

ZASTOSOWANIE SZTUCZNYCH SIECI NEURONOWYCH DO PROGNOZOWANIA CEN NA GIEŁDZIE ENERGII

Adrian MILLER¹, Paweł BUĆKO²

1. Politechnika Gdańska, ul. G. Narutowicza 11/12, 80-952 Gdańsk
e-mail: amiller@ely.pg.gda.pl
2. Politechnika Gdańska, ul. G. Narutowicza 11/12, 80-952 Gdańsk
e-mail: p.bucko@ely.pg.gda.pl

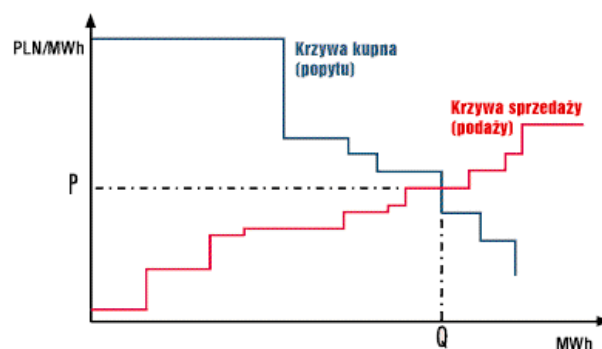
Streszczenie: Jednym z aktualnych zagadnień gospodarczych jest prognozowanie cen na Giełdzie Energii. Autorzy publikacji proponują wykorzystanie do tego celu sztucznych sieci neuronowych i oprogramowania MATLAB. Opisują narzędzie do prognozy krótkoterminowej Rynku Dnia Następnego. W artykule przedstawiono sposób kalibracji danych. Opisano również użyte funkcje aktywacji w warstwach neuronowych wraz z konsekwencjami z nich wynikającymi dla samego procesu nauczania. W końcowej części artykułu porównano prognozę sieci neuronowej z rzeczywistym przebiegiem notowań na Giełdzie Energii.

Słowa kluczowe: prognoza, sztuczne sieci neuronowe, rynek energii, MATLAB.

1. WSTĘP

Głównym celem analiz było stworzenie skutecznego narzędzia do krótkoterminowej godzinowej prognozy ceny na Giełdzie Energii. Zdecydowano się, że w tym celu wykorzystane zostaną sztuczne sieci neuronowe. Za narzędzie wspomagające kreowanie sieci, wybrano pakiet Neural Network ToolboxTM programu MATLAB 2008a.

Jednym z podstawowych obszarów działania Giełdy Energii jest Rynek Dnia Następnego (RDN). Jak sama nazwa wskazuje, notowania na RDN mają miejsce na dzień przed fizyczną dostawą energii. Notowania prowadzone są codziennie przez cały rok. Na Rynek Dnia Następnego składają się 24 godzinowe okresy rozliczeniowe, w których zawierane są transakcje kupna i sprzedaży energii elektrycznej. Polecenia kupna i sprzedaży są wysyłane przez uczestników RDN, następnie dla każdej godziny doby handlowej ze wszystkich dyspozycji zakupu powstaje krzywa popytu, natomiast z dyspozycji sprzedaży powstaje krzywa podaży [8]. Punkt przecięcia krzywych podaży oraz popytu jest podstawą do wyznaczenia ceny (kursu jednolitego). Przy wyznaczeniu wcześniej wspomnianej ceny stosuje się kolejno dwie główne zasady, tj. „maksymalizacji wolumenu obrotu oraz minimalizacji różnicy między skumulowanym wolumenem energii elektrycznej w zleceniach kupna i zleceniach sprzedaży możliwym do zrealizowania po określonym kursie”. Ogólny sposób ustalania ceny równowagi oraz wolumenu obrotu przedstawia poniższy rys. 1.



Rys. 1. Zasada wyznaczania ceny równowagi [4]

2. SKALOWANIE DANYCH

By przystąpić do pracy ze sztucznymi sieciami neuronowymi uprzednio należało zgromadzić odpowiednio liczny zbiór danych opisujących przeszłość procesu zmienności ceny na RDN. Minimalna, potrzebna ilość danych, jaką się przyjmuje do prawidłowego nauczania sieci, to dane z 12 miesięcy [1]. Zbiór danych powinien uwzględnić cykliczne zmienności zapotrzebowania w systemie (sezonowość procesu). Okres ten autorzy uznali jako ograniczenie przy tworzeniu modelu prognostycznego. Do badań wykorzystano zbiór danych z 2010 roku zaczerpnięty ze strony internetowej www.tge.pl [4].

