

## **Ukryte modele Markowa w analizie wyników testu koniunktury gospodarczej**

### **Streszczenie**

W pracy zbadana została możliwość wykorzystania algorytmu Viterbiego do analizy sald odpowiedzi respondentów na pytania testu koniunktury w przemyśle, prowadzonego przez Instytut Rozwoju Gospodarczego Szkoły Głównej Handlowej w Warszawie. W badaniu rozważane były pytania dotyczące oceny stanu obecnego. Do analizy wykorzystane zostały ukryte modele Markowa z warunkowymi rozkładami normalnymi. Pod uwagę brane były modele, w których łańcuchy Markowa mają dwuelementową i trójelementową przestrzeń stanów. Uzyskane wyniki zostały skonfrontowane z pochodzącymi z różnych źródeł datowaniami punktów zwrotnych cyklu koniunkturalnego. Badane modele zostały porównane pod względem skuteczności w wychwytywaniu sygnałów o nadchodzących zmianach w koniunkturze. Przeprowadzone analizy przemawiają za stosowaniem modeli z trzystanowymi łańcuchami Markowa. Wyniki badania sugerują ponadto, iż należy brać pod uwagę opóźnienia między odpowiedziami respondentów a zmianami klimatu koniunktury.

Słowa kluczowe: ukryte modele Markowa, algorytm Viterbiego, test koniunktury, punkty zwrotne cyklu koniunkturalnego

Kody klasyfikacji JEL: E37

---

<sup>‡</sup> Instytut Ekonometrii, Szkoła Główna Handlowa

Michał Bernardelli<sup>±</sup>, Monika Dędyś<sup>±</sup>

## **Hidden Markov Models in Analysis of Results of Business Tendency Surveys**

### **Abstract**

The paper considers the possibility of using the Viterbi algorithm to analyse results of the RIED WSE business surveys in the manufacturing industry. The analysis was focused on the state balances. The hidden Markov models with conditional normal distributions were applied. There were considered models with two-state and three-state Markov chains. The results were compared with the timing of turning points taken from other sources. The tested models were compared in terms of effectiveness in detecting of coming changes in economic conditions. The analysis suggests models with three-state Markov chains be used. The results also suggest that it is necessary to take into account a delay between the opinions of survey respondents and changes in economic climate.

Keywords: hidden Markov models, Viterbi algorithm, business tendency surveys, business cycle turning points

JEL classification: E37

---

<sup>±</sup> Institute of Econometrics, Warsaw School of Economics

## 1. Wprowadzenie

Jednym z problemów pojawiających się w analizie wyniku testów koniunktury jest wyodrębnienie okresów, w których respondenci oceniają sytuację dotyczącą określonych kategorii ekonomicznych jako lepszą lub gorszą. Taka segmentacja szeregów czasowych powstałych na podstawie odpowiedzi na pytania ankiety może być pomocna w ocenie zgodności informacji dostarczanych przez respondentów z obserwowanymi zmianami realnych zmiennych ekonomicznych. Może też dostarczać sygnałów o nadchodzących zmianach klimatu koniunktury. Wydaje się, że ukryte modele Markowa (*hidden Markov models*, HMM), stanowią idealne narzędzie do tego typu analiz. W niniejszej pracy stosujemy HMM do przetworzenia sygnałów zawartych w danych z testu koniunktury w przemyśle, prowadzonego przez Instytut Rozwoju Gospodarczego SGH.

Ukryte modele Markowa były wielokrotnie stosowane do badania tego typu szeregów czasowych (np. Abberger i Nierhaus, 2010; Skrzypczyński, 2008). W pracy proponujemy jednak inne podejście do znajdowania oceny realizacji ukrytego łańcucha Markowa. Stosujemy mianowicie algorytm Viterbiego. Od dziesięcioleci algorytm ten jest obecny między innymi w zastosowaniach HMM w rozpoznawaniu mowy (Rabiner, 1989). Jednak w analizie makroekonomicznych szeregów czasowych jest, według naszej wiedzy, używany niezwykle rzadko. Zaproponowany sposób segmentacji szeregów czasowych pochodzących z testu koniunktury dostarcza interesujących informacji o trafności ocen wydawanych przez respondentów i stwarza nadzieje na opracowanie efektywnego systemu prognozowania punktów zwrotnych cyklu koniunkturalnego.

## 2. Charakterystyka narzędzi

Rozważania rozpoczniemy od przytoczenia definicji ukrytego modelu Markowa. Więcej informacji na temat tych modeli można znaleźć m.in. u Cappé, Moulinesa i Rydén (2005). Częściowo obserwowalny proces stochastyczny  $(X_t, Y_t)_{t=1}^{\infty}$  nazywamy ukrytym modelem Markowa, jeśli spełnione są następujące warunki:

- 1) nieobserwowalny proces  $(X_t)_{t=1}^{\infty}$  jest jednorodnym łańcuchem Markowa ze skończoną przestrzenią stanów  $S_X$  i macierzą prawdopodobieństw przejścia

$$P = [p(x_i, x_j)]_{i,j=1}^k,$$

gdzie  $k$  jest liczbą stanów.

- 2) dla każdego  $t \geq 0$ , zmienne losowe  $Y_1, Y_2, \dots, Y_t$ , pod warunkiem  $(X_1, X_2, \dots, X_t)$  są niezależne i warunkowy rozkład prawdopodobieństwa zmiennej  $Y_i$  dla  $i = 1, 2, \dots, t$ , zależy tylko od  $X_i$ , przy czym

$$P(Y_i \leq y_i | X_i = x_i) = H(y, \theta_{x_i}),$$

gdzie  $H$  jest dystrybuantą indeksowaną parametrami  $\theta_{x_i}$ .

Proces  $(X_t)_{t=1}^{\infty}$  nazywamy ukrytym łańcuchem Markowa. Jeśli przestrzeń stanów ukrytego łańcucha Markowa jest  $k$ -elementowa, to będziemy mówić o  $k$ -stanowym HMM. W pracy rozważamy dwu- i trójstanowe modele z warunkowymi rozkładami normalnymi.

HMM należą do klasy przełącznikowych modeli Markowa. W ekonometrii wykorzystuje się je najczęściej do analizy szeregów finansowych i makroekonomicznych. Zdarza się czasem, że stany składowej nieobserwowalnej modelu mają interpretację ekonomiczną; wydają się powiązane z okresami hossy i bessy, ekspansji i recesji. W tych przypadkach wnioskowanie na temat zachowania się zmiennej nieobserwowalnej może dostarczyć nieocenionych informacji na temat sytuacji ekonomicznej.

