

**Mariusz KOZAKIEWICZ¹, Marek KWAS¹,
Karolina MUCHA-KUŚ², Maciej SOŁTYSIK²**

¹Szkoła Główna Handlowa, ²TAURON Polska Energia SA

ZASTOSOWANIE EKONOMETRYCZNYCH MODELI PROGNOSTYCZNYCH W TRANSAKCJACH PROPRIETARY TRADING

Uczestnictwo w strukturach wspólnotowych Unii Europejskiej nakłada na kraje stowarzyszone pewien wachlarz obowiązków, w szczególności konieczność zagwarantowania praw do swobody w przepływie towarów, przedsiębiorczości oraz świadczenia usług, które możliwe są do osiągnięcia jedynie przy otwartych rynkach. Daleko idąca liberalizacja rynków, mająca swoją podstawę w dyrektywach i rozporządzeniach wykonawczych, zapewnia wzrost bezpieczeństwa dostaw towarów i efektywności rynków implikując dążenie do ujednoczenia schematów działań i stworzenia jednego, wspólnotowego rynku. Taki ujednoczony model ściśle wpływa na parametryzację produktową towarów giełdowych ograniczając możliwości spekulacyjnej i arbitrażowej gry między rynkami.

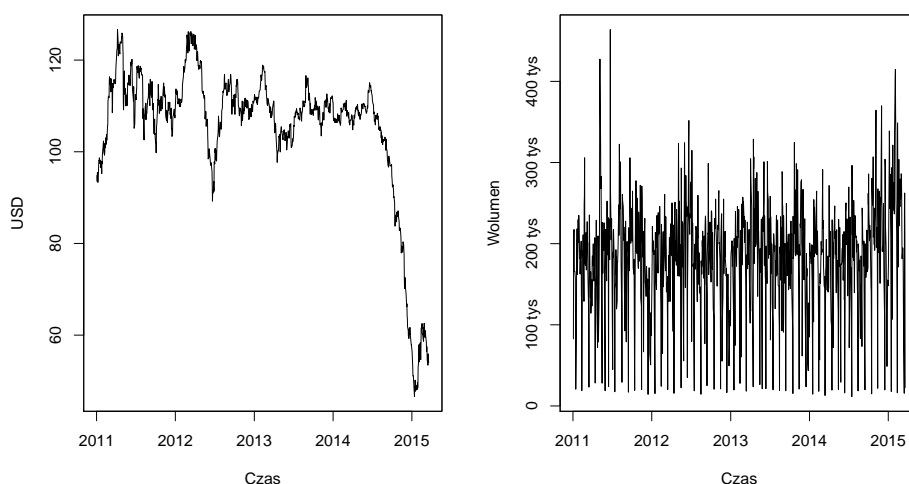
Likwidacja rynków w wymiarze lokalnym skutkująca wzrostem płynności z jednoczesnym ograniczeniem zmienności i stabilizacją poziomów cen na rynku globalnym, wymusza na graczach rynkowych poszukiwanie nowych obszarów aktywności. Działania te koncentrują się na poszukiwaniu nisz rynkowych i produktowych oraz ich asymilacji w ramach zupełnie nowych domen handlowych. Przykładowo, spółki obrotu dotychczas aktywne głównie na rynku energii elektrycznej, chcąc zapewnić istotną dla swoich akcjonariuszy stabilność wzrostu wartości, angażują się w obrót różnego rodzaju towarami pokrewnymi.

Komplementarność rynków energii elektrycznej i gazu oraz synergie występujące z produktami powiązаныmi jak ropa naftowa czy węgiel nadają atrakcyjności rozwojowi ukierunkowanemu w stronę rynków towarowych. Rynki tych produktów dzięki ich globalnemu charakterowi cechują się dużą płynnością, co powoduje że możliwa i biznesowo uzasadniona staje się próba skonstruowania modelu do prowadzenia gry spekulacyjnej. W referacie zaprezentowane są wyniki analiz modeli bazujących na notowaniach produktu ropa Brent w ramach rynków intraday i day ahead.

