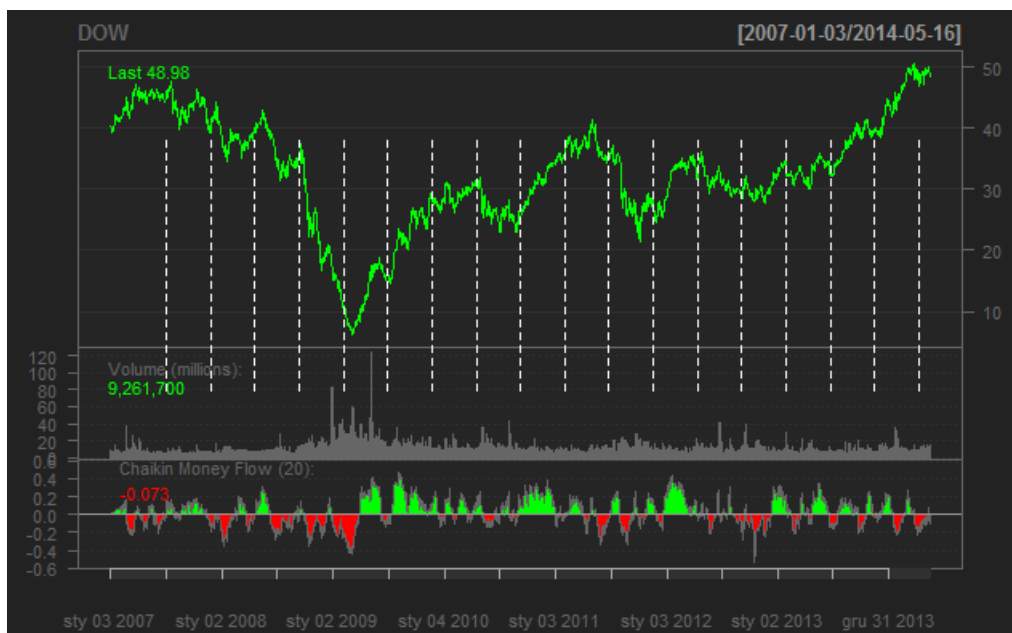
2.6 Aroon



Wykres 11. Indykator Aroona z agregacją tygodniową.

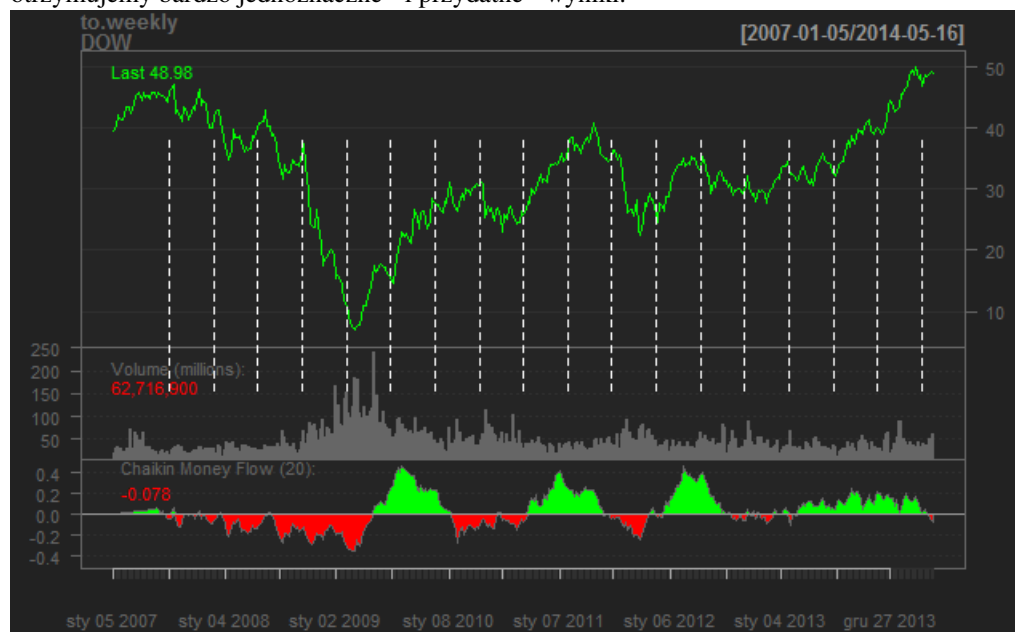
Podobnie jak powyżej, agregacja miesięczna i brak agregacji zaciemniały obraz. Agregacja tygodniowa pozwoliła wyróżnić istotne zmiany trendu z rosnącego na malejący. Zwłaszcza cenne okazało się wczesne (choć można powiedzieć: przedwczesne) ostrzeżenie przed sygnałem spadkowym okresu (ii). Z pewnym opóźnieniem przychodzi sygnał do kupowania akcji po zakończeniu okresu (ii), ale za to spadki w lipcu 2011 i rosnący trend ostatniego okresu zostały rozpoznane bezbłędnie.

2.7 CMF



Wykres 12. CMF bez agregacji, n domyślne (20).

