

Andrzej Romanowski*, Krzysztof Grudzień*, Dominik Sankowski*

Przegląd możliwości wykorzystania łańcuchów Markowa w rozwiązaniu tomograficznego problemu odwrotnego metodami Monte Carlo

1. Wprowadzenie

W tomografii, celem metod przetwarzania danych tomograficznych jest zwykle zrekonstruowanie obrazu rozkładu materiału w przestrzeni pomiarowej (tomogramu). Aby było to możliwe, konieczne jest rozwiązanie tzw. problemu odwrotnego. Problem, czy też zagadnienie odwrotne oznacza odtworzenie stanu badanego zjawiska na podstawie przeprowadzonych zewnętrznych obserwacji [2, 4]. W praktyce oznacza to wyznaczenie szukanych parametrów danego obiektu w oparciu o zebrane wartości pomiarowe dotyczące jego samego lub też towarzyszącego mu otoczenia [1]. Natomiast rekonstrukcja obrazu w tomografii polega na takim rozwiązaniu zagadnienia odwrotnego, aby na podstawie zebranych wartości pomiarowych wyznaczyć wartości, jakie należy przypisać pikselom, odpowiadającym przestrzennie zrekonstruowanemu rozkładowi mierzonej wartości własności fizycznej badanej przestrzeni. W przypadku tomografii pojemnościowej (ECT – *Electrical Capacitance Tomography*) będzie to wartość stałej dielektrycznej, bądź po prostu wartość przenikalności elektrycznej.

Zrekonstruowane obrazy pozwalają m.in. zrozumieć naturę badanych procesów, których przy użyciu innych metod nie udałoby się zobrazować, unikając jednocześnie ich zaburzenia. Obrazy pozwalają również wizualizować wnętrza reaktorów, dając wgląd w bieżącą fazę procesu. Jakość obrazów, rozumiana jako dokładność odwzorowania graficznego przestrzeni pomiarowej, czyli pośrednio dokładność rozwiązania problemu odwrotnego, jest w większości zastosowań tomografii bardzo istotna; szczególnie w przypadku badania natury danego zjawiska.

Problem odwrotny w przypadku ECT można scharakteryzować kilkoma cechami. Mianowicie **jest to zagadnienie nieliniowe, źle uwarunkowane, źle postawione oraz pod-określone** [70]. Nieliniowość polega na nieliniowej zależności zachowania pola elektrycznego pod wpływem właściwości oraz rozkładu substancji obecnej w przestrzeni po-

* Katedra Informatyki Stosowanej, Politechnika Łódzka

miarowej. Cecha ta jest związana z występowaniem tzw. efektu miękkiego pola (ang. soft-field effect). Złe uwarunkowanie i postawienie problemu jest następstwem faktu, że nawet duże zmiany w rozkładzie właściwości powodują nieznaczne zmiany odczytu wartości pomiarowych. Definicja źle postawionego (źle uwarunkowanego) zagadnienia (*ill-posed problem*, [3]) wymaga potwierdzenia przynajmniej jednego z trzech warunków: brak rozwiązania, rozwiązanie istnieje lecz nie jest unikalne, rozwiązanie nie jest ciągłą funkcją dla dziedziny problemu. W ECT mamy zwykle do czynienia ze spełnieniem drugiego warunku, jako że różne rozkłady przenikalności dielektrycznej mogą powodować takie same odczyty pomiarowe, najczęściej obciążone szumem dodatkowo fałszującym wyniki obserwacji. Wreszcie w związku z tym, że liczba wartości pomiarowych, czyli danych jest znacznie mniejsza od liczby szukanych, czyli najczęściej liczby pikseli w obrazie, problem odwrotny w ECT jest pod-określony.

W celu zrekonstruowania obrazu tomograficznego, a przy okazji przewyciężenia powyższych trudności, stosowanych jest szereg klasycznych algorytmów rekonstrukcji obrazów, zarówno prostych, iteracyjnych, liniowych i złożonych [7].

W związku ze wspomnianymi trudnościami matematycznymi, klasyczne metody rekonstrukcji obrazów tomograficznych opierają się w dużej mierze na aproksymacji i uproszczeniach. Co za tym idzie, wprowadzają one do rozwiązania trudne do oszacowania błędy powodując rozmycia obrazów i wygładzając graficzne rozwiązanie problemu odwrotnego w postaci tomogramu.

2. Podejście statystyczne

Alternatywne do metod klasycznych podejście do rozwiązywania problemu odwrotnego istnieje przy użyciu statystyki Bayesa oraz metod Monte Carlo łańcuchów Markowa. Z perspektywy twierdzenia Bayesa [70] nie istnieje zasadnicza różnica pomiędzy własnościami rozpatrywanego systemu, a danymi wyjściowymi, czyli obserwacjami; wszystkie parametry są rozpatrywane jako zmienne losowe X . Analogicznie można potraktować system pomiarowy ECT, co szczegółowo zostało opisane w [8].