1. Motywacja biznesowa

Ropa Brent jest mieszaniną ropy wydobywanej z kilkunastu pól na Morzu Północnym. Do głównych determinant wpływających na poziom jej cen należą

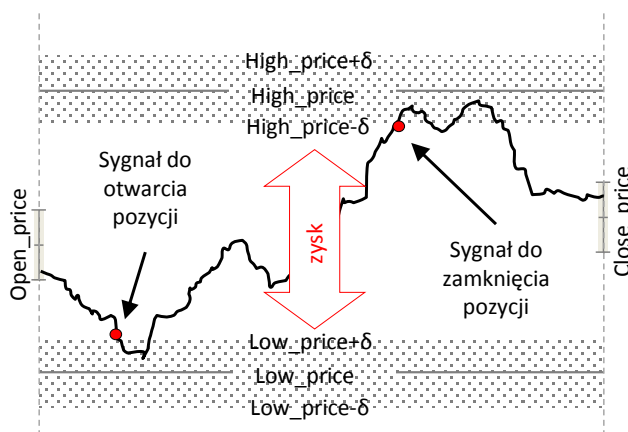
zaliczyć przede wszystkim relacje podaży-popytowe, wahania kursu dolara, sezonowość oraz sytuację geopolityczną. Najważniejszymi instytucjami odpowiedzialnymi za globalny obrót ropą są giełdy New York Mercantile Exchange oraz Intercontinental Exchange, której domeną jest ropa Brent. Obrót ropą odbywa się zarówno w ramach kontraktów terminowych z fizyczną dostawą, jak również w oparciu o instrumenty finansowe na rynku futures. Transakcje realizowane są całodobowo, w ramach rynku OTC oraz na platformach obrotu zorganizowanych przy udziale giełdowych grup kapitałowych.



Rys. 1. Przebieg notowań ropy Brent w okresie 11-04-2011 – 12-03-2015.

Analiza notowań historycznych ropy Brent dla lat 2011 – 2014, por. rys. 1, wskazuje na występowanie sezonowości, trendów oraz statystycznie istotnej zmienności notowań zarówno w ramach doby, jak i w okresach dłuższych. Dodatkowo obserwuje się duży wolumen obrotu, dzięki czemu na rynek ten mniejsze przełożenie ma handel spekulacyjny i arbitraż między rynkami. Potencjalnie możliwe staje się zatem skonstruowanie modelu, który w oparciu o analizę zależności fundamentalnych między determinantami, jak również w oparciu o statystyczną analizę szeregów czasowych notowań, pozwoli w sposób automatyczny na badanie zależności między zmiennymi cenowymi i ich predykcjami oraz wskaże rekomendowane kierunki działań handlowych.

Przebiegi notowań ciągłych towarów giełdowych sparametryzowane są co najmniej czterema wartościami cen: ceną otwarcia (*open price*), ceną zamknięcia (*close price*), ceną minimalną (*low price*) oraz ceną maksymalną (*high price*). Proponowany model bazuje na prognozach tych cen, w oparciu o które, w sposób

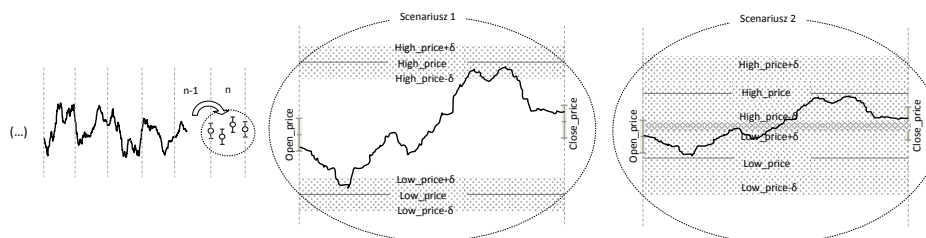


Rys. 2. Mechanizm otwierania i zamykania pozycji.

automatycznie generowane są sygnały do otwarcia i zamknięcia pozycji maksymalizujące spread między nimi, co w sposób schematyczny ilustruje rys. 2.