Wskaźnik ten bez agregacji dał fałszywe sygnały, przede wszystkim bardzo wczesny sygnał sprzedaży przed okresem (ii) oraz sygnał sprzedaży we wrześniu 2012 (wykres znacznie poniżej poziomu -0.25). Zasadniczo jednak nie przysparzał strat, lecz zyski. Agregując dane tygodniowo, otrzymujemy bardzo jednoznaczne - i przydatne - wyniki.



Wykres 13. CMF, agregacja tygodniowa, n domyślne (20).

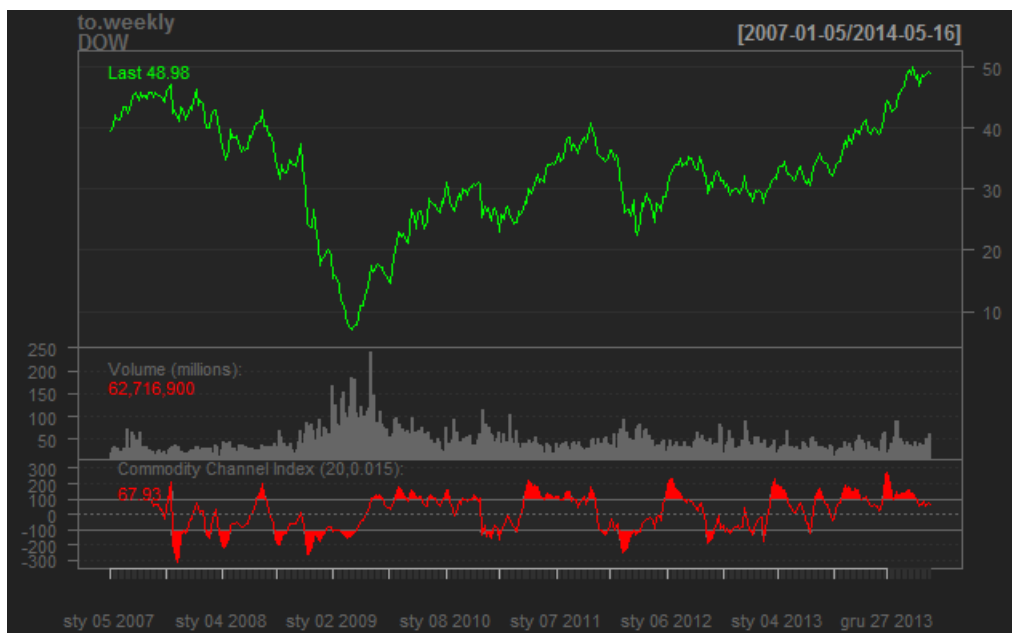
2.8 MFI



Wykres 14. MFI, agregacja tygodniowa

Wykres 14. jest wyjątkowy w porównaniu do poprzednich. Proponowana przez ekspertów interpretacja wskaźnika MFI nie sprawdziłaby się w żadnym z krytycznych przypadków. Linia MFI przebiega poziomo 80 i 20 w momentach zgoła dla inwestora zgubnych; jedynie pod koniec okresu (ii) sygnał kupowania daje oczekiwaną wskazówkę. Najwyraźniej jest to wskaźnik krótkookresowy - sygnał sprzedaży w lutym 2013, jakkolwiek wyraźny, pozwoliłby na bardzo niewielki zysk - i tylko w krótkim okresie.

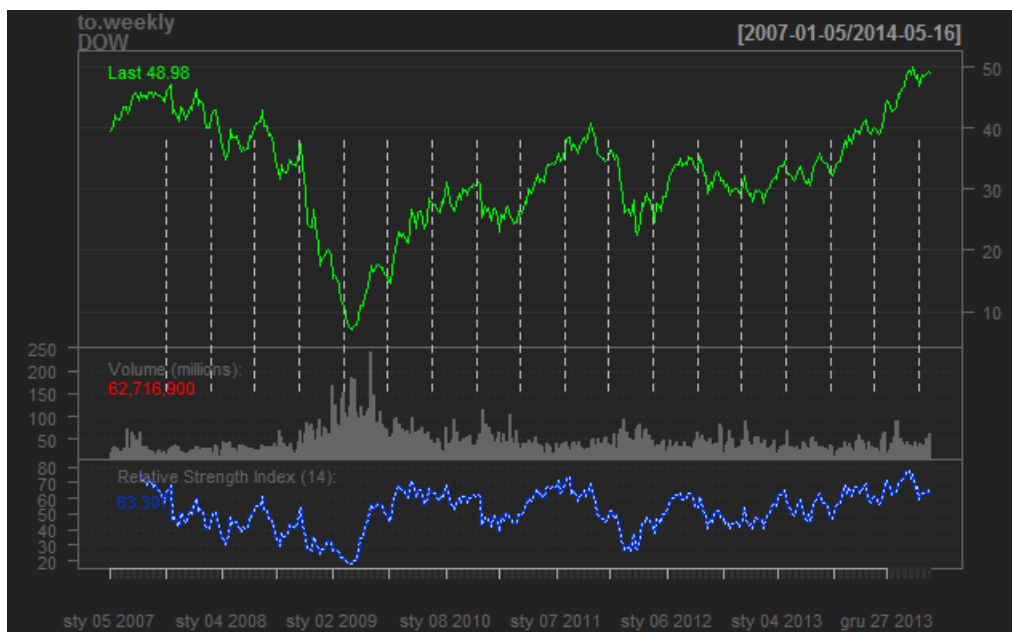
2.9 CCI



Wykres 15. CCI z agregacją tygodniową.