W badaniach ekonometrycznych (Hamilton, 1994) do znalezienia oceny nieobserwowalnej zmiennej  $X_t$  w dowolnym momencie czasu  $t$ , gdzie  $t \leq T$  na ogół wykorzystuje się:

- prawdopodobieństwa przefiltrowane, czyli prawdopodobieństwa warunkowe

$$P(X_t = x_t | Y^t = (y_1, y_2, \dots, y_t))$$

- lub prawdopodobieństwa wygładzone

$$P(X_t = x_t | Y^T = (y_1, y_2, \dots, y_T)).$$

Zazwyczaj, gdy  $S_X = \{0, 1\}$  oraz  $P(X_t = 1 | Y^T = (y_1, y_2, \dots, y_T)) > m$ , gdzie  $m$  jest liczbą wybraną arbitralnie przez badacza, przyjmuje się, że w chwili  $t$  ukryty łańcuch Markowa znajdował się w stanie 1. Analogiczne wnioskowanie można przeprowadzić na podstawie prawdopodobieństw przefiltrowanych. Ocenę stanów przeprowadza się więc niejako dla każdej chwili  $t$  oddzielnie i na podstawie tych ocen cząstkowych podaje się ocenę ścieżki ukrytego łańcucha Markowa w okresie od  $t=1$

do  $t=T$ . Wydaje się, że warto wziąć pod uwagę bardziej całościowe podejście do problemu; zamiast zajmować się każdą chwilą oddzielnie można poszukiwać najbardziej prawdopodobnej ścieżki ukrytego łańcucha Markowa w całym okresie od  $t=1$  do  $t=T$ .

Taką ścieżkę wyznaczyć można, stosując algorytm zaproponowany przez Viterbiego w 1967 roku. Umożliwia on znalezienie tak zwanej ścieżki Viterbiego, to jest takiego ciągu stanów, który jest najbardziej prawdopodobny, biorąc pod uwagę zaobserwowane wartości w danym okresie czasu. Formalnie rzecz ujmując, algorytm Viterbiego znajduje ścieżkę

$$\tilde{X}^T = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_T) \in (S_X)^T$$

taką, że

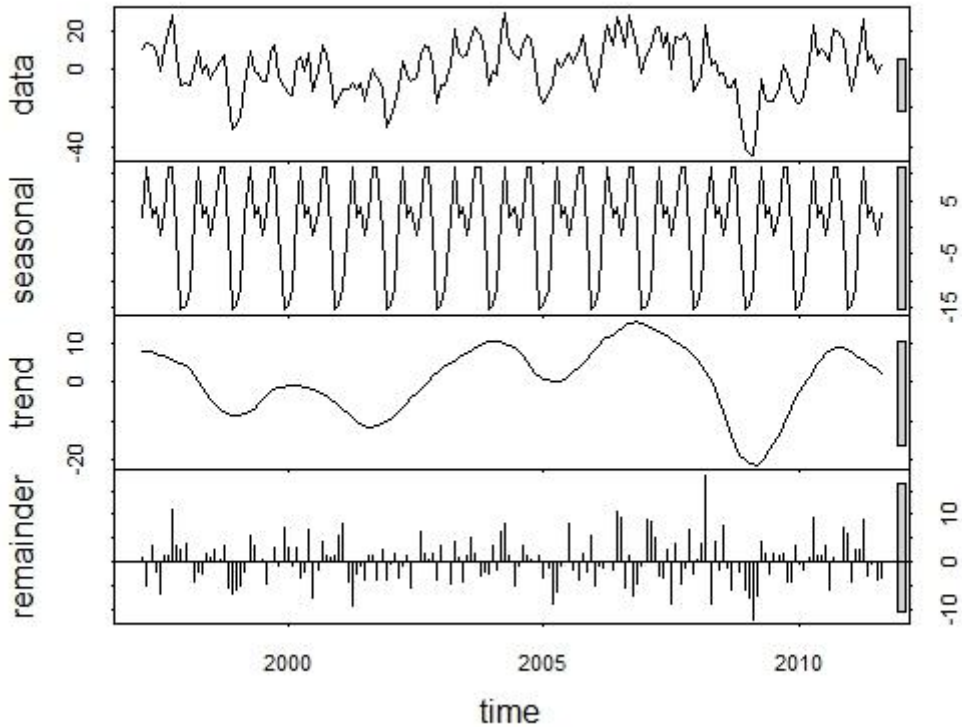
$$P(X_1 = \tilde{x}_1, X_2 = \tilde{x}_2, \dots, X_T = \tilde{x}_T | Y^T = (y_1, y_2, \dots, y_T)) \\ = \max_{(x_1, x_2, \dots, x_T) \in (S_X)^T} P(X_1 = x_1, X_2 = x_2, \dots, X_T = x_T | Y^T = (y_1, y_2, \dots, y_T))$$

### 3. Charakterystyka danych empirycznych

W pracy poddajemy analizie szeregi sald pochodzących z testu koniunktury w przemyśle z okresu od marca 1997 do sierpnia 2011 roku. Badaniem zostały objęte wszystkie pytania testu:

- pytanie 1 - wielkość produkcji,
- pytanie 2 - portfel zamówień ogółem,
- pytanie 3 - portfel zamówień eksportowych,
- pytanie 4 - poziom zapasów produkowanych wyrobów,
- pytanie 5 - ceny produkowanych wyrobów,
- pytanie 6 - poziom zatrudnienia,
- pytanie 7 - sytuacja finansowa przedsiębiorstwa,
- pytanie 8 - ogólna sytuacja gospodarki polskiej.

Uwagę skupiamy na ocenach stanu bieżącego, przy czym szeregi sald zostały oczyszczone z wahań sezonowych i losowych. Do dekompozycji szeregów wykorzystana została procedura STL z pakietu komputerowego R. Zastosowano w niej iteracyjny algorytm z wykorzystaniem regresji lokalnie ważonej, tak zwaną metodę „loess” (Cleveland, Cleveland i McRae, 1990). Dekompozycja szeregu sald odpowiedzi na pytanie o wielkość produkcji przedstawiona została na rysunku 1. Wyeksponowane zostały poszczególne składowe: sezonowa (oznaczona jako *seasonal*), trend i losowa (*remainder*).



Rysunek 1. Dekompozycja szeregu sald odpowiedzi na pytanie o wielkości produkcji z wykorzystaniem procedury STL

Źródło: opracowanie własne na podstawie danych IRG SGH.

Wyniki segmentacji szeregów czasowych konfrontujemy z datowaniami punktów zwrotnych cyklu koniunkturalnego dokonanyimi przez OECD (cykl odchyień) i Drozdowicz-Bieć (2012) (cykl klasyczny). Dodatkowo do analizy sygnałów przetworzonych za pomocą trójstanowych HMM posłużyliśmy się datowaniami punktów zwrotnych cyklu odchyień, określonymi przez Drozdowicz-Bieć (2008). W badaniu sald dotyczących wielkości produkcji wzięliśmy także pod uwagę wyróżnione w pracy (Adamowicz i in., 2012) górne i dolne punkty zwrotne składnika cyklicznego produkcji przemysłowej w Polsce oraz składnika cyklicznego PKB. Na ich podstawie opracowaliśmy zerojedynekowy szereg referencyjny, w którym 0 odpowiada fazie spadkowej, a 1 fazie wzrostu.

#### 4. Wyniki badania empirycznego

Oczyszczone szeregi sald odpowiedzi na pytania o stan aktualny traktujemy jako realizacje składowej obserwowalnej  $(Y_t)_{t=1}^{\infty}$  HMM. W przypadku dwustanowego modelu HMM przyjmujemy, że  $S_X = \{0, 1\}$ . Stan 0 dotyczy okresów, w których respondenci oceniają sytuację jako gorszą, zaś stan 1 wiążemy z sytuacją ocenianą jako lepszą. Formalnie stan 0 utożsamiamy z rozkładem warunkowym o niższej wartości średniej.