W ogólnej postaci możemy zapisać, że przeciwdziedzina funkcji prawdopodobieństwa *a posteriori* jest proporcjonalna do przeciwdziedziny funkcji hipotezy oraz prawdopodobieństwa *a priori*. Ogólna postać w ten sposób sformułowanego twierdzenia mogłaby mieć następującą formę:

$$P(a \text{ posteriori}) \propto P(\text{wiarygodność hipotezy}) * P(a \text{ priori})$$

Różne dostępne miary funkcji gęstości prawdopodobieństwa *a posteriori* są bardzo użyteczne w analizie bayesowskiej, ponieważ mogą stanowić bezpośrednie oszacowanie wartości parametrów rozwiązania odwrotnego w ECT. W związku z tym, interesujące mia-

ry mogą być wyliczone jako oczekiwane wartości poprzez zastosowanie całkowania funkcji *a posteriori*. Niestety całkowanie to bywa źródłem podstawowych trudności w statystyce Bayesa; w większości zastosowań rozwiązanie analityczne nie jest możliwe [32]. Alternatywne podejścia stosują przybliżenia numeryczne, które cechuje niedokładność, przybliżenia analityczne, jak np. przybliżenie Laplace'a [41], które sprawdza się w niektórych zastosowaniach, oraz całkowanie Monte Carlo [42], w tym w szczególności metody Monte Carlo łańcuchów Markowa (MCMC), jak pokazano w [43]. MCMC oznacza zasadniczo całkowanie Monte Carlo przy użyciu łańcuchów Markowa. W statystyce Bayesa, mając dane, szukana jest całka po rozkładzie *a posteriori* parametrów modelu. Całkowanie metodą Monte Carlo polega na wyborze próbki z żadanego rozkładu, a następnie sformułowaniu średniej próbek, w celu aproksymacji wartości oczekiwanych. Natomiast metody Monte Carlo łańcuchów Markowa wyróżniają się tym, że wybierają wspomniane próbki przy wykorzystaniu odpowiednio skonstruowanego łańcucha Markowa, prowadząc próbkowanie niejako „wzdłuż” tego łańcucha przez odpowiednio długi okres. Jeśli chodzi o praktykę, to dziedziną, dla której zdefiniowana jest funkcja, bywa zwykle przedział czasowy lub ewentualnie obszar przestrzeni. Dla tego drugiego przypadku używana jest najczęściej nazwa losowego pola Markowa (Mrf). Metody bazujące na Mrf wykorzystywane są często w przetwarzaniu i analizie obrazów, m.in. do analizy tekstur [69]. Bardziej szczegółowe podstawy koncepcji łańcuchów Markowa można znaleźć w [66], a ogólną teorię w [49, 50]. Opisanie rozważania są stosunkowo ogólne, ale jeżeli dodatkowo wiadomo, że X stanowią dyskretne zmienne losowe, to operacje całkowania zastępuje operacja sumowania. Istnieją różne sposoby konstruowania tego typu łańcuchów, ale generalnie wszystkie odmiany, w tym m.in. algorytm *heat bath*, nazywany inaczej próbkowaniem Gibbsa [31], są specjalnymi przypadkami ogólnego algorytmu Metropolisa [44].

Na potrzeby ogólnych rozważań łańcuch Markowa możemy zdefiniować, jako szereg zmiennych losowych $\{X_0, X_1, X_2, \dots\}$, taki, że w każdej chwili czasu $t \geq 0$, następny stan X_{t+1} jest próbkowany z rozkładu $P(X_{t+1}|X_t)$, który jest zależny jedynie od bieżącego stanu łańcucha, X_t . Innymi słowy, mając dany stan X_1 , kolejny stan X_{t+1} jest zależny tylko i wyłącznie od bieżącego stanu X_t , a nie zależy od historii łańcucha $\{X_0, X_1, \dots, X_{t-1}\}$ [70].

2.1. Metody Monte Carlo łańcuchów Markowa

Całkowanie Monte Carlo wylicza daną miarę E funkcji f wektora zmiennych losowych X poprzez próbkowanie $\{X_t, t = 1, \dots, n\}$ rozkładu *a posteriori*, który możemy oznaczyć jako $\pi(\cdot)$, aproksymując go następnie zgodnie ze wzorem (1):

$$E[f(X)] \approx \frac{1}{n} \sum_{t=1}^n f(X_t) \quad (1)$$

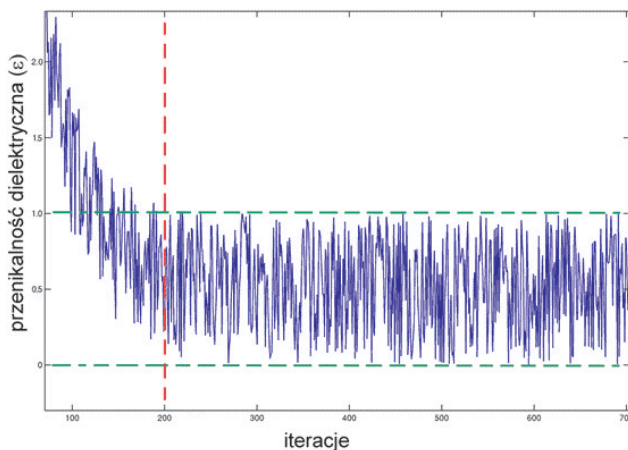
Zgodnie z powyższą zależnością, średnia populacji $f(X)$ jest oszacowana za pomocą średniej próbek. Jeżeli próbki $\{X_t\}$ są niezależne, to prawo wielkich liczb zapewnia, że aproksymacja może zostać dokonana, z wymaganą precyzją, poprzez odpowiednie zwią-

szenie rozmiaru zbioru próbek n [45]. Warto w tym miejscu zauważyć, że n jest pod kontrolą przeprowadzającego analizę; innymi słowy nie jest to stały, ustalony rozmiar zbioru próbek. Generalnie losowanie próbek z rozkładu *a posteriori* w taki sposób, aby były od siebie niezależne, nie zawsze jest możliwe, ponieważ rozkład ten może być niestandardowy [46]. Jednakże $\{X_t\}$ nie muszą koniecznie być niezależne. W literaturze znaleźć można opisy, z których wynika, że $\{X_t\}$ będą akceptowalne jeśli będą generowane przez jakikolwiek proces próbkujący rozkład *a posteriori* w odpowiednich proporcjach [47]. Proporcje te natomiast można zapewnić przy użyciu łańcucha Markowa, posiadającego rozkład *a posteriori* w formie rozkładu stacjonarnego [48].