Tabela 1. Dane charakteryzujące cenę energii w 2010 roku

| WIELKOŚĆ | WARTOŚĆ [PLN/MWh] |
|---|-------------------|
| Minimalna godzinowa cena energii | 101,15 |
| Średnia godzinowa cena energii | 191,78 |
| Maksymalna godzinowa cena energii | 499,86 |
| Minimalna dobową średnia cena energii | 146,08 |
| Maksymalna dobową średnia cena energii | 302,12 |
| Minimalna rozpiętość dobową ceny energii | 28,63 |
| Średnia rozpiętość dobową ceny energii | 92,78 |
| Maksymalna rozpiętość dobową ceny energii | 357,26 |

Dane podzielono na zbiór uczący (treningowy) i testujący. Do zbioru testującego wybrano co 7 tygodni roku, tworząc zbiór charakterystyczny dla różnych sezonów. Zaproponowany podział pozwalał na sprawdzenie początkowych i końcowych wyników uczenia sieci. Dodatkową zaletą takiego rozwiązania było również sprawdzanie wyników dla nietypowych dni w ciągu roku, takich jak dni przed- i poświąteczne.

Przygotowanie danych treningowych, polegające na ich odpowiednim przeskalowaniu, może poprawić zarówno działanie sieci, jak i jej dokładność [2, 3]. Podstawowym etapem było sprowadzenie danych do przedziału od 0 do 1. Dokonano tego stosując proste skalowanie:

$$A = \frac{B}{B_{\max}} \quad (1)$$

gdzie: A – wartość wejściowa przeskalowana do przedziału (0,1), B – wartość rzeczywista ceny [zł/MWh], B_{\max} – maksymalna wartość ceny w zbiorze danych [zł/MWh].

By dokonać przeliczenia wartości znormalizowanych na rzeczywiste należało skorzystać z zależności:

$$B = A \cdot B_{\max} \quad (2)$$

Po zakończeniu podstawowego etapu skalowania i wyłonieniu najlepszych struktur neuronowych, postanowiono zastosować inne skalowanie i sprawdzić jaki ma ona wpływ na błąd prognozy. Tym razem wejścia wykorzystanych sieci zostały skalowane poprzez zastosowanie liniowego odwzorowania ekstremów zmiennej na ekstrema funkcji aktywacji w warstwie ukrytej. Dane po przeskalowaniu posiadały wartości z przedziału (-1,1). Do przeliczenia wejść została zastosowana zależność liniowa:

$$y = ax + b \quad (3)$$

gdzie: y – wartość wejściowa przeskalowana, x – wartość rzeczywista przeliczanej składowej, a, b – współczynniki funkcji.

Przy czym wartości współczynników to :

$$a = \frac{2}{x_{\max}}, b = -1 \quad (4)$$

gdzie: x_{\max} – wartość maksymalna zmiennej wejściowej.

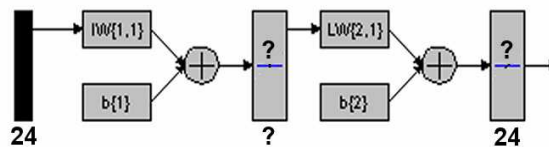
By dokonać przeliczenia wartości znormalizowanych na rzeczywiste należało skorzystać z zależności:

$$x = \frac{x_{\max}}{2} (y + 1) \quad (5)$$

3. BUDOWA STRUKTURY SIECI

W badaniach stosowano wyłącznie sieć z wielowarstwowym wejściem i wyjściem. Zastosowana sieć zawierała w warstwie wyjściowej 24 neurony. Każde wyjście odpowiadało jednej godzinie doby. Na wejście zdecydowano się podawać również 24 informacje. Symulacje rozpoczęto od testowania najprostszych sieci, gdzie w warstwie ukrytej znajdował się jeden neuron. Następnie stopniowo rozbudowywano sieć. W symulacjach rozpatrzono różne kombinacje funkcji aktywacji, jak

również przetestowano sieci pod względem różnych metod uczenia. Tryb postępowania był następujący: na wejścia sieci wprowadzano 24 wielkości godzinowe ceny z dnia poprzedniego (d-1), przeskalowane do przedziału (0,1). W wyniku testów wyłoniono najlepsze struktury neuronowe, które potem poddawano próbom ulepszenia w postaci zmian sposobu skalowania danych jak i zmiany danych wejściowych, które mogły by mieć wpływ na stan nauczania sieci.