W obliczu kryzysu wskaźnik dał sygnał niedowartościowania ceny, co było informacją mylną. Jednocześnie zaobserwować można sygnały sprzedaży w czasie umacniania się trendu rosnącego.

2.10 RSI



Wykres 16. RSI z agregacją tygodniową.

Ustalmy poziom wyprzedania na 20. Wówczas wskaźnik wyraźnie daje sygnał do kupowania tuż po zakończeniu okresu (ii). Niestety wszystkie sygnały sprzedaży, nawet przy małym poziomie

wykupienia (rzędu 60-65), okazują się fałszywe w rozważanym przypadku.

2.11 Podsumowanie

Analizowaliśmy notowania akcji na przestrzeni 7 lat, co spowodowało, że pod uwagę braliśmy trendy miesięczne i roczne. Może to z tego względu większość wskaźników dawała bardziej pożądane sygnały (pozwalające zminimalizować stratę lub zmaksymalizować zysk) przy agregacji tygodniowej cen. Agregacja miesięczna i brak agregacji sprawiały, że wskazania rozważanych wskaźników nieraz wprowadzały w błąd. Stosowano najczęściej parametry domyślne dla wskaźników analizy technicznej - najczęściej bowiem okazywały się optymalne; ich zmiana powodowała nieraz zwiększenie lub zmniejszenie tolerancji na ryzyko. Część wskaźników wprowadzała w błąd niezależnie od dobranych parametrów (w tym agregacji) lub nie dawała żadnych sygnałów.

Trzeba pamiętać, że wskaźniki analizy technicznej najczęściej nie są estymatorami zgodnymi szacowanych wielkości; ich stosowanie opiera się na wieloletnich obserwacjach i założeniu, że "historia lubi się powtarzać". Za szeregiem czasowym reprezentującym notowania kryją się żywi ludzie, a jakkolwiek zjawiska socjologiczne i ekonomiczne w skali makro całkiem dobrze dają się modelować matematycznie, jednak zawsze model będzie tylko symplifikacją rzeczywistości do zjawisk przewidywalnych.

W kontekście analizy szeregów czasowych badane wskaźniki miały na celu najczęściej estymację trendu lub wariancji szeregu czasowego. Abstrahując od interpretacji giełdowych, ciekawymi estymatorami wydają się Indykator Aroona, Wstęga Bollingera, czy MACD - zwłaszcza te dwa ostatnie opierają się na podstawowych metodach stosowanych w ASC.

3 Dopasowywanie modelu autoregresji - teoria

Definicja 3.1. Niech $\{X_t\}$ - stacjonarny szereg czasowy rzędu II. Wówczas definiujemy funkcję autokorelacji cząstkowej jako

$$PACF(h) = \text{Corr}(X_{h+1} - \text{Proj}_{\text{lin}\{1, X_2, X_3, \dots, X_h\}} X_{h+1}, X_1 - \text{Proj}_{\text{lin}\{1, X_2, X_3, \dots, X_h\}} X_1)$$

Twierdzenie 3.1. Jeżeli X_t jest z modelu $AR(p)$, to $PACF(h) = 0$ dla $h > p$.

Reguła 1. Jeżeli próbkowa $PACF$ znajduje się pomiędzy $\pm\sqrt{n}$ dla $h > p$ to sugerowany jest model $AR(p)$.

Definicja 3.2.

3.1 Estymatory Yule'a-Walkera

Rozważmy model $AR(p)$:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + Z_t \quad (1)$$

Równaniami Yule-Walkera nazwiemy

$$\gamma_X(0) = \alpha_1 \gamma_X(1) + \alpha_2 \gamma_X(2) + \dots + \alpha_p \gamma_X(p) + \sigma_Z^2$$

$$\gamma_X(h) = \alpha_1 \gamma_X(h-1) + \alpha_2 \gamma_X(h-2) + \dots + \alpha_p \gamma_X(h-p) + \sigma_Z^2, \quad h = 1, 2, \dots, p$$

W postaci macierzowej:

$$\gamma_X(0) - \underline{\alpha}_p^T \gamma_{X,p} = \sigma_Z^2$$

$$\Gamma_{X,p} \underline{\alpha}_p = \underline{\gamma}_{X,p}$$

przy czym:

$$\Gamma_{X,p} = [\gamma_X(i-j)]_{i,j=1}^p$$

$$\underline{\alpha}_p = (\alpha_1, \alpha_2, \dots, \alpha_p)^T$$

$$\underline{\gamma}_{X,p} = (\gamma_X(1), \gamma_X(2), \dots, \gamma_X(p))^T$$

Podzielmy obie strony przez $\gamma_X(0)$:

$$\sigma_Z^2 = \gamma_X(0)(1 - \underline{\alpha}_p^T \underline{\rho}_{X,p})$$

$$R_{X,p} \underline{\alpha}_p = \underline{\rho}_{X,p}$$

gdzie:

$$R_{X,p} = [\rho_X(i-j)]_{i,j=1}^p$$

$$\underline{\rho}_{X,p} = (\rho_X(1), \rho_X(2), \dots, \rho_X(p))^T$$

Zastępując nieznane funkcje autokowariancji i autokorelacji procesu AR(p) estymatorami próbkowymi autokowariancji i autokorelacji otrzymujemy estymator wektora $\underline{\alpha}_p$ parametrów modelu AR(p) postaci:

$$\hat{\underline{\alpha}}_p = \hat{R}_{X,p}^{-1} \hat{\underline{\rho}}_{X,p}$$

oraz estymator wariancji białego szumu Z_t postaci:

$$\hat{\sigma}_Z^2 = \hat{\gamma}_X(0)(1 - \hat{\underline{\alpha}}_p^T \hat{\underline{\rho}}_{X,p})$$