2.2. Punkt startowy i przejściowy okres początkowy

Biorąc pod uwagę, że estymacja metodami MCMC polega zwykle na zebraniu określonej liczby próbek, po czym zastosowaniu do tak zebranego zbioru pewnej miary (np. średniej), to naturalnym sposobem implementacji takich założeń są numeryczne algorytmy iteracyjne. W tym miejscu powstaje naturalne pytanie: jaki wpływ na X_t ma stan początkowy X_0 ? Otóż stosownie do kryteriów regularyzacji [33], łańcuch stopniowo „zapomni” wartość stanu początkowego i w końcu osiągnie zbieżność z unikalnym rozkładem stacjonarnym $\varphi(\cdot)$, który nie jest zależny od t czy od X_0 . Innymi słowy, wraz ze wzrostem wartości t , punkty próbkowania $\{X_t\}$ będą próbkami coraz bardziej związanymi z rozkładem stacjonarnym.

Proces ten jest pokazany na rysunkach 1 oraz 2 na przykładzie szacowania wartości przenikalności dielektrycznej dla rozwiązania problemu odwrotnego pomiarów procesu transportu pneumatycznego. Poziome (zielone) linie kreskowane ograniczają wartości rozkładu stacjonarnego $\varphi(\cdot)$, pionowa (czerwona) linia kreskowana oznacza zakończenie przejściowego okresu początkowego łańcucha Markowa.



Rys. 1. Schemat procesu zbieżności algorytmu MCMC [70]

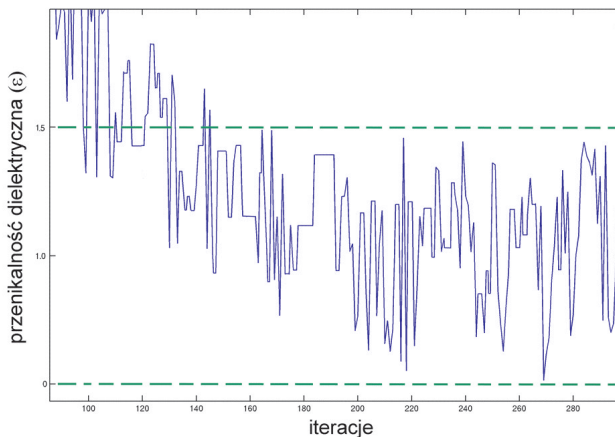
Dlatego też, po odpowiednim początkowym okresie przejściowym (*burn-in period*) trwającym m iteracji, punkty $\{X_t; t = m+1, \dots, n\}$ będą zależnymi punktami z rozkładu stacjonarnego $\varphi(\cdot)$. Próbkę z okresu przejściowego są zwykle odrzucane, tzn. nie bierze się ich pod uwagę przy obliczeniach miar rozkładu oczekiwanego. Poniższy wzór (2) ilustruje konstrukcję estymatorów, czyli wskazuje, jak można użyć łańcucha Markowa do estymacji $E[f(X)]$, gdy ten osiągnie swój rozkład stacjonarny $\varphi(\cdot)$:

$$\bar{f} = \frac{1}{n-m} \sum_{t=m+1}^n f(X_t) \quad (2)$$

Taka postać nosi miano średniej ergodycznej. Zbieżność do wymaganego rozkładu oczekiwanego zapewniona jest przez twierdzenie ergodyczne [49].

2.3. Algorytm Metropolis–Hastings

Ostatecznie, aby zastosować przedstawiony sposób rozumowania do rozwiązania danego problemu, należy się upewnić, że skonstruowany łańcuch Markowa posiada swój rozkład stacjonarny tożsamy z rozkładem poszukiwanym, czyli $\pi(\cdot)$ – rozkładem *a posteriori*. W tym miejscu można użyć generalizacji algorytmu Metropolisa wprowadzonej przez Hastingsa [42]. Otóż w algorytmie Metropolis–Hastings dla każdej chwili czasowej t , następny stan X_{t+1} , wybierany jest przez próbkowanie punktu – kandydata Y wybieranego z rozkładu proponowanego $q(\cdot|X_t)$. Warto w tym miejscu zauważyć, że rozkład proponowany może być zależny od stanu bieżącego X_t . Następnie proponowany punkt zostaje zaakceptowany lub odrzucony zgodnie z odpowiednim prawdopodobieństwem. Jeżeli proponowana wartość zostanie zaakceptowana, to następny stan będzie równy: $X_{t+1} = Y$. Jeżeli natomiast kandydujący punkt jest odrzucony, to łańcuch pozostaje w tym samym punkcie, tzn. $X_{t+1} = X_t$.