Dla każdego z ośmiu szeregów czasowych wyznaczamy najbardziej prawdopodobną ścieżkę ukrytego łańcucha Markowa. Ścieżkę tę traktujemy jako ocenę sytuacji w kategorii objętej pytaniem, dokonaną przez respondentów w okresie objętym badaniem. Ocenę tę zestawiamy z szeregiem referencyjnym. Ściśle rzecz ujmując, w przypadku dwustanowych HMM porównujemy dwa szeregi zerojedynkowe. Zliczamy okresy, w których w obydwu szeregach występuje 0 („trafione 0”) oraz okresy, w których w obydwu szeregach występuje 1 („trafione 1”). Wyniki obliczeń dla szeregu referencyjnego wyznaczonego na podstawie datowań OECD zamieszczono w tabeli 1, zaś dla szeregu referencyjnego związanego z cyklem klasycznym w tabeli 2. Dodatkowo, na rysunkach 2-10 umieszczono najbardziej prawdopodobne ścieżki dla szeregów sald na tle szeregu referencyjnego OECD, zaś na rysunkach 11-18 na tle szeregu referencyjnego przedstawiającego cykl klasyczny.

Pobieżna analiza wyników prowadzi do wniosku, że sygnały generowane przez szeregi sald odpowiedzi na pytanie o ceny produkowanych wyrobów (rysunki 7 i 15) niezbyt dobrze korelują z szeregami referencyjnymi. Nieco lepiej, ale nadal niesatysfakcjonująco jest w przypadku portfela zamówień eksportowych (rysunki 4 i 13). Do szeregu referencyjnego OECD (cykl odchyień) dobrze dopasowane są salda odpowiedzi na pytanie o wielkość produkcji (rysunki 2 i 11) i poziom zapasów produkowanych wyrobów<sup>1</sup> (rysunki 5-6 i 14). Ocena sytuacji finansowej przedsiębiorstwa (rysunki 9 i 17) oraz ocena sytuacji gospodarki polskiej (rysunki 10 i 18) wydają się mieć związek z datowaniami punktów zwrotnych w cyklu klasycznym. Nieco gorsze dopasowanie do obydwu szeregów referencyjnych wykazują szeregi zerojedynkowe związane z pytaniem o portfel zamówień ogółem (rysunki 3 i 12) i poziom zatrudnienia (rysunki 8 i 16). Zważywszy na duże opóźnienia

---

<sup>1</sup> W przypadku poziomu zapasów przedstawione zostały dwa wykresy. Najbardziej prawdopodobną ścieżkę przedstawiono na rysunku 6. Odpowiedzi respondentów korespondują z faktem, iż poziom zapasów jest zmienną antycykliczną. W celu ilustracji tego faktu na rysunku 5 przedstawiono szereg z zamienionymi stanami, tj.  $\{0,1\} \rightarrow \{1,0\}$ .

w identyfikacji punktów zwrotnych w cyklu klasycznym, należy uznać, że większy związek mają one z cyklem odchyień OECD. Niestety, sygnały o zmianie klimatu koniunktury emitowane przez przefiltrowane szeregi są bardzo często opóźnione względem datowań punktów zwrotnych w szeregach referencyjnych.

Tabela 1. Porównanie zgodności najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytania ankietowe z szeregiem referencyjnym OECD

Nr pyt.	Treść pytania	Liczba odpowiedzi zgodnych ze stanem 0	Liczba odpowiedzi zgodnych ze stanem 1	Liczba zgodnych odpowiedzi	Liczba niezgodnych odpowiedzi	Średnia liczba trafionych odpowiedzi
1	wielkość produkcji	72	78	150	24	86,2 %
2	portfel zamówień ogółem	72	77	149	25	85,6 %
3	portfel zamówień eksportowych	55	74	129	45	74,1 %
4	poziom zapasów produkowanych wyrobów	80	61	141	33	81,0 %
5	ceny produkowanych wyrobów	55	51	106	68	60,9 %
6	poziom zatrudnienia	70	58	128	46	73,6 %
7	sytuacja finansowa przedsiębiorstwa	72	73	145	29	83,3 %
8	ogólna sytuacja gospodarki polskiej	64	77	141	33	81,0 %

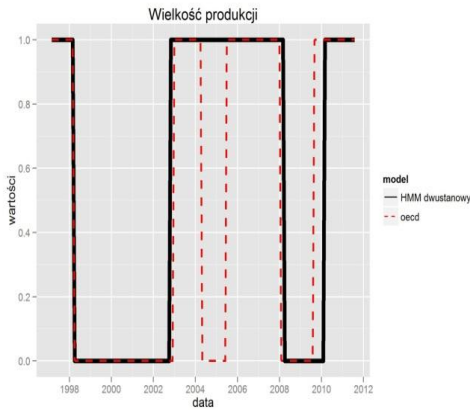
Źródło: obliczenia własne na podstawie danych IRG SGH i OECD.



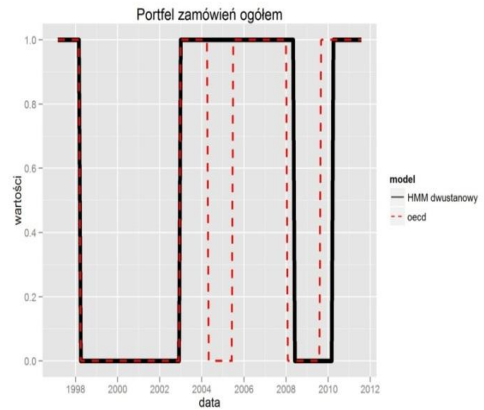
Tabela 2. Porównanie zgodności najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytania ankietowe z szeregiem referencyjnym przedstawiającym cykl klasyczny

Nr pyt.	Treść pytania	Liczba odpowiedzi zgodnych ze stanem 0	Liczba odpowiedzi zgodnych ze stanem 1	Liczba zgodnych odpowiedzi	Liczba niezgodnych odpowiedzi	Średnia liczba trafionych odpowiedzi
1	wielkość produkcji	58	92	150	24	86,2 %
2	portfel zamówień ogółem	56	89	145	29	83,3 %
3	portfel zamówień eksportowych	45	92	137	37	78,7 %
4	poziom zapasów produkowanych wyrobów	61	70	131	43	75,3 %
5	ceny produkowanych wyrobów	32	58	90	84	51,7 %
6	poziom zatrudnienia	56	72	128	46	73,6 %
7	sytuacja finansowa przedsiębiorstwa	56	85	141	33	81,0 %
8	ogólna sytuacja gospodarki polskiej	48	89	137	37	78,7 %

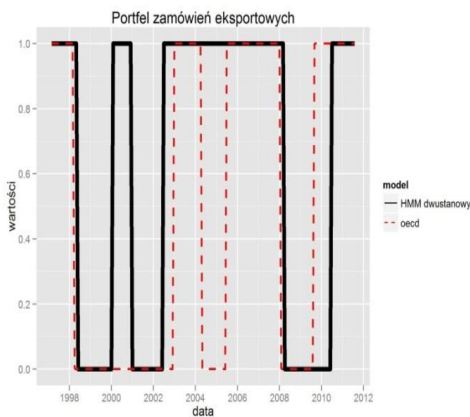
Źródło: obliczenia własne na podstawie danych IRG SGH i Drozdowicz-Bieć (2012).