3.2 FPE (Final Prediction Error)

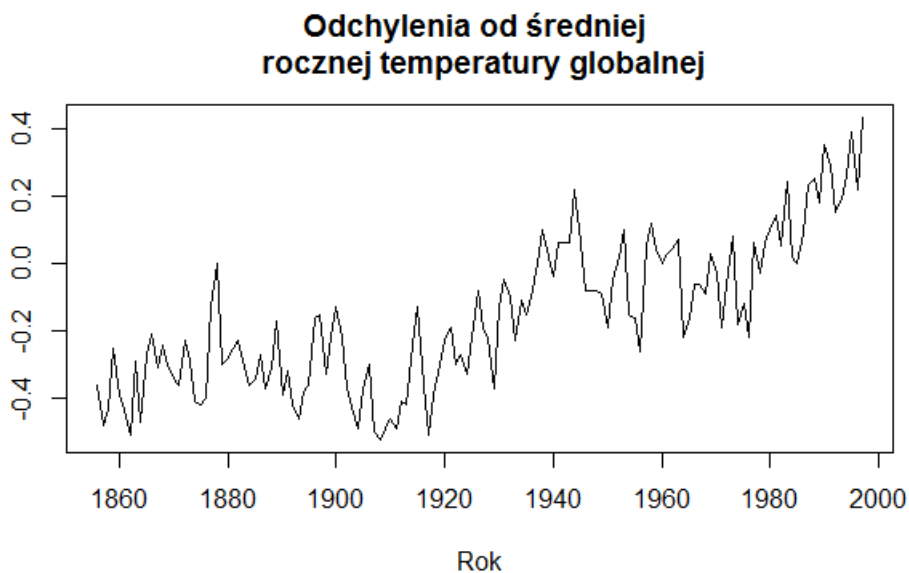
$$FPE(p) = \hat{\sigma}_n^2 \frac{n+p}{n-p}$$

Estymatorem \hat{p} rzędu modelu AR(p) jest p^* takie, że

$$FPE(p^*) = \min_p FPE(p)$$

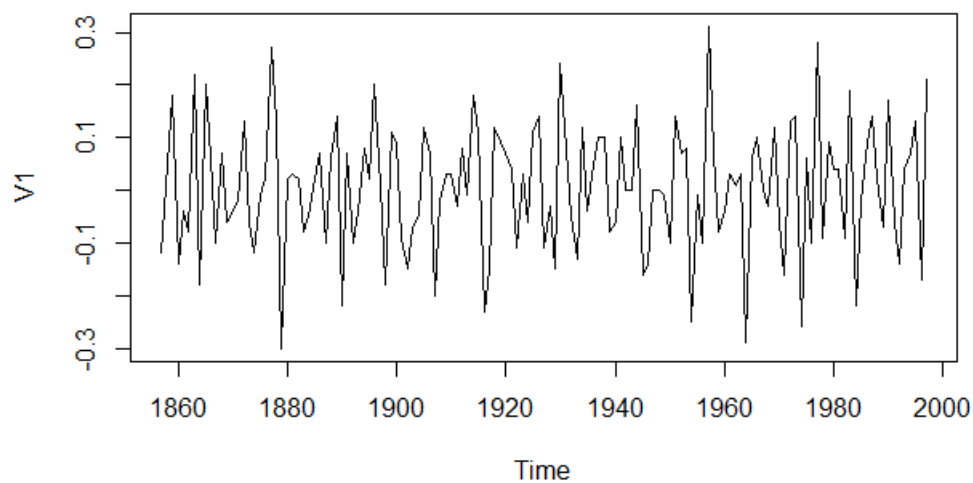
4 Dopasowanie modelu autoregresji - przykład

Dane pochodzą ze strony www.stat.pitt.edu/stoffer/tsa.html (dostęp 10 maja 2014), zawarte są w pliku globtemp.dat. Przedstawiają odchylenia średniej rocznej temperatury od temperatury średniej globalnej od roku 1856.