Rys. 2. Schemat procesu tzw. *mieszania* łańcucha algorytmu MCMC [70]

Rysunek 2 ilustruje łańcuch, który wielokrotnie pozostawał w tym samym punkcie. Poziome segmenty oznaczają momenty, w których wartość proponowana została odrzucona i łańcuch nie zmienił wartości przez kilka iteracji

Zgodnie z powyższym opisem, najprostsza postać algorytmu Metropolis–Hastings mogłaby na etapie implementacji przyjąć postać, którą w formie pseudokodu przedstawia wycinek zamieszczony na (rys. 3).

```
Inicjalizacja  $X_0$ ;
Przypisanie  $t=0$ ;
Powtórz {
  Próbuj w punkcie  $Y$  z rozkładu  $q(\cdot|X_t)$ ;
  Próbuj z rozkładu Jednorodnego  $(0,1)$ 
  zmienna losowa  $U$ ;
  Jeżeli  $U = a(X_t, Y)$ 
    przypisz  $X_{t+1}=Y$ ;
    w innym przypadku przypisz:  $X_{t+1}= X_t$ ;
  Zwiększ  $t$ ;
}
```

Rys. 3. Schemat algorytmu Metropolis–Hastings w postaci wycinka pseudokodu [70]

Warto w tym miejscu odnotować, że proponowany rozkład $q(\cdot|.)$ może mieć dowolną postać, a rozkład stacjonarny łańcucha będzie w formie $\pi(\cdot)$.

3. Implementacja

Z algorytmami MCMC związanych jest szereg problemów implementacyjnych. Najważniejsze zagadnienia dotyczą:

- szybkości zbieżności, doboru losowego kroku algorytmu, blokowania, kolejności uaktualnienia składników,
- liczby łańcuchów,
- doboru wartości początkowych,
- długości okresu przejściowego i momentu zakończenia,
- specjalnych odmian algorytmu (próbkiwanie niezależne, algorytmy jedno- oraz wieloskładnikowe, próbkiwanie Gibbsa, itp.),
- analizy wyników.

Podstawowym problemem związanym częściowo z implementacją, a częściowo będącym problemem sprzętowym, jest wysoka iteracyjność algorytmów MCMC. Zebranie od-

powiedniej ilości próbek, w zależności od danego zastosowania powiązanej jeszcze z dodatkowymi obliczeniami dla każdej z nich, wymaga przeprowadzenia obliczeń w pętłach iteracyjnych, często powtarzanych nawet dziesiątki tysięcy razy [70]. Istnieje jednak szereg metod przyspieszania wykonywania obliczeń, szczególnie poprzez zastosowanie dedykowanej implementacji sprzętowej [9]. Metody te były już stosowane dla niektórych zastosowań metod Monte Carlo [13]. Należy jednak zwrócić uwagę, że przed próbą implementacji sprzętowej, pożądane jest wcześniejsze zoptymalizowanie algorytmu dla konkretnego zastosowania.

3.1. Formy kanoniczne rozkładu proponowanego

Porównując różne hipotetyczne łańcuchy, można wyciągnąć wniosek, że pomimo dyskutowanego już powyżej faktu, iż dowolny proponowany rozkład dostarczy ostatecznie próbek zżądanego rozkładu $\hat{d}(\cdot)$, to jednakże szybkość zbieżności z rozkładem stacjonarnym zależeć będzie głównie od relacji pomiędzy $q(\cdot)$ oraz $\pi(\cdot)$. Co więcej, osiągnięcie zbieżności nie gwarantuje tzw. dobrego mieszania (*to mix*) łańcucha, tzn. nie zapobiega powolnemu poruszaniu się łańcucha wokół rozkładu stacjonarnego. Sytuację tę dobrze ilustrują rysunki 1 oraz 2. Na rysunku 1 można zaobserwować, jak łańcuch szybko osiąga zbieżność, nawet pomimo dosyć przypadkowej i odległej od $\pi(\cdot)$ wartości początkowej. Dodatkowo widać, że po osiągnięciu zbieżności, łańcuch miesza się równie szybko. Odmianą sytuację zilustrowano na rysunku 2, gdzie można zauważyć powolne poruszanie się łańcucha. W tym przypadku nawet pomimo osiągnięcia zbieżności z rozkładem stacjonarnym, łańcuch często pozostaje przy poprzednio ustalonej wartości. Skutkiem powolnego poruszania się łańcucha jest konieczność dłuższego działania w celu otrzymania podobnie wiarygodnej estymacji, co w przypadku zamieszczonym na rysunku 1.

Algorytm Metropolisa, stanowiący najprostszy przypadek w swojej formie kanonicznej, rozpatruje jedynie symetryczne rozkłady proponowane postaci: $q(Y|X) = q(X|Y)$ dla wszystkich X i Y . W związku z tym bywają przypadki, gdy wygodnie jest wybrać rozkład proponowany, który będzie generował każdy z komponentów Y warunkowo niezależnie mając dane X_t (np. dla X ciągłego, $q(\cdot|X)$ może stanowić rozkład normalny wielu zmiennych ze stałą macierzą kowariancji Σ). Dzięki temu uzyskujemy uproszczoną postać procedury wyliczania prawdopodobieństwa akceptacji i co za tym idzie uproszczenie algorytmu [31]. Specyficzną odmianą tego algorytmu jest algorytm Metropolisa kroku losowego (*random-walk*), gdzie $q(Y|X) = q(|X - Y|)$. Wyniki przedstawione na rysunku 1 zostały otrzymane właśnie przy użyciu danych wygenerowanych przez algorytm kroku losowego. Dobór skali rozkładu proponowanego (np. Σ) powinien być nieprzypadkowy [29]. Ostrożnie dobrany rozkład dający małe kroki $Y - X_t$ daje w efekcie wysoki stopień akceptacji, ale mimo to będzie się cechował słabym mieszaniem [26]. Z drugiej strony, dobierając rozkład dający duże kroki, można spowodować niski stopień akceptacji, co skutkować będzie częstymi przestojami łańcucha, jak to pokazano na rysunku 2. Idealne skalowanie rozkładu proponowanego powinno pozwolić uniknąć obydwu z opisanych przypadków ekstremalnych [30].