Proponowany mechanizm zakłada otwarcie pozycji w momencie przecięcia się bieżących notowań produktu z linią prognozowanej ceny *low price*, dodatkowo powiększonej o wartość niepewności wynikającej z analiz ex-post. Sygnał do zamknięcia pozycji jest tożsamy z momentem osiągnięcia przez bieżący poziom notowanych cen, prognozowanej ceny maksymalnej dobowej *high price* pomniejszonej o wartość niepewności ex-post. Model celowo uwzględnia historyczne poziomy niepewności zawężając przestrzeń do realizacji wyniku, co jednocześnie przekłada się na minimalizację ryzyka wygenerowania straty. Analiza danych historycznych w konfrontacji z danymi z predykcji prowadzi jednak do wniosku, że algorytm postępowania w zaprezentowanej formie nie przynosi zakładanych efektów. Głównymi przyczynami takiego stanu rzeczy są: (i) niewielka amplituda zmian w ramach notowań dobowych, (ii) wielomodalność notowań, (iii) zmienny poziom niepewności ex-post, (iv) pokrywanie się obszarów niepewności dla zmiennych *low price* i *high price*, (v) brak uwzględniania w prognozach cen sygnałów z otwarcia notowań w dobie realizacyjnej. Różnorodność możliwych zachowań rynku i notowań ilustrują dwa skrajne scenariusze przedstawione na rys. 3.

Ze względu na fundamentalne trudności w predykcji cen *low price* i *high price* oraz momentów ich osiągnięcia w trakcie trwania notowań, modelowanie oparto na cenach *open price* i *close price*. Charakter tych cen jest stabilniejszy niż zmiennych i silnie fluktuujących notowań w ramach doby. Te odniesienia cenowe stanowią zatem pewien element analizy fundamentalnej i są punktem odniesienia



Rys. 3. Scenariusze kształtowania się obszaru zysku/straty w ramach notowań dobowych.

w składanych przez graczy rynkowych zleceniach. Występujące różnice w cenach zamknięcia w dobie n i otwarcia w dobie $n + 1$, a także różnice między prognozowanym poziomem ceny otwarcia, a jej rzeczywistą wartością, są pochodnymi nastrojów rynkowych i zmiennych zachowań graczy i powinny być uwzględniane w modelu. Ponadto, decyzje transakcyjne oparte na predykcji wartości *open price* i *close price* mają dobrze zdefiniowane momenty otwarcia/zamknięcia pozycji, przez co unikamy ich predykcji, obciążonej zwykle dużym błędem.

2. Algorytm

Rozważmy transakcje na instrumencie BRENT polegające na otwarciu i zamknięciu pozycji w danym dniu po cenach *open price* i *close price*. Oczywiście transakcja taka przynosi zysk wyłącznie wtedy, gdy różnica między *close price* a *open price* jest dodatnia. Algorytm zaprezentowany w tym rozdziale generuje rekomendacje realizacji takich transakcji przez odpowiednią prognozę wspomnianej różnicy cen. Istotnym elementem algorytmu jest predykcja cen instrumentu BRENT wykorzystująca historyczne wartości kursu EUR/USD. Wyniki predykcji są użyte do budowy decyzyjnego drzewa regresyjnego, które generuje ostateczne rekomendacje. Zastosowanie tej techniki umożliwiło istotne zmniejszenie strat w okresach spadku cen.

2.1. Oznaczenia

W dalszej części referatu zostały przyjęte następujące oznaczenia i założenia.