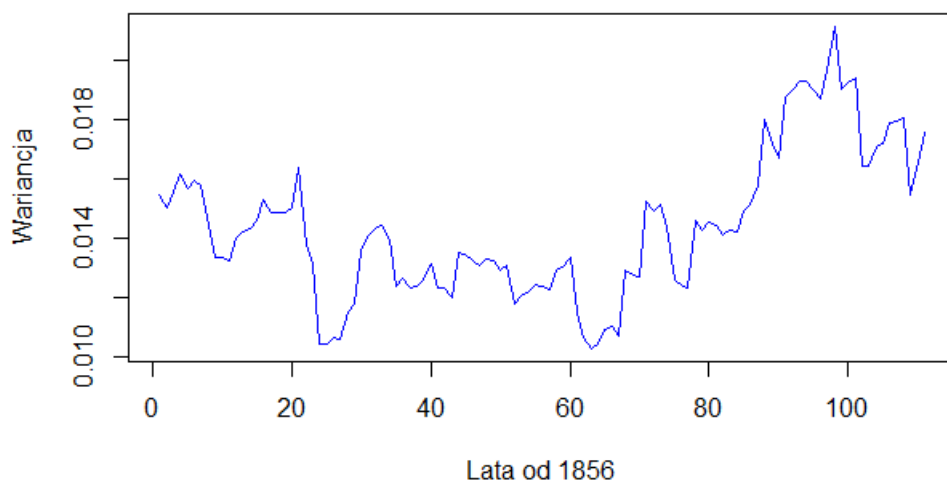
Wykres 17. Dane z pliku globtemp.dat

Patrząc na wykres, da się wyróżnić wyraźny trend. Z tego względu dokonamy jednokrotnego różnicowania z opóźnieniem 1.



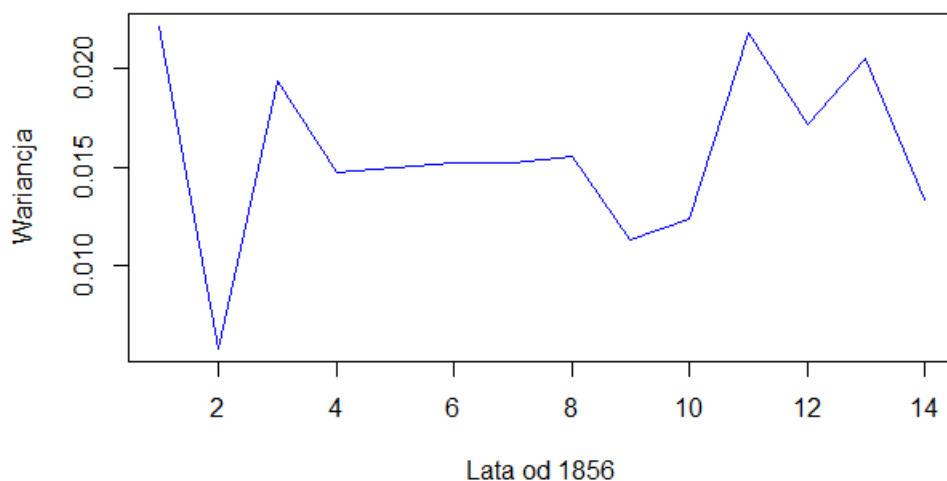
Wykres 18. Dane globtemp po jednokrotnym zróżnicowaniu z opóźnieniem 1.

Wariancja wygląda na jednorodną, ale sprawdźmy, jak kształtuje się ona w każdym trzydziesto-leciu (badamy wszystkie podprzedziały $[t, t + 30]$ i liczymy dla nich wariancję).



Wykres 19. Wariacje podprzedziałów trzydziestoletnich.

Różnice wydają się małe. Wariancja wszystkich otrzymanych wyników jest rzędu jednej milionowej. Ograniczmy się do podprzedziałów rozłącznych (dekady):

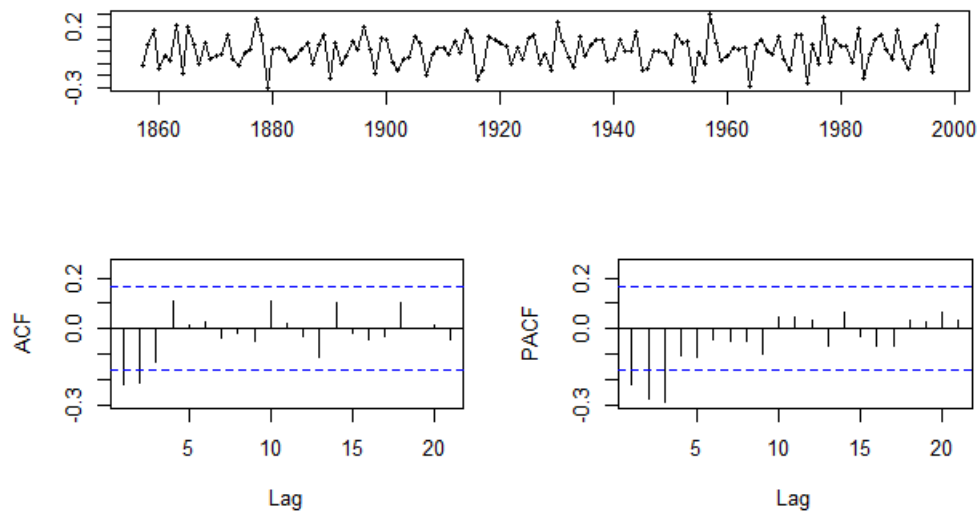


Wykres 20. Wariacje podprzedziałów rozłącznych, dekady.