Kolejną odmianą Metropolisa–Hastingsa jest próbkowanie niezależne, w którym rozkład proponowany nie zależy od X , czyli $q(Y|X) = q(Y)$. Ogólnie w literaturze istnieją opisy, z których wynika, że takie algorytmy działają albo perfekcyjnie, albo bardzo słabo [50]. Próbkowanie niezależne do dobrego działania wymaga, żeby $q(\cdot)$ było dobrą aproksymacją $\pi(\cdot)$ [51].

Jak pokazano w [28], zamiast aktualizacji całego wektora zmiennych losowych X jednocześnie, często wygodniej jest podzielić X na komponenty $\{X_0, X_1, \dots, X_k\}$ o możliwie różnych rozmiarach, a następnie kolejno aktualizować poszczególne z nich, co nazywane jest algorytmem pojedynczego składnika. Dozwolone jest także aktualizowanie więcej niż jednego składnika [15]. Tego typu strategię zwiększając liczbę iteracji, ale w zależności od zastosowania często bywa tak, że procedura wyliczenia prawdopodobieństwa akceptacji upraszcza się znacząco, co może stanowić zaletę obliczeniową równoważącą dodatkowe iteracje szczególnie, gdy w przypadku aktualizacji X *en bloc*, stopień akceptacji możliwy do osiągnięcia jest znikomy [34]. Specjalnym przypadkiem tego typu algorytmu jest próbkowanie Gibbsa, która to nazwa została wprowadzona do statystyki w latach osiemdziesiątych dla tzw. algorytmu *heat-bath*, znanego z fizyki teoretycznej. Ten typ algorytmu jest jak dotąd najszerzej używany w zastosowaniach MCMC [67]. Unikalną cechą próbkowania Gibbsa jest to, że rozkład proponowany, przy aktualizacji i -tego składnika X stanowi (równanie (3)):

$$E[f(X)] \approx \frac{1}{n} \sum_{t=1}^n f(X_t) \quad (3)$$

z którego wynika, że prawdopodobieństwo akceptacji jest równe 1 [54]. Oznacza to naturalnie, że w próbkowaniu Gibbsa próbki kandydujące są zawsze akceptowane! W równaniu (3) wielkość po prawej stronie to pełen rozkład warunkowy, więc można stwierdzić, że ten typ algorytmu polega na próbkowaniu z takich właśnie rozkładów [27].

3.2. Łączenie komponentów, porządek aktualizacji oraz liczba łańcuchów

W opisanych powyżej algorytmach możliwe jest także łączenie składników szczególnie, gdy są one silnie skorelowane wzajemnie w przestrzeni rozkładu stacjonarnego $\pi(\cdot)$. W takim przypadku blokowanie komponentów w jeden wyższego rzędu może usprawnić proces mieszania, zależnie od rozkładu proponowanego, choć dane literaturowe na ten temat są sprzeczne [47].

W dyskutowanych wcześniej metodach pojedynczego składnika oraz próbkowaniu Gibbsa zwykle aktualizacje X_t dokonywane są w porządku z góry ustalonym, jakkolwiek nie jest to wymagane. Dozwolone są także losowe permutacje kolejności aktualizacji. Co więcej, nie jest wymagane także aktualizowanie wszystkich składników w każdej iteracji. Przykładowo, można aktualizować jeden ze składników w każdej iteracji, wybierając komponent i z pewnym prawdopodobieństwem $s(i)$. Sugeruje się także częstszą aktualizację bardziej ze sobą skorelowanych składników w celu poprawy mieszania łańcucha [70].

Kolejną dopuszczalną modyfikacją jest uruchomienie więcej niż jednego łańcucha Markowa. Wskazania literaturowe w tym przypadku niestety nie są zgodne. W różnych źródłach sugerowane są różne rozwiązania: od wielu krótkich łańcuchów, poprzez kilka długich, aż do jednego bardzo długiego [57]. Ogólna konkluzja, jaka ustaliła się na przestrzeni lat, prowadzi do tego, że raczej odrzuca się ideę wielu krótkich łańcuchów jako błędną. Natomiast debata pomiędzy zwolennikami drugiego i trzeciego rozwiązania nie jest zakończona. Część badaczy uważa, że porównanie kilku równoległych łańcuchów nie zawsze ma sens [42], a inni wręcz przeciwnie [35]. Niektórzy twierdzą nawet, że to dobry test na stwierdzenie, czy łańcuch osiągnął już rozkład stacjonarny [37]. Generalnie, można skłonić się do stwierdzenia, że dysponując kilkoma procesorami uruchomienie po jednym łańcuchu na każdym z nich może być korzystne.

3.3. Wartości początkowe i początkowy okres przejściowy

Jeżeli łańcuch jest nieredukowalny [51], to wybór konkretnego punktu startowego X_0 nie ma wpływu na rozkład stacjonarny. Szybko mieszające się łańcuchy, jak np. ten z rysunku 1, szybko też odnajdą drogę nawet startując od wartości ekstremalnych. Dla wolniej mieszających się łańcuchów zaleca się bardziej uważny dobór wartości początkowych, by uniknąć zbyt długiego początkowego okresu przejściowego. Generalnie w przypadku aktualizacji wielu wartości i/lub, przy ewentualnym wykonywaniu dodatkowych operacji w poszczególnych iteracjach, rozsądne wydaje się dobieranie, o ile to możliwe, wartości początkowych bliższych oczekiwanemu rozkładowi. Czasami jednak sugeruje się stosowanie wartości początkowych nawet skrajnie rozproszonych, odmiennych od oczekiwanych, uruchamiając wiele łańcuchów równoległych, by ocenić i porównać ich zbieżność [55, 56].