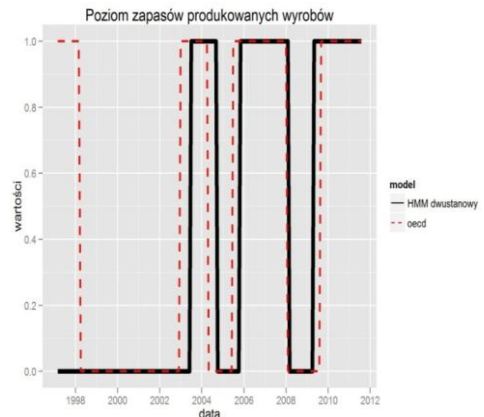
Rysunek 2. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym OECD



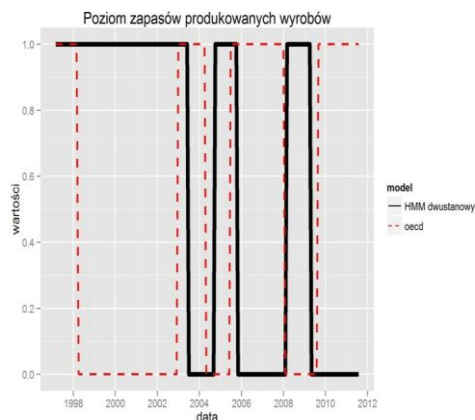
Rysunek 3. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień ogółem (pytanie 2) z szeregiem referencyjnym OECD



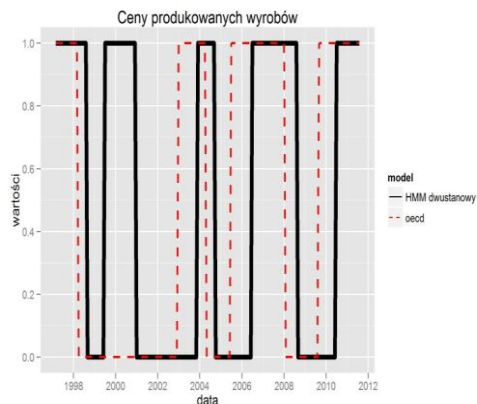
Rysunek 4. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień eksportowych (pytanie 3) z szeregiem referencyjnym OECD



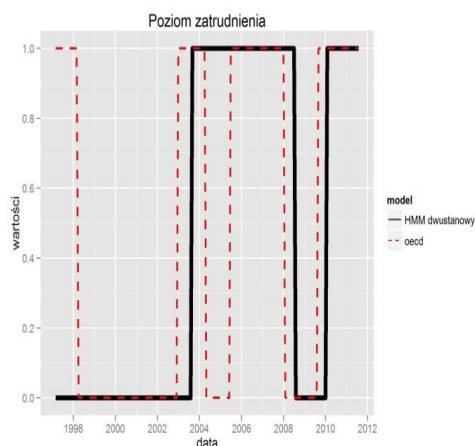
Rysunek 5. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o poziom zapasów produkowanych wyrobów (pytanie 4) z szeregiem referencyjnym OECD



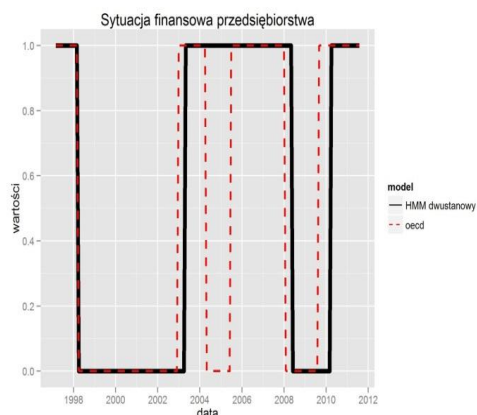
Rysunek 6. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o poziom zapasów produkowanych wyrobów (pytanie 4) z szeregiem referencyjnym OECD



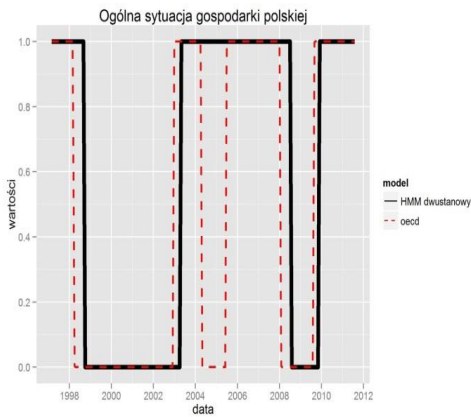
Rysunek 7. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o ceny produkowanych wyrobów (pytanie 5) z szeregiem referencyjnym OECD



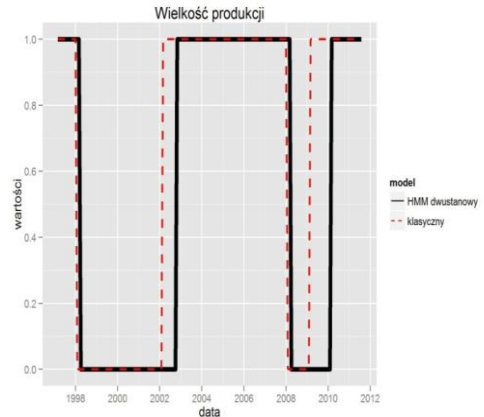
Rysunek 8. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o poziom zatrudnienia (pytanie 6) z szeregiem referencyjnym OECD



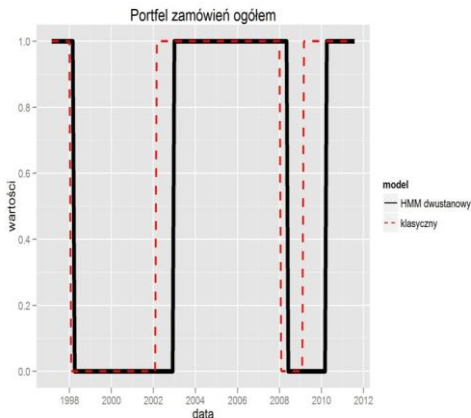
Rysunek 9. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o sytuację finansową przedsiębiorstwa (pytanie 7) z szeregiem referencyjnym OECD



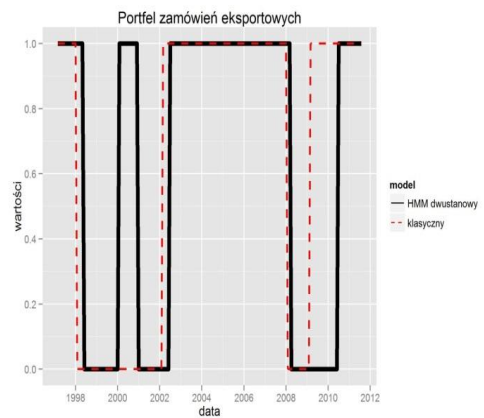
Rysunek 10. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o ogólną sytuację gospodarki polskiej (pytanie 8) z szeregiem referencyjnym OECD