Wariancja w tym przypadku jest na poziomie jednej stutysięcznej, a więc akceptowalna. Może się wydawać dziwne, że badamy wariację dla wariacji w podprzedziałach wariacji globalnej temperatury, ale w celu rzetelnego dopasowania modelu autoregresji trzeba upewnić się co do stacjonarności zróżnicowanego szeregu.

Dodatkowo wykonano na zróżnicowanym szeregu *globtemp* formalny test ADF. *p*-wartość okazała się mniejsza od 0,01, co pozwala przyjąć hipotezę o stacjonarności badanego szeregu.

4.1 Metoda PACF

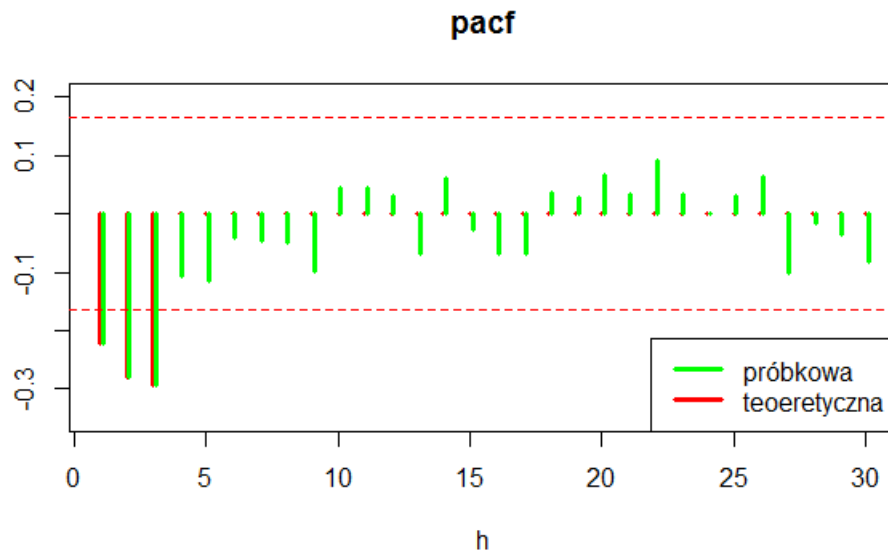


Wykresy 21-23. Wykresy uwzględniające ACF i PACF zróżnicowanego szeregu.

Powyższy wykres dodatkowo przekonuje nas o stacjonarności szeregu. Na podstawie wykresu PACF sugerowany model AR ma rząd 3. Spróbujmy dopasować model.

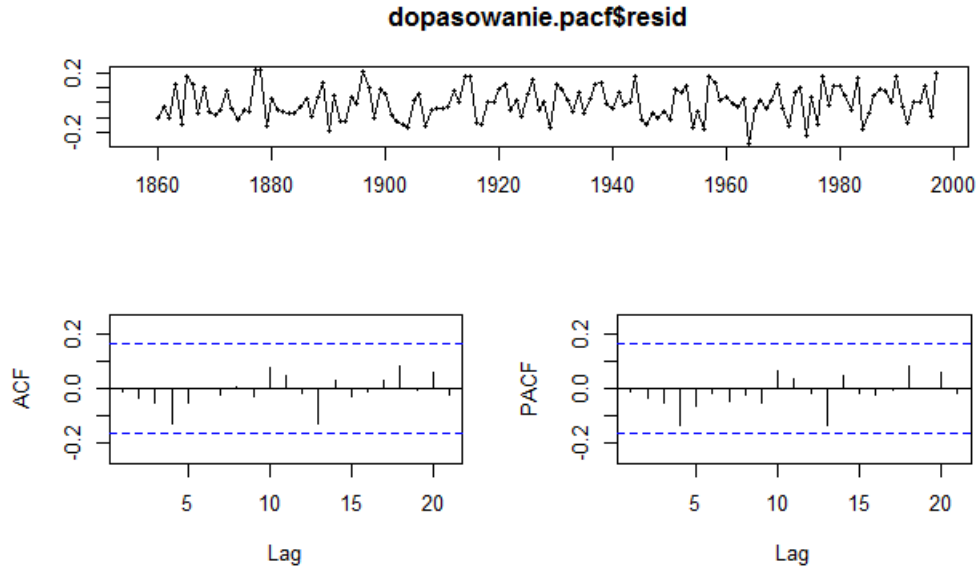
współczynnik	wartość
α_1	-0.3669430
α_2	-0.3638472
α_3	-0.2927388
σ^2	0.01201