Długość przejściowego okresu początkowego m zależy od:

- X_0 , szybkości zbieżności $P^{(t)}(X_t|X_0)$ do $\pi(X_t)$, oraz od
- wymaganego stopnia podobieństwa pomiędzy $P^{(t)}(\cdot|\cdot)$ oraz $\pi(X_t)$.

Teoria mówi, że jeśli tylko zdefiniowane zostanie kryterium „wystarczającego podobieństwa“, to m może być wyznaczone analitycznie. Niestety, w większości przypadków tego typu obliczenia nie są możliwe praktycznie [33]. W związku z tym najczęstszą metodą wyznaczania okresu początkowego jest graficzna interpretacja tzw. wykresów wyjściowych Monte Carlo, czyli wykresów funkcji zmiennej losowej $\{X_t, t = 1, \dots, n\}$. Zaproponowano także inne metody wyznaczania m , które określane są mianem diagnostyki zbieżności. Stosują one zwykle rozmaite metody teoretyczne, uproszczenia i aproksymacje w odniesieniu do pojedynczych lub wielokrotnych łańcuchów, ale wszystkie w ten lub inny sposób opierają się na analizie danych wyjściowych Monte Carlo [58]. Grupa metod opierających się na monitorowaniu dowolnej funkcji $f(X_t)$ wyjściowej jest łatwa do zastosowania, jakkolwiek może prowadzić do błędnych interpretacji, ponieważ $f(X_t)$ może wydawać się zbieżna z rozkładem w danej iteracji m , podczas gdy inna niemonitorowana funkcja $g(X_t)$ może jeszcze nie osiągnąć zbieżności [56]. Jakiegokolwiek funkcji $f(\cdot)$ nie poddano by

monitorowaniu, to mogą istnieć inne, które zachowują się w odmienny sposób [61]. Kolejna grupa opiera się na porównywaniu globalnym rozkładu $\pi(\cdot)$ z pełnym rozkładem łącznym $P^{(t)}(\cdot)$. W celu uniknięcia potrzeby zajmowania się bezpośrednio $P^{(t)}(\cdot)$, wypracowano metody uzyskiwania z niego próbek poprzez uruchomienie łańcuchów równoległych i użycia macierzy przejścia $P(\cdot)$ [62, 63]. Jednakże okazuje się, że dla zachowania stabilności proceduralnej, należałoby uruchamiać bardzo dużą liczbę łańcuchów równoległych, co w przypadku, gdy zbieżność jest powolna, może stanowić poważne ograniczenie praktyczne [48]. Uruchamianie równoległych łańcuchów zwiększa w oczywisty sposób obciążenie obliczeniowe, ale może być bardzo użyteczne, choćby w celu diagnozy wolnej zbieżności. Gelman i Rubin [55] wskazali, jak kilka równoległych łańcuchów, które indywidualnie wydawały się osiągnąć zbieżność, w trakcie porównania ujawniły znaczące różnice pozornie stacjonarnych rozkładów. Z praktycznego punktu widzenia, metody bazujące na danych wyjściowych MC są najbardziej wygodne, pozwalając na ocenę zbieżności bez konieczności uciekania się do analizy przejściowej $P(\cdot)$. Istnieje także podejście sugerujące brak konieczności wyznaczania długości początkowego okresu przejściowego. W tym podejściu zakłada się, że okres ten ma najprawdopodobniej mniej niż 1% całkowitej długości łańcucha n prowadzonego wystarczająco długo, by uzyskać zadowalającą precyzję dla estymatorów ze wzoru (2) [57]; jakkolwiek odrzucanie nawet 2% długości uważa się za wskazane, szczególnie przy uruchamianiu łańcucha od wartości ekstremalnych.

3.4. Wyznaczanie długości – przerwanie łańcucha

Ważnym zagadnieniem praktycznym jest także decyzja, kiedy przerwać łańcuch. Jak wspomniano wcześniej, celem łańcucha jest działanie wystarczająco długie, by uzyskać odpowiednią precyzję estymacji (2). Szacowanie wariancji estymatorów \bar{f} (wariancji Monte Carlo) jest utrudnione z powodu braku zależności w iteracjach $\{X_t\}$. Najbardziej oczywistą, nieformalną, metodą wyznaczania długości n jest uruchomienie kilku łańcuchów równoległe, z różnymi wartościami startowymi, a następnie porównanie estymacji \bar{f} . Jeżeli wartości te nie są ze sobą wystarczająco zgodne, n musi zostać zwiększone. Istnieją także bardziej formalne metody wyznaczania długości łańcucha [64]. Zaproponowano je dla różnych szczególnych przypadków i zapewniają zadowalające oszacowanie, np. niektórych miar rozkładu *a posteriori* [60, 65], lub wektora wybranych parametrów w tym rozkładzie za pomocą szeregów binarnych [59]. Inną formalną metodą wyznaczania długości jest porównanie modeli łańcuchów Markowa kolejnych rzędów pomocniczo obliczając stosunek wiarygodności hipotezy w celu porównania tzw. kryterium BIC (*Bayesian Informarmation Criterion*) obu modeli [38, 39].