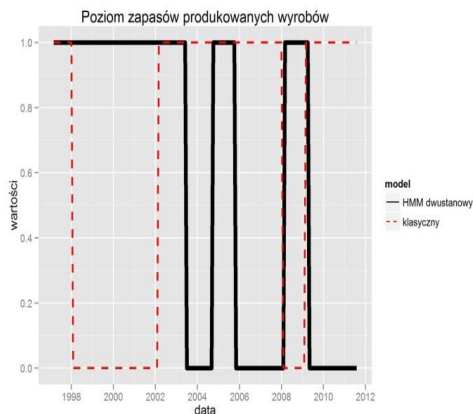
Rysunek 11. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym klasycznym



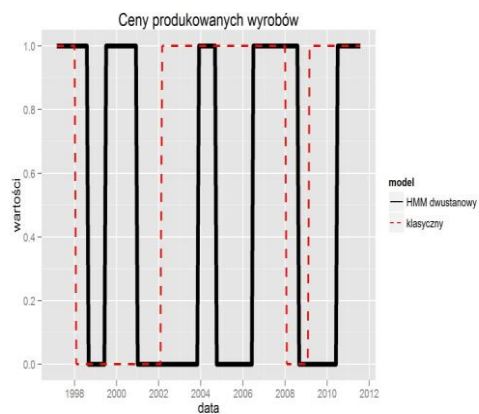
Rysunek 12. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień ogółem (pytanie 2) z szeregiem referencyjnym klasycznym



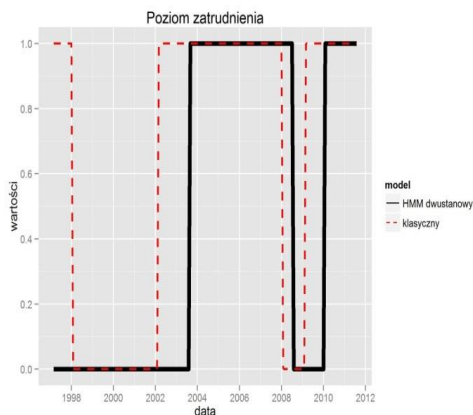
Rysunek 13. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień eksportowych (pytanie 3) z szeregiem referencyjnym klasycznym



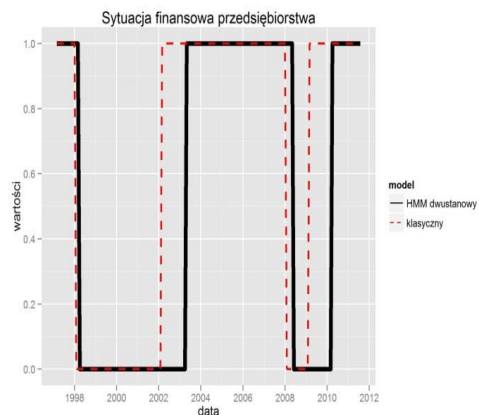
Rysunek 14. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o poziom zapasów produkowanych wyrobów (pytanie 4) z szeregiem referencyjnym klasycznym



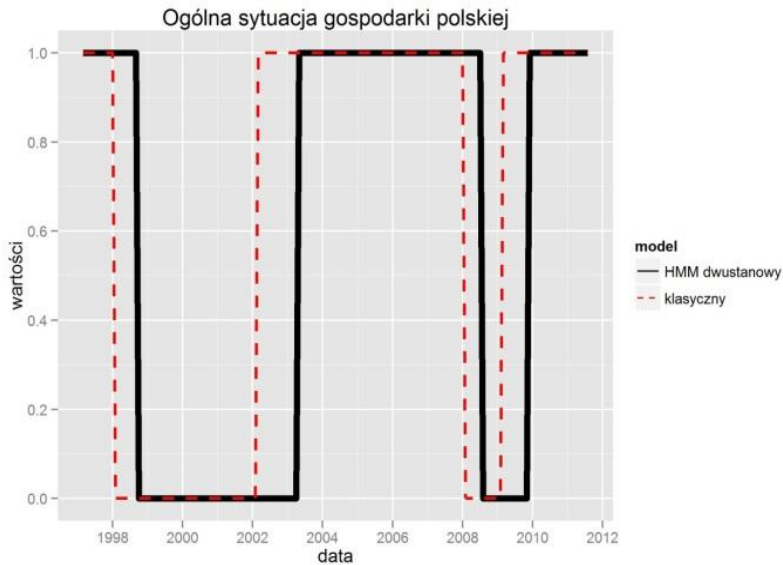
Rysunek 15. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o ceny produkowanych wyrobów (pytanie 5) z szeregiem referencyjnym klasycznym



Rysunek 16. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o poziom zatrudnienia (pytanie 6) z szeregiem referencyjnym klasycznym



Rysunek 17. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o sytuację finansową przedsiębiorstwa (pytanie 7) z szeregiem referencyjnym klasycznym



Rysunek 18. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o ogólną sytuację gospodarki polskiej (pytanie 8) z szeregiem referencyjnym klasycznym

Źródło: opracowanie własne na podstawie danych IRG SGH, OECD i Drozdowicz-Bieć (2012).

Interesujące wyniki uzyskujemy, badając możliwości zastosowania trójstanowych HMM do analizy szeregów sald. Poza sytuacjami, w których przefiltrowane salda jednoznacznie wskazują na wzrost lub spadek, chcemy uwzględnić okresy charakteryzujące się dużą niepewnością ocen stanu koniunktury. W zbiorze  $S_X = \{0, \frac{1}{2}, 1\}$  stan  $\frac{1}{2}$  miał być takim stanem przejściowym, niepewnym. Tak jak w przypadku dwustanowych HMM, stanowi 0 przypisaliśmy rozkład warunkowy o najmniejszej wartości oczekiwanej, zaś stanowi 1 rozkład o największej wartości oczekiwanej. Wprowadzenie stanu pośredniego sprawia, że model HMM inaczej przetwarza sygnały ukryte w odpowiedziach respondentów, jest bardziej plastyczny i tym samym daje nadzieję na szybsze wychwytywanie informacji o zmianie klimatu koniunktury. Wyniki obliczeń trójstanowych HMM dla szeregu referencyjnego powstałego na podstawie datowań OECD zamieszczono w tabeli 3, zaś dla szeregu referencyjnego przedstawiającego cykl klasyczny w tabeli 4.

Tabela 3. Porównanie zgodności najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytania ankietowe z szeregiem referencyjnym OECD

Nr pyt.	Treść pytania	Liczba odpowiedzi zgodnych ze stanem 0	Liczba odpowiedzi zgodnych ze stanem 1	Liczba zgodnych odpowiedzi	Liczba niezgodnych odpowiedzi	Średnia liczba trafionych odpowiedzi
1	wielkość produkcji	46	72	118	56	67,8 %
2	portfel zamówień ogółem	48	63	111	63	63,8 %
3	portfel zamówień eksportowych	51	51	102	72	58,6 %
4	poziom zapasów produkowanych wyrobów	50	58	108	66	62,1 %
5	ceny produkowanych wyrobów	37	39	76	98	43,7 %
6	poziom zatrudnienia	55	54	109	65	62,6 %
7	sytuacja finansowa przedsiębiorstwa	72	39	111	63	63,8 %
8	ogólna sytuacja gospodarki polskiej	63	49	112	62	64,4 %

Źródło: obliczenia własne na podstawie danych IRG SGH i OECD.