Tablica 1: Estymowane wartości współczynników w modelu AR(3) metodą Yule-Walkera



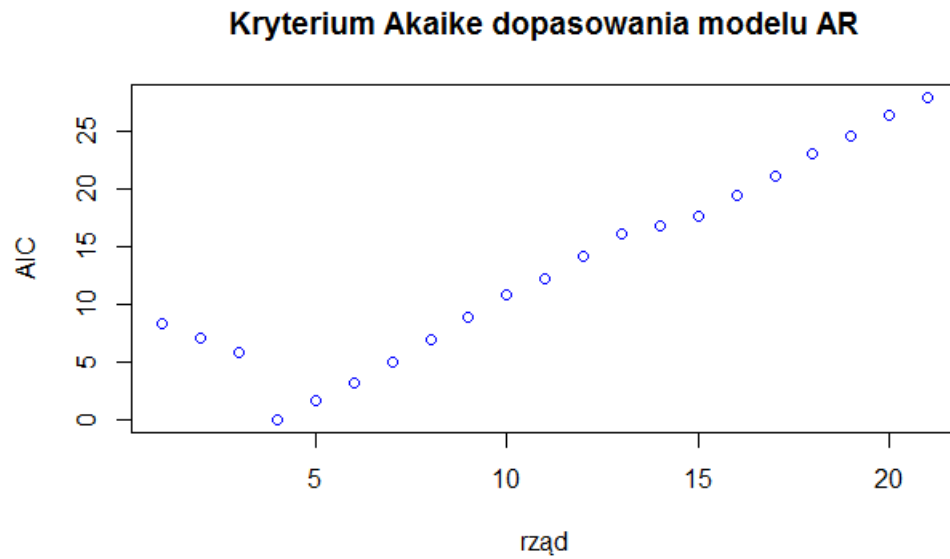
Wykres 24. Teoretyczna i próbkowa PACF dla modelu AR(3)

Test Box-Pierce'a wykazał (z p -wartością równą 0.7176) białoszumowość reszt. Dla testu Ljunga-Boxa p -value = 0.7146, a więc jest niewiele mniejsze. Graficzna metoda również nie dała żadnych niepokojących sygnałów:



Wykresy 25-27. Reszty po dopasowaniu modelu AR(3). Wykonano test istotności współczynników na poziomie istotności $\alpha = 0,001$. Wszystkie współczynniki okazały się istotne.

4.2 Metoda AIC



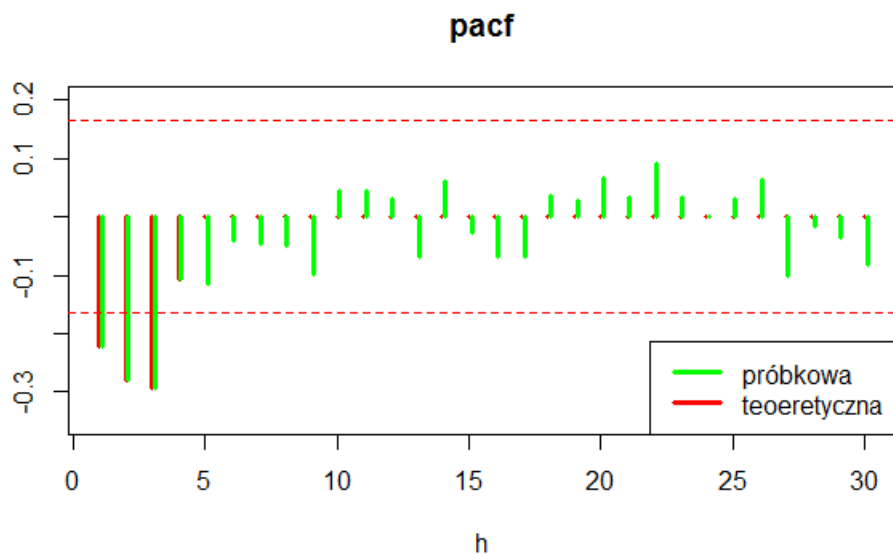
Wykres 28. Wartość funkcji AIC w zależności od rzędu p dla globtemp po zróżnicowaniu.

Funkcja AIC osiąga minimum dla $p = 4$. Po dopasowaniu modelu metodą Yule-Walkera, otrzymujemy następujące oszacowania parametrów modelu:

współczynnik	wartość
α_1	-0.3983924
α_2	-0.4029360
α_3	-0.3321601
α_4	-0.1074318
σ^2	0.01468

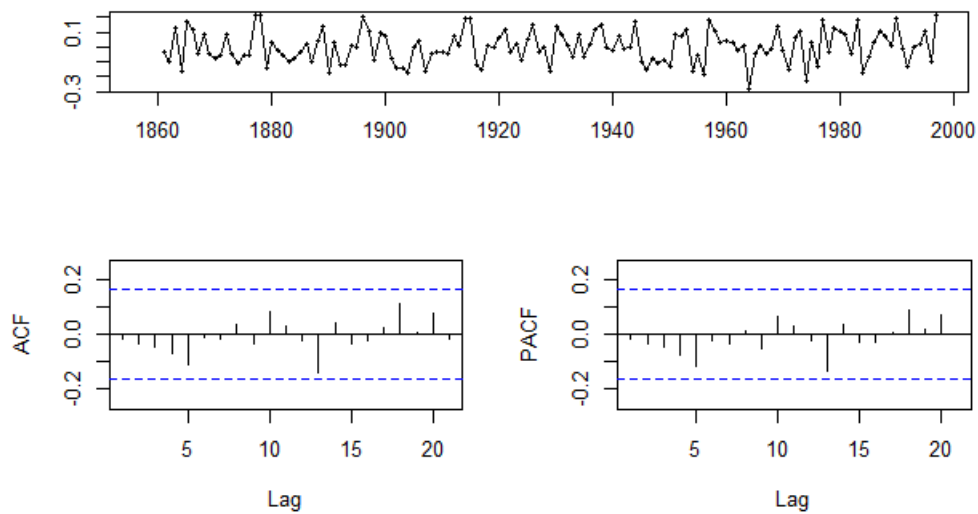
Tutaj również współczynniki podobne.