3.5. Analiza danych wyjściowych

W ramach wnioskowania Bayesa, rozkład *a posteriori* $\hat{d}(\cdot)$ jest często rozpatrywany przy pomocy takich narzędzi, jak: średnie, odchylenia standardowe, korelacja, przedziały pewności czy rozkłady brzegowe komponentów $\{X_t\}$. Średnie, odchylenia standardowe

i korelacje mogą być estymowane przy pomocy ich próbkowych odpowiedników z danych wyjściowych Monte Carlo $\{X_{ti}, t = m+1, \dots, n\}$ zgodnie z równaniem (2). Przykładowo średnia brzegowa oraz wariancja X_{ti} mogą być obliczane za pomocą wzorów (4) i (5).

$$\overline{X}_i = \frac{1}{n-m} \sum_{t=m+1}^n X_{ti} \quad (4)$$

$$S_i^2 = \frac{1}{n-m-1} \sum_{t=m+1}^n (X_{ti} - \overline{X}_i)^2 \quad (5)$$

Warto podkreślić, że estymaty \overline{X}_i i S_i^2 ignorują pozostałe składniki danych wyjściowych Monte Carlo.

Bardzo interesującą miarą, jaka jest do dyspozycji w ramach statystyki Bayesa i MCMC, są przedziały ufności [40]. Przedział $[c_p, c_{1-p}]$ ufności w stopniu $100(1-2p)\%$ dla komponentu skalarne X_i może być oszacowany poprzez ustalenie c_p równego kwantylowi p -tego rzędu dla $\{X_{ti}, t = m+1, \dots, n\}$, oraz c_{1-p} równego kwantylowi $(1-p)$ -tego rzędu. Przykładowe obliczenia przedziałów ufności dla dwóch i więcej wymiarów można znaleźć w [52]. Przedziały ufności są o tyle istotne z punktu widzenia zastosowania dla ECT, że prawdopodobnie pozwoliłyby na dokładniejsze obliczenia w przypadku przetwarzania tomogramów w celu np. szacowania wartości strumienia masy, czy określenia reżimu przepływu, niż w przypadku zastosowania klasycznych metod rekonstrukcji i przetwarzania obrazów.

4. Wybór i walidacja modeli Bayesa

Kluczowym problemem, jaki pojawia się w zastosowaniach praktycznych statystyki Bayesa i metod MCMC, jest zaprojektowanie właściwego modelu. Sam proces modelowania można podzielić na dwa zagadnienia. Pierwsze z nich to wybór modelu, a drugie to jego sprawdzenie, czy raczej ocena. W praktyce, poza bardzo rzadkimi przypadkami, specyfikacja modelu nigdy nie bywa idealnie poprawna [23]. Dlatego też nasuwają się pytania o adekwatność rozważanego modelu, w związku z tym należy podjąć decyzję, który z rozważanych modeli jest najlepszy.

Wszystkie złożone modele, cytowane w niniejszym artykule, mogą być postrzegane jako specyfikacja rozkładu łącznego danych, czyli obserwacji, oraz danych nieobserwowalnych (np. parametry modelu czy brakujące dane). Często modele tego typu mogą być przedstawione w postaci kierunkowych grafów acyklicznych [22], dla których możliwe jest określenie wspomnianych gęstości łącznych. W innych przypadkach, np. w modelowaniu przestrzennym [34], rzeczony rozkład musi być unikalnie wyznaczony na podstawie odpowiedniej charakterystyki lokalnej. Bez względu jednak na rodzaj modelowania, to właśnie ten rozkład łączny powinien zostać sprawdzony pod kątem zgodności wyznaczonego

modelu z rozpatrywanym zastosowaniem fizycznym metody. Wnioskowanie bayesowskie bazuje na cechach rozkładu warunkowego parametrów nieobserwowalnych modelu pod warunkiem znajomości danych obserwacyjnych. Z drugiej strony sprawność modelu jest oceniana przez pryzmat cech rozkładu marginalnego danych obserwowalnych, a w szczególności przez porównanie tego, co model przewiduje z tym, co zaobserwowano. Jednym z tego typu porównań są wykresy Q-Q (kwantyl-kwantyl), czyli graficzna technika porównania 2 zbiorów danych, którego wynikiem jest odpowiedź na pytanie czy dane pochodzą z populacji mogących mieć wspólny rozkład. Test dobrego dopasowania (*goodness-of-fit*) pozwala agregować wzmiankowane porównania, dając w efekcie globalną miarę rozbieżności [20]. Podobne są miary odchyłkowe (potocznie „obserwacje minus oczekiwania“), średniokwadratowe, odchyłki, miara Mallowsa, statystyka PRESS, itp. [21].

Natomiast klasyczne metody wyboru spośród możliwych modeli uznawane są w literaturze za niezadowalające i pozostawiają nadal pole do dalszego rozwoju. Zaliczają się do nich zarówno testy wiarygodności hipotezy, włączające aproksymacje asymptotyczne, czyli różnice odchyłek, które mogą być wysoce błędne dla małych wielkości próbek [18], jak i te dla niezagnieżdżonych modeli, które istnieją tylko w szczególnych, dobrze zdefiniowanych przypadkach [19]. Kolejne, nowsze metody, takie jak m.in. zastosowanie czynnika Bayesa wraz z kalibracją [17] są bardziej uznane, chociaż także mocno krytkowane [16] jako: trudne (często niemożliwe do praktycznej implementacji), obciążone paradoksem Lindleya i ogólnie niepraktyczne. Istnieją także modyfikacje wspomnianych metod, które próbowano zastosować w wybranych przypadkach. Możliwe jest także wspomaganie wyboru modelu metodami MCMC [10]. Zagadnieniem poruszonym w literaturze w oderwaniu od wyboru samego modelu, istotnym dla praktyków przy tworzeniu modeli, jest decyzja co do wyboru zmiennych, potrzebnych do ujęcia w modelu. Pośrednio mogą w tym pomóc także klasyczne metody wymienione w bieżącym akapicie. Dodatkowo istnieje też, zyskująca popularność, metoda stochastycznego wyszukiwania selekcji zmiennych SSVS (*stochastic search variable selection*) [12].