Rys. 2. Model sieci wykorzystywanej neuronowej

W celu przeprowadzenia symulacji wykorzystano sieć trójwarstwową, złożoną z warstwy wejściowej, ukrytej oraz wyjściowej. Zrezygnowano z symulacji sieci czterowarstwowych. Powodem był fakt, iż informacje znajdujące się w literaturze przedmiotu [6, 7] wskazywały, że sieci te wykazują się błędem nieco wyższym od sieci z jedną warstwą ukrytą. Dodatkowo uczenie bardziej rozbudowanej sieci jest znacznie trudniejsze, gdyż występuje o wiele większa liczba powiązań między kolejnymi warstwami neuronów.

W celu dokonania prognozy zamodelowano 7 sztucznych sieci neuronowych. Jedna sieć pozwalała na wykonanie prognozy dla jednego dnia tygodnia. Wszystkie sieci były wielowarstwowym preceptorem z algorytmem uczenia metodą propagacji wstecznej błędu.

Tabela 2. Wyniki testowania sieci dla wybranej prognozy (środa)

| Symbol sieci | Dane wejściowe przeskalowane według zwykłej normalizacji (0,1). | Dane wejściowe przeskalowane liniowym odwzorowaniem ekstremów (-1,1). |
|--------------|---|---|
| | Błąd prognozy w [%] | |
| | MAPE _{h,d} av | MAPE _{h,d} av |
| 5 LP rb | 4,83 | 5,10 |
| 9 LP rb | 5,30 | 6,23 |
| 7 LT rb | 5,28 | 6,10 |
| 9 LT rb | 5,17 | 7,14 |
| 5 PP rb | 4,05 | 6,13 |
| 7 PP rb | 4,21 | 5,23 |
| 9 PT rb | 5,19 | 7,06 |
| 4 TP rb | 4,92 | 6,59 |
| 10 TP qn | 4,85 | 5,04 |
| 7 TT rb | 4,53 | 6,94 |
| 9 TT rb | 5,08 | 5,16 |

Oznaczenie: *Cyfra*, *Literal*, *Litera2*, *Oznaczenie 2-literowe* określa symbol struktury neuronowej gdzie:

Cyfra - liczba neuronów w warstwie ukrytej, *Literal* - Funkcja aktywacji w warstwie ukrytej (L - logistyczna, T - Tangensoidalna, P - Liniowa), *Litera2* - Funkcja aktywacji w warstwie wyjściowej (L - logistyczna, T - Tangensoidalna, P - Liniowa), *Oznaczenie 2-literowe* - metoda uczenia sieci (rb - propagacji wstecznej Resilient, qn - propagacji wstecznej quasi-Newtona),

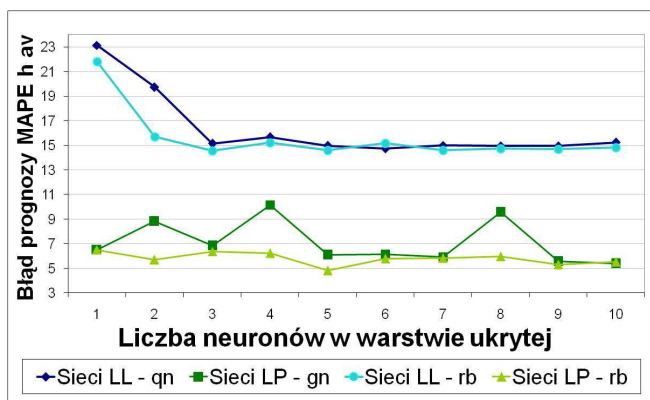
MAPE_{h,d} av - średni bezwzględny błąd procentowy godzinowy.

W tabeli 2 przedstawiono wyniki testowania trójwarstwowych struktur do prognozowania ceny energii dla środy. W zestawieniu wzięły udział wszystkie sieci dające najmniejsze błędy prognozy.