Tablica 2: Estymowane wartości współczynników w modelu AR(4) metodą Yule-Walkera



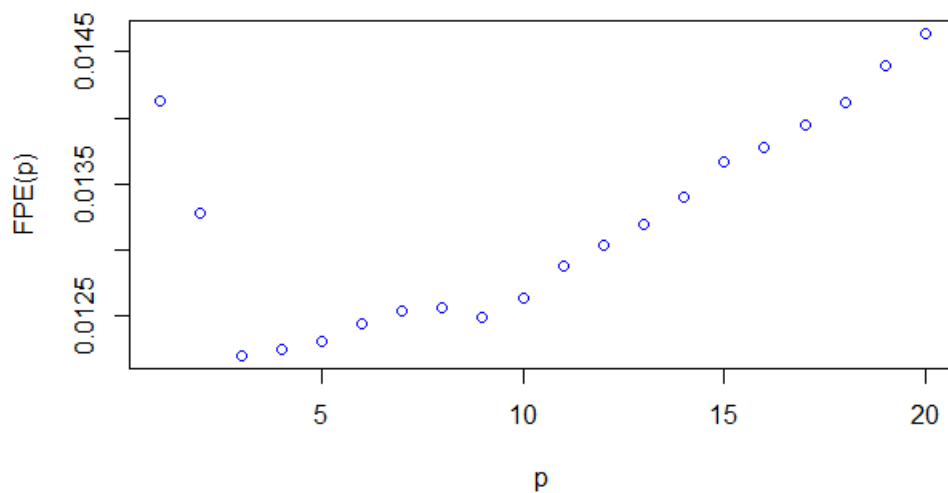
Wykres 29. Teoretyczna i próbkowa PACF dla modelu AR(4)

p -wartość w teście Boxa-Pierce'a białoszumowości reszt po dopasowaniu powyższego modelu przyjmuje wartość 0.7715, natomiast w teście Ljunga-Boxa - 0.8411.



Wykres 30. Reszty po dopasowaniu modelu AR(4) Wykonano test istotności współczynników na poziomie istotności $\alpha = 0,001$. Wszystkie współczynniki poza ostatnim okazały się istotne.

4.3 Metoda FPE



Wykres 31. Wartość funkcji FPE w zależności od rzędu p dla globtemp po zróżnicowaniu.

Kryterium FPE sugeruje model AR(3). Model ów dopasowywaliśmy wcześniej, wartości przeprowadzonych testów można więc znaleźć wyżej.

4.4 AR(3) i AR(4) ze współczynnikami wyestymowanymi innymi metodami

Dotychczas estymowaliśmy współczynniki metodą Yule-Walkera. Poniżej w tabelach podano wartości estymatorów największej wiarygodności. Wyniki bardzo podobne, jak przy estymatorach

współczynnik	wartość
α_1	-0.3720
α_2	-0.3638
α_3	-0.2999
σ^2	0.0116

Tablica 3: Estymowane wartości współczynników w modelu AR(3) metodą największej wiarygodności

Yule-Walkera.

współczynnik	wartość
α_1	-0.4061
α_2	-0.4035
α_3	-0.3384
α_4	-0.1102
σ^2	0.01146

Tablica 4: Estymowane wartości współczynników w modelu AR(4) metodą największej wiarygodności

4.5 Wnioski

Do badanego szeregu czasowego po jednokrotnym zróżnicowaniu z opóźnieniem $h = 1$ za pomocą trzech metod dopasowywano model autoregresji. Szereg ów (zróżnicowany) wydaje się stacjonarny, co daje nam mandat do podjęcia takich prób dopasowania. Rząd modelu wyznaczono na trzy sposoby: wizualnie za pomocą PACF, a także formalnie za pomocą FPE i AIC. Pierwsza oraz trzecia metoda sugerowały model AR(3), natomiast kryterium Akaike - AR(4). Współczynniki do modelu danego rzędu zostały wyestymowane metodą Yule-Walkera, a potem metodą największej wiarygodności. PACF dla modelu AR(4) wydaje się lepiej dopasowane od AR(3), choć na ich podstawie trudno orzec jednoznacznie o postaci modelu. Estymowana wariancja białego szumu zależała od zastosowanej metody estymacji współczynników - najmniejsza była dla modelu AR(4) z estymatorami największej wiarygodności, zaś największa - dla modelu AR(4) z estymatorami Yule-Walkera. ACF i PACF reszt było podobne zarówno dla modelu AR(3), jak i AR(4). p -value przy badaniu białoszumowości za pomocą testów Boxa-Pierce'a i Ljunga-Boxa reszt okazała się większa dla modelu AR(4), choć w obu modelach wartości te były duże. Wykonane testy istotności współczynników na poziomie istotności $\alpha = 0,001$ sugerują, że w modelu AR(4) ostatni współczynnik jest nieistotny. Testy na poziomie $\alpha = 0,0001$ wykluczają ostatni współczynnik w obydwu modelach, zaś test na poziomie $\alpha = 0,01$ również wyklucza ostatni współczynnik z modelu AR(4). Stąd model AR(3) wydaje się nieco lepiej dopasowany.