- Przedział czasowy, na którym operuje algorytm składa się z dni transakcyjnych (bez uwzględnienia weekendów i świąt), oznaczanych liczbami naturalnymi. Jeżeli nie jest określone inaczej, dzień bieżący oznaczamy przez n . Jest to dzień, do którego włącznie znamy historyczne wartości cen. Dzień $n + 1$ rozpoczyna horyzont prognozy.
- Wartości szeregów czasowych dla instrumentu BRENT w dniu k oznaczamy

przez

$$\text{BRENT}_k^O, \text{BRENT}_k^L, \text{BRENT}_k^H, \text{BRENT}_k^C,$$

gdzie indeksy O, L, H, C oznaczają odpowiednio ceny *open price*, *low price*, *high price*, *close price*.

- Wartości szeregów czasowych dla kursu EUR/USD w dniu k oznaczamy przez

$$\text{EUR/USD}_k^O, \text{EUR/USD}_k^L, \text{EUR/USD}_k^H, \text{EUR/USD}_k^C,$$

gdzie, jak wyżej, indeksy O, L, H, C oznaczają odpowiednio poziomy *open*, *low*, *high*, *close*.

- Wartości prognozowane oznaczamy przez “ $\widehat{}$ ”.

2.2. Struktura algorytmu

Algorytm, por. tab. 1, generuje rekomendację otwarcia/zamknięcia pozycji w dniu $n + 2$ na podstawie cen historycznych do dnia n i, opartej na nich, krótkoterminowej prognozy różnicy cen *close price* - *open price* $\text{BRENT}_{n+2}^C - \text{BRENT}_{n+2}^O$. Do wyznaczenia tej prognozy zostały wykorzystane historyczne wartości kursu EUR/USD, por. roz. 3. Ponadto, alternatywna bardziej rozbudowana wersja algorytmu, por. tab. 1, krok (b*), używa dodatkowo modeli ARIMA, por. [2], do poprawy wspomnianej prognozy.

Istotnym elementem algorytmu jest użycie decyzyjnego drzewa regresyjnego do wygenerowania ostatecznych rekomendacji. Zmienne objaśniające, na podstawie których jest budowane drzewo zostały dobrane tak aby wyczerpująco scharakteryzować lokalną monotoniczność, wklęsłość oraz zmienność wartości BRENT_n oraz predykcji $\widehat{\text{BRENT}}_{n+1}$, $\widehat{\text{BRENT}}_{n+2}$, $\widehat{\text{BRENT}}_{n+3}$. Dodatkowo, testy empiryczne wykazały istotną zależność wyników strategii od typów dnia tygodnia i dnia miesiąca.

Tabela 1

Algorytm generujący rekomendację dla otwarcia/zamknięcia pozycji w dniu $n + 2$.

<p>1. Prognozowanie $BRENT^X$ z wykorzystaniem EUR/USD^X, $X \in \{O, L, H, C\}$.</p> <p>(a) Dla każdego wskaźnika $X \in \{O, L, H, C\}$ tworzony jest niezależny model regresji liniowej LM^X</p> $\Delta BRENT^X \sim \Delta EUR/USD^X$ <p>estymowany na wartościach różnic $\Delta BRENT^X$ z dni $n - 5, \dots, n$ i odpowiednich, przesuniętych o 10 dni transakcyjnych wstecz, wartościach różnic $\Delta EUR/USD^X$.</p> <p>(b) Dla każdego wskaźnika $X \in \{O, L, H, C\}$, z modelu LM_n^X, wyznaczane są prognozy $\widehat{BRENT}_{n+1}^X, \widehat{BRENT}_{n+2}^X, \widehat{BRENT}_{n+3}^X$.</p> <p>(b*) Dla każdego wskaźnika $X \in \{O, L, H, C\}$ do reszt regresji modelu LM_n^X dopasowany zostaje model ARIMA, który wraz z modelem LM_n^X został wykorzystany do wyznaczenia prognoz $\widehat{BRENT}_{n+1}^X, \widehat{BRENT}_{n+2}^X, \widehat{BRENT}_{n+3}^X$.</p> <p>2. Generowanie rekomendacji otwarcia/zamknięcia pozycji w dniu $n + 2$.</p> <p>(a) Skonstruowanie drzewa regresyjnego $TREE_n$ na danych z okresu $n - 35, \dots, n - 1$ dla</p> <ul style="list-style-type: none"> • zmiennej objaśnianej Y_k przyjmującej wartość 1 gdy zachodzi $BRENT_{k+2}^C - BRENT_{k+2}^O > 0$ oraz -1 w przeciwnym przypadku, • zmiennych objaśniających będących funkcjami historycznych i prognozowanych wartości $BRENT$ oraz typów dni tygodnia i miesiąca. <p>(b) Wygenerowanie wartości zmiennych objaśniających dla dnia n, a następnie, przy użyciu drzewa $TREE_n$, predykcji \hat{Y}_n. Wartość 1 oznacza pozytywną rekomendację dla otwarcia/zamknięcia pozycji w dniu $n + 2$, wartość -1 – rekomendację negatywną.</p>