Obok modelu, w którym nie narzucamy żadnych warunków na prawdopodobieństwa przejścia ukrytego łańcucha Markowa, rozważamy też taki, w którym  $p(1, 0)=p(0, 1)=0$ . Blokując przejścia w jednym kroku między stanami 0 oraz 1, wymuszamy niejako na modelu, aby pomiędzy okresami rozpoznawanymi jako lepsze i gorsze uwzględniał fazę przejściową. Możliwe jest również występowanie fazy przejściowej między okresami tego samego typu. Taką przejściową nieznaczną zmianę możemy interpretować jako fałszywy sygnał o zmianie klimatu koniunktury. Warto zauważyć, że dla większości analizowanych pytań najbardziej prawdopodobne ścieżki otrzymane z modelu z zablokowanymi przejściami i z modelu z pełną macierzą prawdopodobieństw przejścia są niemal identyczne.

Tabela 4. Porównanie zgodności najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytania ankietowe z szeregiem referencyjnym przedstawiającym cykl klasyczny

Nr pyt.	Treść pytania	Liczba odpowiedzi zgodnych ze stanem 0	Liczba odpowiedzi zgodnych ze stanem 1	Liczba zgodnych odpowiedzi	Liczba niezgodnych odpowiedzi	Średnia liczba trafionych odpowiedzi
1	wielkość produkcji	37	78	115	78	66,1 %
2	portfel zamówień ogółem	38	70	108	66	62,1 %
3	portfel zamówień eksportowych	42	59	101	73	58,0 %
4	poziom zapasów produkowanych wyrobów	35	65	100	74	57,5 %
5	ceny produkowanych wyrobów	29	44	73	101	42,0 %
6	poziom zatrudnienia	46	68	114	60	65,5 %
7	sytuacja finansowa przedsiębiorstwa	57	46	103	71	59,2 %
8	ogólna sytuacja gospodarki polskiej	47	61	108	66	62,1 %

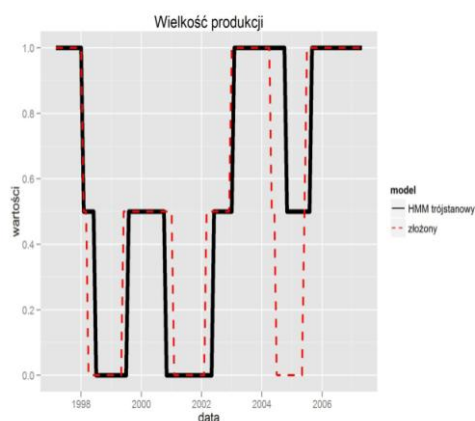
Źródło: obliczenia własne na podstawie danych IRG SGH i Drozdowicz-Bieć (2012).

Odrębny problem stanowi porównanie najbardziej prawdopodobnej ścieżki ukrytego trójstanowego łańcucha Markowa z ocenami stanu koniunktury pochodzącymi z innych źródeł. Wydaje się, że okresy, w których różni badacze mają odmienne zdanie na temat stanu koniunktury, mogą być utożsamiane z przebywaniem ukrytego łańcucha Markowa w „niepewnym” stanie  $\frac{1}{2}$ . W badaniu porównujemy datowania punktów zwrotnych cyklu odchyłeń, dokonane przez OECD i Drozdowicz-Bieć (2008). Ze względu na to, że druga z tych analiz kończy się w maju 2007 roku ograniczamy zakres badania do tego okresu. Na podstawie datowań pochodzących ze wspomnianych źródeł tworzymy szereg referencyjny w następujący sposób. Otóż, jeśli w okresie  $t$  badacze zgodnie wskazują na fazę wzrostu, to w szeregu referencyjnym pojawia się wartość 1. Jeśli badacze zgodnie wskazują fazę spadku, to w chwili  $t$  w szeregu

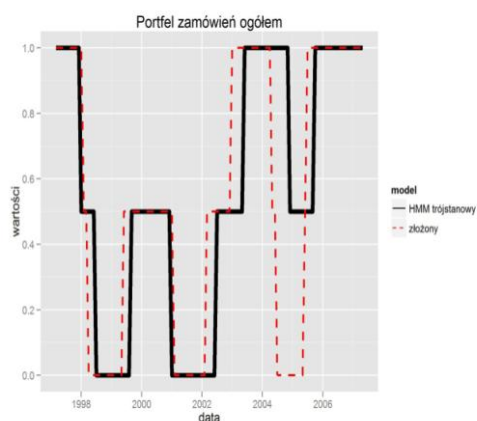


referencyjnym mamy wartość 0. W przypadku braku zgodności badaczy w szeregu pojawia się wartość  $\frac{1}{2}$ .

Okazuje się, że w złożony szereg referencyjny szczególnie dobrze wpisują się najbardziej prawdopodobne ścieżki ukrytych łańcuchów Markowa związanych z pytaniem o wielkość produkcji i portfel zamówień ogółem (rysunki 19 i 20).



Rysunek 19. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) ze złożonym szeregiem referencyjnym

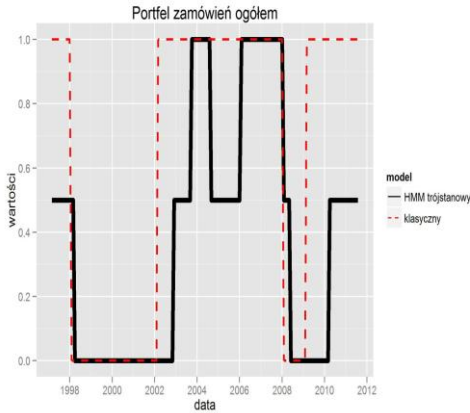


Rysunek 20. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień ogółem (pytanie 2) ze złożonym szeregiem referencyjnym

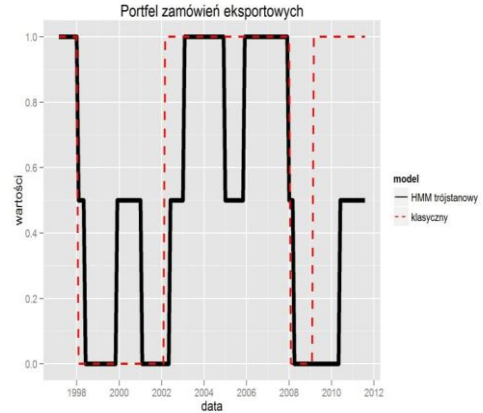
Źródło: opracowanie własne na podstawie danych IRG SGH, OECD i Drozdowicz-Bieć (2008).