4. TESTOWANIE SIECI

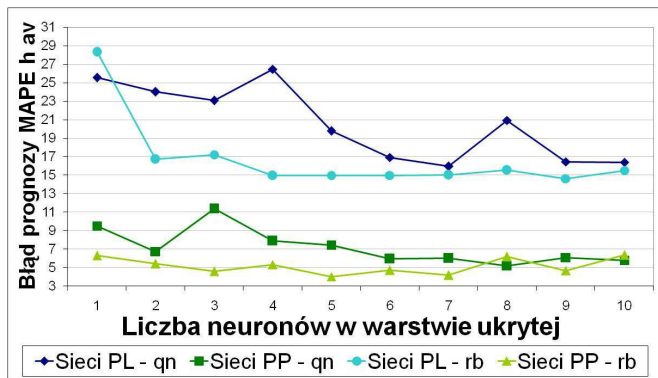
W trakcie testowania dokonano 720 symulacji. Na podstawie badań wyłoniono sieci obarczone najmniejszym błędem prognozy. W toku pracy brano pod uwagę takie czynniki jak zmiana funkcji aktywacji, zmiana sposobu skalowania danych uczących czy zmiana liczby neuronów w warstwie ukrytej.

Aby zobrazować wpływ niektórych z wyżej wymienionych czynników, dokonano zestawień błędów prognozy przy różnych stosowanych strukturach sieci (rys. 2–3).



Rys. 2. Zestawienie wyników sieci neuronowych – funkcje logistyczne

Na podstawie wyników testowanych sieci stwierdzono, że liniowa funkcja aktywacji pojawia się najczęściej w strukturach najlepiej oddających stan rzeczywisty. Nie dostrzeżono by wzrost liczby neuronów prowadził do pogorszenia istotnego jakości prognozy. Kombinacje funkcji aktywacji, w której jedną z nich jest funkcja liniowa, prowadziło do stosowania mniej rozbudowanych struktur sieci, a tym samym oszczędzało czas w nauczaniu sieci. Analizując wykresy, można zauważyć, że struktury z logistyczną funkcją aktywacji warstwy wyjściowej charakteryzowały się największym błędem prognozy w stosunku do pozostałych kombinacji.



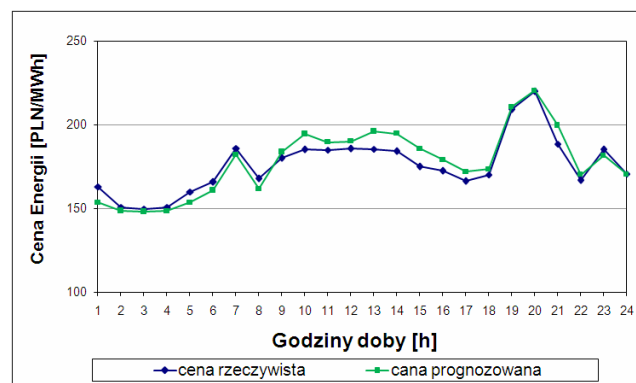
Rys. 3. Zestawienie wyników sieci neuronowych – funkcje liniowe

Analizując wykresy, można zauważyć, że struktury z logistyczną funkcją aktywacji warstwy wyjściowej charakteryzowały się największym błędem prognozy w stosunku do pozostałych kombinacji. Sieci z tangensoidalną warstwą aktywacji na wyjściu wykazywały lepszy

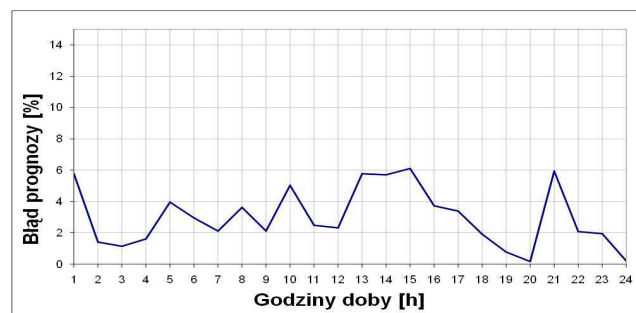
poziom skuteczności wraz z rozbudowywaniem sieci neuronowej. Problemem takich sieci było określenie granicznej liczby neuronów gdyż trend skuteczności sieci był bardzo chaotyczny. Najlepszą funkcją aktywacji warstw wyjściowych okazała się funkcja liniowa, która wykazuje dwie zalety: duża zbieżność wyników z wartościami rzeczywistymi oraz skoordynowany i ustabilizowany poziom błędów w przypadku zmian ilości neuronów w sieci.