3. Związek między BRENT a EUR/USD

Algorytm z tab. 1 wykorzystuje krótkoterminową prognozę cen instrumentu BRENT do generowania rekomendacji. W prognozie tej wykorzystany został kurs EUR/USD, wybrany z szerokiego katalogu dostępnych instrumentów. Charakteryzuje się on wysoką płynnością oraz dużym wolumenem transakcji, ponadto wartości EUR/USD są mocno skorelowane z wartościami BRENT. Dla ustalenia uwagi, zostały zaprezentowane statystyki absolutnych błędów względnych (absolute percentage error – APE) prognoz dla cen *close price* z doby $n + 3$, por. tab. 2. Wyniki dla cen *open price*, *high price* i *low price* są porównywalne. Podkreślimy, że przedstawione są tutaj błędy prognoz out-of-sample, czyli prognozy zostały

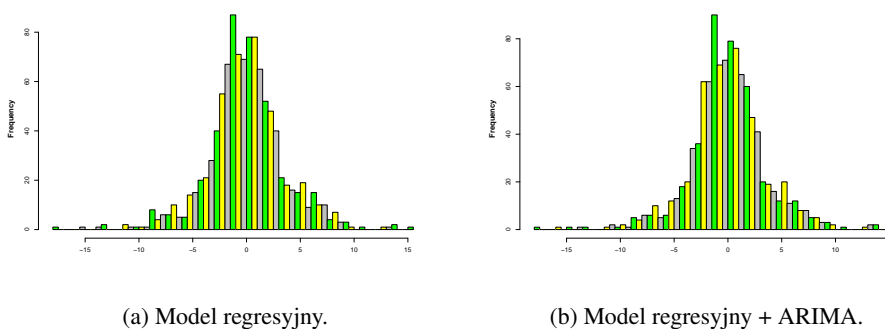
wyznaczone dla wartości przyszłych względem danych, na których estymowano model a procedura została wykonana krocząco dla danych z okresu testowego 11-04-2011 – 12-03-2015.

Tabela 2

Wartości APE (w procentach) dla prognoz $BRENT_{n+3}^C$ z modelu regresyjnego, z dodatkowym użyciem modeli ARIMA. Okres testowy 11-04-2011 – 09-12-2014.

	min	1. kw.	med.	MAPE	3. kw.	max
mod. reg.	0,0015	0,8172	1,7120	2,6160	3,3830	36,6800
mod. reg. + ARIMA	0,0015	0,8190	1,6950	2,5690	3,2840	36,6800