Porównanie ocen stanów ukrytego łańcucha Markowa dla trójstanowych HMM z datowaniami punktów zwrotnych cyklu odchyień OECD i cyklu klasycznego prowadzi do wniosku, że najbardziej prawdopodobne ścieżki łańcuchów Markowa zdają się lepiej wpisywać w cykl klasyczny. Sytuację taką obserwujemy dla szeregów czasowych związanych z pytaniem o portfel zamówień ogółem (rysunek 21), portfel zamówień eksportowych (rysunek 22) oraz poziom zapasów produkowanych wyrobów (rysunek 23) i poziom zatrudnienia (rysunek 24). Warto zauważyć, że w przypadku pytania o poziom zapasów produkowanych wyrobów najbardziej prawdopodobna ścieżka uzyskana dla modelu dwustanowego (rysunek 5) nieco lepiej wpisywała się w cykl

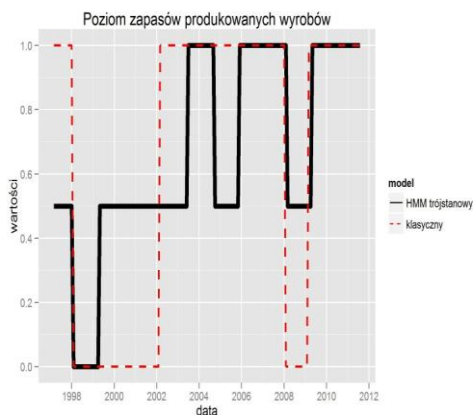
odchylen OECD. Tak też było w przypadku pytania o portfel zamówień ogółem (rysunki 3 i 21). Z kolei w przypadku pytania o poziom zamówień eksportowych (rysunki 4 i 22) dwustanowy HMM nie generował satysfakcjonującej dekompozycji szeregu sald. Trzeba podkreślić, że najbardziej prawdopodobna ścieżka ukrytego trójstanowego łańcucha Markowa związana z pytaniem o wielkość produkcji komponuje się dobrze zarówno z cyklem odchylen jak i z cyklem klasycznym (rysunki 25 i 26).



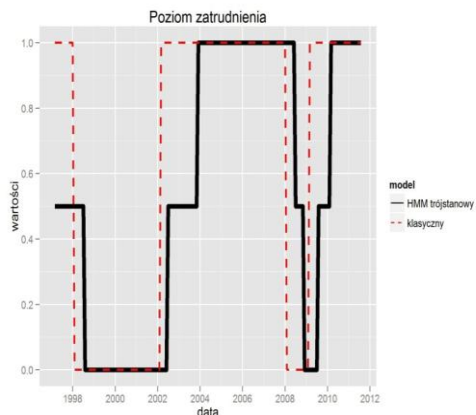
Rysunek 21. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień ogółem (pytanie 2) z szeregiem referencyjnym klasycznym



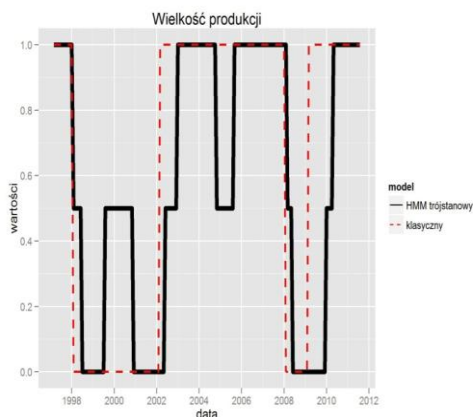
Rysunek 22. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień eksportowych (pytanie 3) z szeregiem referencyjnym klasycznym



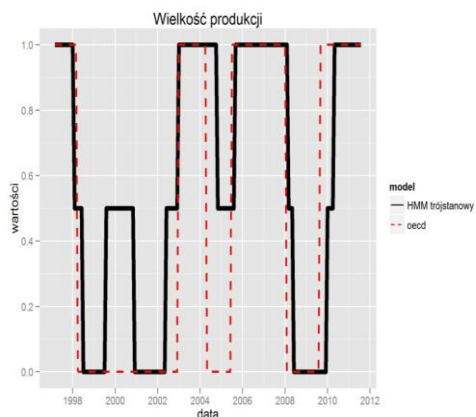
Rysunek 23. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o poziom zapasów produkowanych wyrobów (pytanie 4) z szeregiem referencyjnym klasycznym



Rysunek 24. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o poziom zatrudnienia (pytanie 6) z szeregiem referencyjnym klasycznym



Rysunek 25. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym klasycznym



Rysunek 26. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym OECD

Źródło: opracowanie własne na podstawie danych IRG SGH, OECD i Drozdowicz-Bieć (2008).

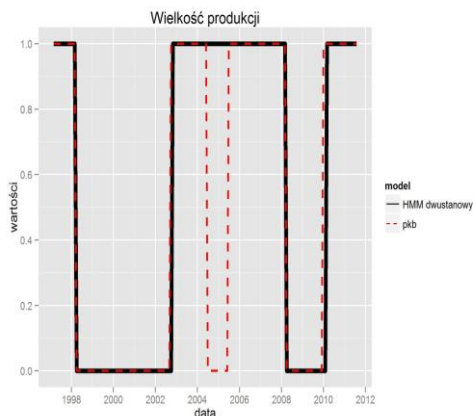
Pojawienie się stanu  $\frac{1}{2}$  w najbardziej prawdopodobnej ścieżce może być interpretowane jako sygnał o nadchodzącej zmianie koniunktury. Okazuje się, że w pewnych sytuacjach trójstanowe HMM wcyedniej generują sygnały o zmieniającej się koniunkturze niż modele dwustanowe. Z sytuacją taką mamy do czynienia w przypadku cyklu klasycznego i pytań o poziom zatrudnienia (rysunki 16 i 24), wielkość produkcji (rysunki 11 i 15) oraz poziom zamówień ogółem (rysunki 12 i 21).

Przetworzone przez HMM informacje z ankiet odnosimy również do wielkości realnych. Najbardziej prawdopodobne ścieżki dla dwu- i trójstanowych modeli HMM porównujemy z datowaniami górnych i dolnych punktów zwrotnych składnika cyklicznego PKB i składnika cyklicznego produkcji dokonanyymi w pracy Adamowicz i in. (2012). W przypadku datowań kwartalnych PKB przypisaliśmy wartość 1 wszystkim miesiącom kwartału, w którym wystąpił górny punkt zwrotny w szeregu referencyjnym (o częstotliwości miesięcznej), a wartość 0 tym, w których wystąpił dolny punkt zwrotny (rysunki 27-32). Stosunkowo najlepszym dopasowaniem do szeregu referencyjnego PKB charakteryzują się wielkość produkcji (rysunki 27 i 28), portfel zamówień ogółem (rysunki 29 i 30) oraz portfel zamówień eksportowych. W każdym z wymienionych przypadków trójstanowe HMM wydają się generować bardziej adekwatne sygnały. Z analogiczną sytuacją mamy do czynienia, porównując model odpowiadający wielkości produkcji z szeregiem referencyjnym uzyskanym ze składnika cyklicznego produkcji (rysunki 31 i 32). Także w tym przypadku model trójstanowy daje lepsze dopasowanie niż model dwustanowy.