5. PODSUMOWANIE

Stworzona prognoza dla poszczególnych dni takich jak: poniedziałek, wtorek, środa, sobota oraz niedziela spełniała założone kryteria. Błąd średni prognozy nie przekraczał 5%. Przykładem może być wykres porównawczy jednego z prognozowanych dni (rys. 4).

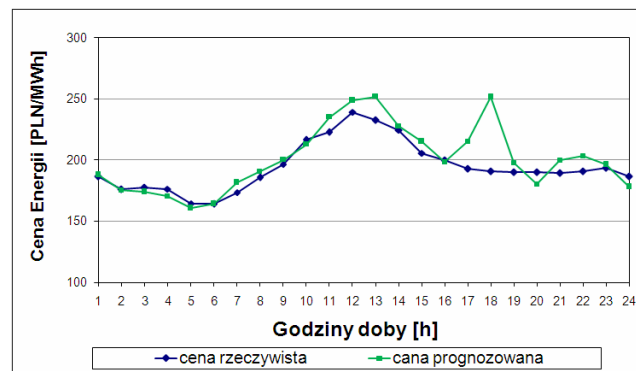


Rys. 4. Przebieg ceny rzeczywistej i prognozowanej dla poniedziałku w 9. tygodniu roku 2010

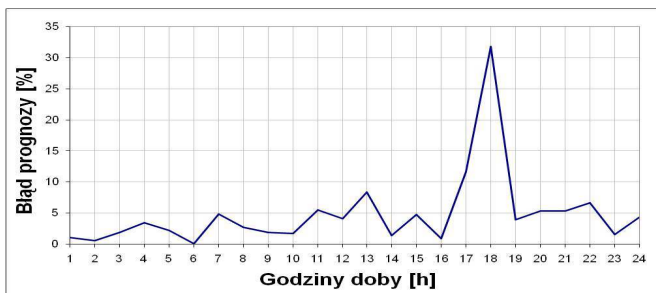


Rys. 5. Błąd bezwzględny prognozy dla poniedziałku w 9. tygodniu roku 2010

Wadą proponowanej metody są jednak pojawiające się nieregularnie duże odchylenia prognozy sięgające powyżej 10% (rys. 6-7). Zbyt mały zbiór danych uczących nie pozwolił przystosować sieci do prawidłowej reakcji w sytuacjach nietypowych. Ogranicza to potencjalne praktyczne zastosowanie metody.

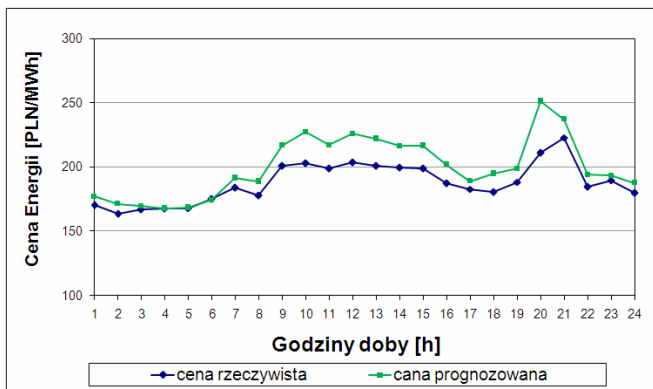


Rys. 6. Przebieg ceny energii rzeczywistej oraz prognozowanej dla soboty w 23. tygodniu roku

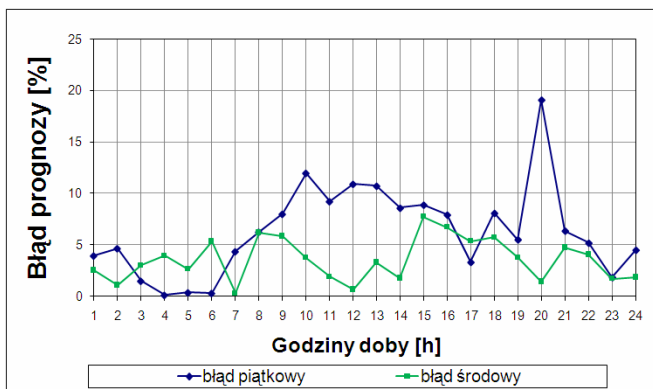


Rys. 7. Błąd bezwzględny prognozy dla soboty w 23. tygodniu roku

Proces uczenia przebiegał podobnie dla wszystkich dni. Pomimo to dla czwartku i piątku (rys.8-9) nie udało się spełnić wymaganych założeń.