Testy empiryczne pokazują, że stosunkowo proste modele regresyjne oparte na wartościach instrumentu EUR/USD z odpowiednim przesunięciem dostarczają prognoz dobrej jakości. Ponadto, dodatkowe użycie modeli ARIMA wydaje się nie wpływać istotnie na poprawę jakości prognozy, por. średnie absolutne błędy względne (MAPE) w tab. 2 oraz rys. 4. Okazuje się jednak, że występuje istotne polepszenie rekomendacji generowanych przez drzewo decyzyjne, por. roz. 4. Mimo, iż histogramy różnic $BRENT^C - \widehat{BRENT}^C$, por. rys. 4, sugerują normalność rozkładów, standardowe testy nakazują taką hipotezę odrzucić. Właściwe dopasowania rozkładów błędów jest kluczowe dla kwestii oszacowania ryzyka przyjętej strategii i będzie przedmiotem dalszych badań.



Rys. 4. Histogramy różnic $BRENT_{n+3}^C - \widehat{BRENT}_{n+3}^C$.

4. Analiza wyników

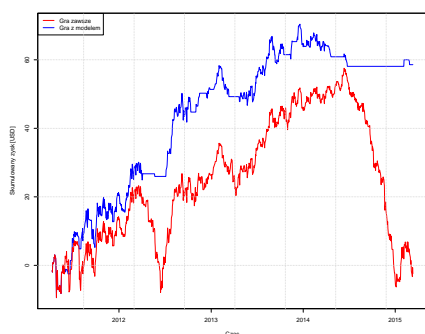
W celu uproszczenia omówienia analizy wyników symulacji zostaną przyjęte następujące skrócone oznaczenia

- AL10 to strategia oparta na rekomendacjach algorytmu z tab. 1 używającego predykcji dla 10-cio dniowego opóźnienia jak w kroku (b).

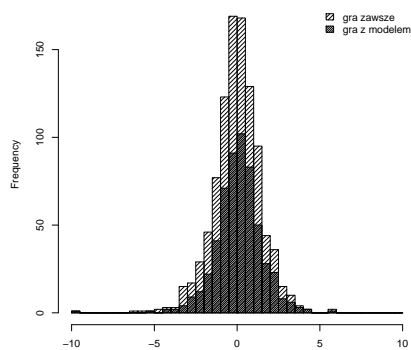
- AL10A to strategia oparta na rekomendacjach algorytmu z tab. 1 używającego predykcji dla 10-cio dniowego opóźnienia z dodatkowym modelem ARIMA dla błędu opisanej w kroku (b*).

Naturalnym odniesieniem dla strategii opartych na rekomendacjach jest strategia “graj zawsze”, w której otwieramy pozycję po cenie *open price* i zamykamy po cenie *close price*. Rys. 5a przedstawia porównanie skumulowanych zysków/strat strategii “graj zawsze” z AL10. Symulacja została przeprowadzona dla okresu 11-04-2011 – 12-03-2015.

Podkreślimy, że uproszczone transakcje polegające na zakupie i sprzedaży jednostki instrumentu BRENT w praktyce nie miałyby zastosowania, jednak są wystarczające do porównania wyników różnych strategii opartych na rekomendacjach. W celu urealnienia scenariusza należałoby prześledzić ewolucję portfela o ustalonym kapitale początkowym dla rzeczywistych parametrów rynkowych, przede wszystkim stóp procentowych, z możliwością odpowiedniej alokacji środków między różne instrumenty, np. BRENT i USD.



(a) Profile zysków/strat.

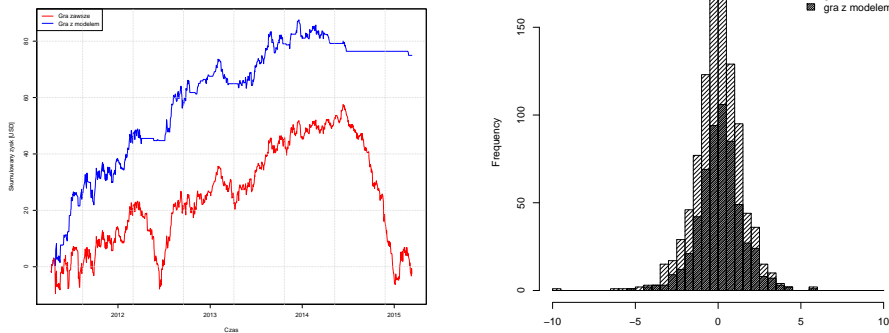


(b) Histogramy zysków/strat (usunięto zera).