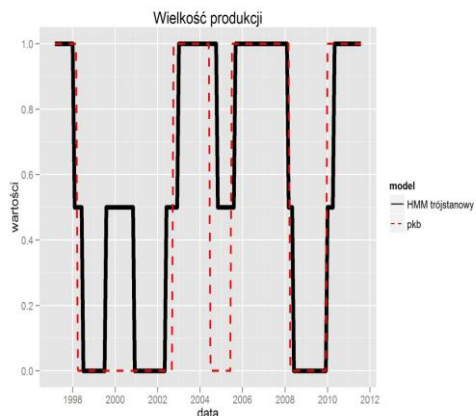
## 5. Wnioski

Algorytm Viterbiego stanowi obiecujące narzędzie analizy szeregów powstałych na podstawie wyników testu koniunktury. Wstępne porównanie prawdopodobieństw przefiltrowanych wyznaczonych w pracy Skrzypczyńskiego (2008) z wyznaczonymi przez nas najbardziej prawdopodobnymi ścieżkami dwustanowych modeli HMM nasuwają przypuszczenie, że te dwie metody mogą prowadzić do różnych wyników. Są to ustalenia wstępne, wymagające dalszych badań.

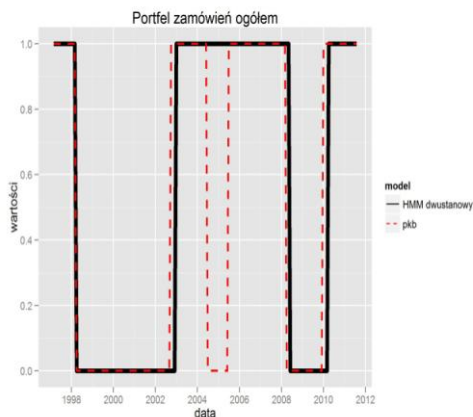
Wyniki testu koniunktury w przemyśle IRG SGH zdają się dobrze dopasowywać do wyników analiz koniunktury pochodzących z innych badań. Sygnały o zmianie koniunktury pojawiają się jednak z pewnym opóźnieniem.



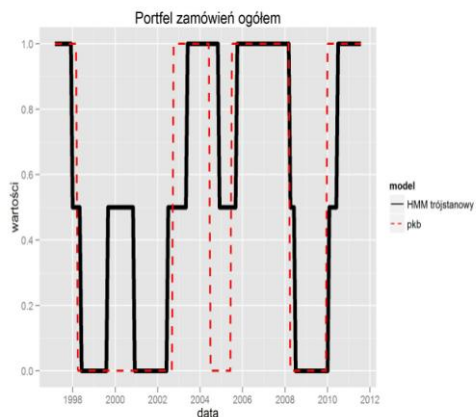
Rysunek 27. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym PKB



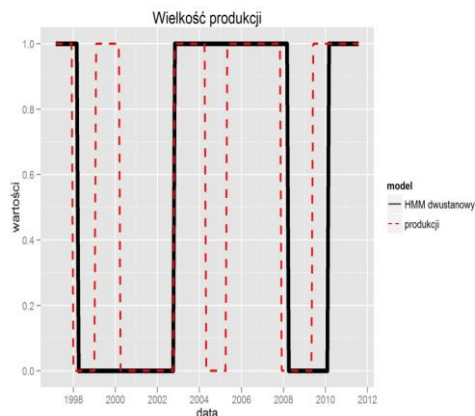
Rysunek 28. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym PKB



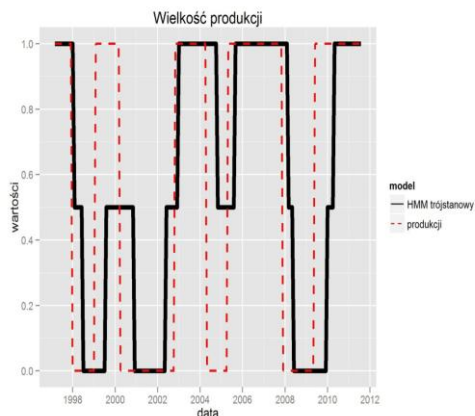
Rysunek 29. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień ogółem (pytanie 2) z szeregiem referencyjnym PKB



Rysunek 30. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o portfel zamówień ogółem (pytanie 2) z szeregiem referencyjnym PKB



Rysunek 31. Porównanie najbardziej prawdopodobnej ścieżki dwustanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym produkcji



Rysunek 32. Porównanie najbardziej prawdopodobnej ścieżki trójstanowego HMM dla szeregu sald odpowiedzi na pytanie o wielkość produkcji (pytanie 1) z szeregiem referencyjnym produkcji

Źródło: opracowanie własne na podstawie danych IRG SGH, OECD i Adamowicz i in. (2012).

Z cyklem odchyleń zdają się współgrać przetworzone przez dwustanowe HMM szeregi sald odpowiedzi na pytanie o wielkość produkcji i poziom zapasów. Punkty zwrotne w cyklu klasycznym zdają się być dobrze sygnalizowane przede wszystkim przez szeregi sald odpowiedzi na pytanie o sytuację finansową przedsiębiorstwa i ocenę sytuacji gospodarki polskiej, a najslabiej – przez szereg zero-jedynkowy sald odpowiedzi na pytanie o ceny produkowanych wyrobów. Stosunkowo najlepsze dopasowanie do szeregu referencyjnego PKB wykazują przetworzone przez dwustanowe HMM szeregi sald pytań o wielkość produkcji, portfel zamówień ogółem i portfel zamówień eksportowych.

Wprowadzenie trójstanowych modeli HMM wzbogaca analizę i może poprawić identyfikację punktów zwrotnych cyklu koniunkturalnego na podstawie wyników badań ankietowych.

## Literatura

Abberger K., Nierhaus W., *Markov-switching and the Ifo business climate: The Ifo business cycle traffic lights*, „Journal of Business Cycle Measurement and Analysis”, nr 2, 2010

- Adamowicz E., Dudek S., Pachucki D., Walczyk K., *Wahania cykliczne w Polsce i strefie euro*, „Prace i materiały Instytutu Rozwoju Gospodarczego SGH”, nr 89, Oficyna Wydawnicza SGH, Warszawa 2012
- Cappé O., Moulines E., Rydén T., *Inference in hidden Markov models*, Springer Series in Statistics, 2005
- Cleveland R. B., Cleveland W. S., McRae J. E., Terpenning I., *STL: A seasonal-trend decomposition procedure based on loess*, „Journal of Official Statistics”, vol. 6, 1990
- Drozdowicz-Bieć M., *Cykle i wskaźniki koniunktury*, Poltext, Warszawa 2012
- Drozdowicz-Bieć M., *Od recesji do boomu. Wahania cykliczne polskiej gospodarki 1990-2007*, w: „Koniunktura gospodarcza – 20 lat doświadczeń Instytutu Rozwoju Gospodarczego SGH”, pr. zb. pod red. E. Adamowicz, „Prace i Materiały Instytutu Rozwoju Gospodarczego SGH”, Oficyna Wydawnicza SGH, Warszawa 2008
- Hamilton J. D., *Time series analysis*, Princeton University Press, Princeton, New Jersey 1994
- Rabiner L. R., *A tutorial on hidden Markov models and selected applications in speech recognition*, „Proceedings of the IEEE”, vol. 77, 1989
- Skrzypczyński G., *Wykorzystanie ukrytych modeli Markowa do wyznaczania punktów zwrotnych koniunktury na podstawie danych ankietowych*, praca niepublikowana, badanie statutowe Instytut Ekonometrii SGH, Warszawa 2008
- Viterbi A., *Error bounds for convolutional codes and an asymptotically optimum decoding algorithm*, „IEEE Transactions on Information Theory”, vol. 13, 1967