Rys. 8. Przebieg ceny energii rzeczywistej oraz prognozowanej dla piątku w 37. tygodniu roku



Rys. 9. Błąd bezwzględny prognozy dla środy oraz piątku w 37. tygodniu roku 2010

Okazało się, że o ile struktura wpływa na jakość uczenia, to głównym i dominującym czynnikiem odpowiadającym za dobre wyniki sieci jest odpowiednio dobrany pakiet danych wejściowych. Zbiór danych wejściowych odgrywa decydującą rolę w procesie nauczania.

USING ARTIFICIAL NEURAL NETWORKS TO FORECASTING THE ENERGY EXCHANGE PRICE

One of the current economic problem is the Energy Exchange price forecasting. Authors' proposal is to implement artificial neural network and MATLAB package for price forecasting. Program for short-time forecast of the Next Day Market price is described. Data normalisation methods are presented in the paper. Neural activation functions are described and the consequences of choosing activation function on neural network learning process are discussed. In the conclusion, comparison of Energy Exchange forecast price and real prices is presented.

Keywords: prediction methods, artificial neural networks, energy market, MATLAB.

Konfrontując wyniki prognozy uzyskanej z zastosowaniem sieci neuronowej, w stosunku do innych metod opisywanych w literaturze przedmiotu [9], autorzy publikacji doszli do wniosku, że prace badawcze związane z prognozowaniem cen na Giełdzie Energii przy pomocy SSN należy kontynuować. Sieć pomimo, że w obecnej strukturze nie spełnia warunku uzyskania mniejszego niż 5% średniego absolutnego błędu procentowego, wykazywała lepsze wyniki prognozy niż modele opierające się na szeregach czasowych. Średni absolutny błąd procentowy metod ARIMA w zależności od konfiguracji wahał się między 8,07% a 21,27% [9].

W przedstawionej formie model może być tylko narzędziem wspomagającym dla pracy wykonywanej przez maklera giełdowego. Model nie jest on zdolny do samodzielnego działania. Potrzebuje on stałej weryfikacji, szczególnie w sytuacjach anormalnych, takich jak dni świąteczne, święta narodowe, wydarzenia ogólnopolskie mające wpływ na produkcję energii elektrycznej czy coraz częściej widocznej gry spekulacyjnej na rynku natychmiastowym energii elektrycznej.

6. BIBLIOGRAFIA

1. Dobrzyński I., Dąsal K., Łyp J., Popławski T., Sowiński J.: Prognozowanie w elektroenergetyce - zagadnienia wybrane. Politechnika Częstochowska 2002.
2. Masters T.: Sieci neuronowe w praktyce. Wydawnictwo Naukowo Techniczne, Warszawa 1996.
3. Klaus R.: Metody uczenia sieci. Materiały wykładowe z witryny internetowej – www.cs.put.poznan.pl/rklaus/assn/uczenie.htm (10 IX 2011).
4. Witryna internetowa: Towarowej Giełdy Energii – www.tge.pl (24 VI 2011).
5. Witryna internetowa: Katedra Inżynierii Komputerowej Politechniki Częstochowskiej – www.kik.pcz.pl/nn/index.php (10 IX 2011)
6. Łyp J.: Prognozowanie cen energii na Rynku Bilansującym z użyciem sztucznych sieci neuronowych. Materiały VII Konferencji Naukowej: Prognozowanie w elektroenergetyce. Częstochowa 2004.
7. Praca magisterska: Pakulski T. :Prognozowanie obciążeń godzinowych i cen na giełdzie energii. Gdańsk 2005.
8. Niedziółka D.: Rynek energii w Polsce. Wydawnictwo Difin, Warszawa 2010.
9. Góra A., Strzała K.: Prognozowanie ceny energii na TGE SA – analiza empiryczna. Journal of Management and Finance 2013.