Rys. 5. Porównanie strategii AL10 ze strategią “graj zawsze”.

Porównanie zachowania obu strategii pokazuje, że rekomendacje AL10 w początkowym okresie przegrywają z strategią “graj zawsze”, dość szybko jednak ujawniają swoją przewagę, gdyż pozwalają uniknąć transakcji w pewnych okresach spadków cen. Algorytm charakteryzuje się jednak długą pamięcią i okresy, w których rekomendowane jest wstrzymanie transakcji, przedłużają się na czas odbudowy rynku po spadkach, co znacznie ogranicza potencjalne zyski.

Interesujące jest porównanie wykresu z rys. 5a z rys. 6a, na których przedstawiono profile zysków/strat dla AL10 i AL10A. Profil z rys. 6a wykazuje widoczną poprawę spowodowaną najprawdopodobniej pewnym wygładzeniem błędów pre-



(a) Profile zysków/strat.

(b) Histogramy zysków/strat (usunięto zera).

Rys. 6. Porównanie strategii AL10A ze strategią “graj zawsze”.

dykcji przez dodatkowy krok modelowania. Efekt ten sugeruje, że na profil zysków/strat ma wpływ nie tylko jakość prognozy (MAPE), por. roz. 3, ale również charakter rozkładu błędów.

Podobne wnioski można wyciągnąć analizując statystyki oraz histogramy zysków/strat dla omawianych strategii. Tab. 3 i 4 uwidaczniają przewagę strategii AL10A nie tylko pod względem średniej zysków/strat ale przede wszystkim pod względem odchylenia standardowego, co ściśle wiąże się z ryzykiem strategii. Ograniczenie zmienności strategii wydaje się zatem równie istotne jak optymalizacja doboru reguł decyzyjnych.

Tabela 3

Liczby trafionych i chybionych rekomendacji oraz odpowiadające im skumulowane zyski/straty dla strategii AL10, AL5 oraz AL10A w porównaniu ze strategią “graj zawsze”.

	gra	rekomendacje			zyski/straty		
		zysk	strata	suma	zysk	strata	suma
AL10	graj	306	256	562	315.09	-256.51	58.58
	nie graj	199	232	431	206.86	-267.85	-60.99
AL10A	graj	312	256	568	320.52	-245.56	74.96
	nie graj	193	232	425	201.43	-278.80	-77.37
“graj zawsze”		505	488	993	521.95	-524.36	-2.41

Tabela 4

Średnie oraz odchylenia standardowe zysków/strat dla strategii “graj zawsze” oraz strategii opartych na rekomendacjach AL10, AL5 oraz AL10A.

	“graj zawsze”	AL10	AL10A
średnia	-0.0024	0.0589	0.0752
odch. std.	1.4263	1.0478	1.0070

Analiza histogramów z rys. 5b i 6b pokazuje w jaki sposób zostały osiągnięte wyżej opisane efekty. Podkreślmy, że wszystkie wykresy zawierają wyłącznie niezerowe wartości. Widoczna jest eliminacja dużej części zarówno zysków jak i strat, jednak istotnie większa w przypadku strat, co spowodowało przesunięcie rozkładów na prawo. Ponadto, ujawnia się proporcjonalnie większy ubytek dużych wartości (szczególnie strat), przez co istotnie zmniejszone zostało odchylenie standardowe.

5. Podsumowanie

W opracowaniu zaprezentowano skuteczny algorytm generowania rekomendacji transakcyjnych na rynku ropy Brent. Algorytm wykorzystuje stosunkowo proste modele ekonometryczne do krótkoterminowej predykcji cen oraz regresyjne drzewa decyzyjne do generowania rekomendacji. Analiza profili zysków/strat dla wybranych wersji algorytmu pokazuje na dużą zależność jego skuteczności od zastosowanych modeli predykcyjnych, w tym ich struktury.

Wyniki zaprezentowane w opracowaniu dają podstawę do kontynuacji badań w tej tematyce. Istotna wydaje się analiza szerszego spektrum modeli i technik predykcyjnych i zidentyfikowanie najważniejszych, dostarczających prognoz nie tylko wysokiej jakości ale również stanowiących podstawę do wygenerowania skutecznych strategii transakcyjnych. Poprawę prognoz można również uzyskać wykorzystując odpowiednio wybrane inne instrumenty rynku energii, w tym instrumenty pochodne. Wnikliwej analizie należy również poddać proces generowania drzew regresyjnych, w szczególności dobór zmiennych objaśniających, zmieniając w razie potrzeby ich katalog.

Do implementacji i przeprowadzenia obliczeń, których wyniki zawiera opracowanie wykorzystano system obliczeń statystycznych R, por. [3], w szczególności pakiety zoo, por. [5]; forecast, por. [1] i rpart, por. [4].

Literatura

- [1] R. J. Hyndman. *forecast: Forecasting functions for time series*, 2011. R package version 3.11.
- [2] G. S. Maddala. *Ekonometria*. PWN, 2006.
- [3] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2011. ISBN 3-900051-07-0.
- [4] Terry Therneau, Beth Atkinson, Brian Ripley. *rpart: Recursive Partitioning and Regression Trees*, 2014. R package version 4.1-8.
- [5] Achim Zeileis, Gabor Grothendieck. zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, 14(6):1–27, 2005.

ZASTOSOWANIE EKONOMETRYCZNYCH MODELI PROGNOSTYCZNYCH W TRANSAKCJACH PROPRIETARY TRADING

W opracowaniu przedstawiony jest algorytm służący do wsparcia decyzji transakcyjnych na rynku ropy Brent. Rynek ropy (w szczególności Brent) charakteryzuje się dużą płynnością oraz wolumenem transakcji i stanowi dogodną platformę zarówno do bezpośredniego inwestowania czy spekulacji, jak i do zarządzania ryzykiem dla podmiotów narażonych na nie ze względu na zależność od rynków energii.

Opisywany algorytm wykorzystuje elementy modelowania ekonometrycznego do krótkoterminowej predykcji cen ropy BRENT oraz technikę regresyjnych drzew decyzyjnych do generowania rekomendacji otwarcia/zamknięcia pozycji. W opracowaniu zawarto porównanie jakości prognozy wybranych modeli predykcyjnych oraz profile zysków/strat dla rekomendacji opartych na tych modelach.

Wyniki zawarte w opracowaniu sugerują, że przyjęte podejście stanowi obiecujący kierunek badawczy z uwagi na jego elastyczność umożliwiającą wykorzystanie różnych technik modelowania, jak również możliwość zaadaptowania do innych instrumentów finansowych.

AN APPLICATION OF ECONOMETRIC PREDICTIVE MODELS TO PROPRIETARY TRADING

The paper presents an algorithm supporting trading decisions on Brent oil market. The oil market, due to its liquidity and the high volume of trading, is a convenient platform not only for direct investment or speculation but also for risk management for entities dependent on volatile energy markets.

The algorithm uses econometric models for short term prediction of Brent oil prices which is subsequently used to construct regression decision trees for generating trading recommendations. The paper contains a comparison of selected prediction techniques and their recommendation profit/loss profiles.

The results suggest that the adopted approach is a promising research direction. Moreover, its flexibility allows applying various modeling techniques and adapting to the wide range of financial instruments.