5. Zastosowania modelowania Bayesa i metod MCMC

Nie sposób wymieniść wszystkich zastosowań statystyki Bayesa w połączeniu z MCMC. Zapytania wpisane w wyszukiwarce Google złożone z kombinacji słów: *bayes*, *mcmc*, *application*, *review*, zależnie od składu, porządku i formy poszczególnych wyrazów, zwracają do ponad 2 milionów wyników (sprawdzono: V 2009 r.). W związku z tym, można tutaj przytoczyć jedynie wybiórczo kilka przykładów.

Najbardziej medialną w ostatnich latach dziedziną wykorzystującą opisywane metody jest genetyka, gdzie stosowane są one zarówno przy dekodowaniu ludzkiego genomu, jak i w wielu innych aspektach – wyczerpujące odniesienia do tego tematu można znaleźć w [11]. Ogólnie w medycynie metody te wykorzystuje się od wielu lat zarówno do przetwarzania danych medycznych, np. w epidemiologii, czy opracowaniu profilaktyki, jak i do

przetwarzania obrazów [68, 69]. Ciekawymi dziedzinami, w których już dawno znalazły zastosowanie metody MCMC, jest inżynieria kosmiczna, astrofizyka i astronomia [10], a także takie dziedziny jak: meteorologia, archeologia, e-learning, sztuczna inteligencja, socjologia, ekonomia i inne; więcej przykładów można znaleźć na stronach cyklicznych konferencji „Uncertainty in Artificial Intelligence” (2006, 2007, 2008). Podjęto także próby zastosowania modelowania przestrzennego w tomografii, głównie medycznej [5] ale i procesowej, również z uwzględnieniem modelowania czasowego [25].

6. Podsumowanie

Jak wynika z przeglądu metod MCMC, zamieszczonego w niniejszym artykule, koncepcja ta jest stosunkowo przejrzysta, biorąc pod uwagę jej potencjalnie duże możliwości. Metodologia MCMC idzie w parze ze statystyką Bayesa, chociaż walidacja modeli Bayesa jest nadal zagadnieniem trudnym i otwartym. Przez dziesiątki lat ubiegłego stulecia trwała debata nad zaletami i wadami podejścia Bayesa, przy czym główną wadą była od zawsze mała dostępność narzędzi – głównie mocy obliczeniowej. Metody te w większości polegają na zebraniu dużej liczby próbek potrzebnych do wyliczenia oszacowania szukanych wartości. W praktyce przekłada się to na implementację wielokrotnie wykonywanych pętli iteracyjnych. Od mniej więcej połowy lat dziewięćdziesiątych, wraz z upowszechnieniem wydajnych komputerów osobistych, stosowność wnioskowania Bayesa oraz metod MCMC rośnie i rozszerza się na kolejne dziedziny nauki i inżynierii, w tym także na tomografię, a w szczególności na ECT.

Podziękowania

Prace prowadzone przez A. Romanowskiego są współfinansowane ze środków Europejskiego Funduszu Społecznego w ramach programu „Innowacyjna dydaktyka bez ograniczeń – zintegrowany rozwój Politechniki Łódzkiej”.

Literatura

- [1] Berryman J.G., *Lecture notes on nonlinear inversion and tomography*. http://sepwww.stanford.edu/sep/berryman/NOTES/lecture_notes.html, 2001.
- [2] Engl H.W., Hanke M., Neubauer A., *Regularization of Inverse Problems*. Kluwer Academic Publishers, 1996.
- [3] Hadamard J., *Sur les problèmes aux dérivées partielles et leur signification physique*. Princeton University Bulletin, 1902, 49–52.
- [4] Twomey S., *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements*. Dover Publications, 1977.
- [5] West R.M., Aykroyd R.G., Meng S, Williams R.A., *MCMC techniques and spatial-temporal modelling for medical EIT*. *Physiol. Meas.*, 25, 2004, 181–194.

- [6] Romanowski A., Grudzień K., *Zestawienie Metod Rekonstrukcji Obrazów i Modelowania Przestrzennego 2D dla Tomografii Procesowej*. Automatyka (półrocznik AGH), vol. 9/3, 2005, 621–630.
- [7] Yang W.Q., Peng L., *Image reconstruction algorithms for electrical capacitance tomography*. Meas. Sci. Technol. 14 pp1–13. Romanowski A., Grudzień K., Williams R.A., Aykroyd R.G., West R.M., Meng S., (2004), *Podejście Bayesa w przemysłowych systemach kontroli i sterowania*, Seminarium „Przetwarzanie i analiza sygnałów w systemach wizji i sterowania” Słok k/Bełchatowa, Automatyka (półrocznik AGH), 8/3, 2003, 217–222.
- [8] Romanowski A., Grudzień K., Williams R.A., Aykroyd R.G., West R.M., Meng S., *Podejście Bayesa w przemysłowych systemach kontroli i sterowania*. Seminarium „Przetwarzanie i analiza sygnałów w systemach wizji i sterowania” Słok k/Bełchatowa, Automatyka (półrocznik AGH), 8/3, 2004, 217–222.
- [9] Wiatr K., *Sprzętowe implementacje algorytmów przetwarzania obrazów w systemach wizyjnych czasu rzeczywistego*. Wydawnictwa Naukowo-Dydaktyczne AGH, 2002.
- [10] Liddle A.R., *Information criteria for astrophysical model selection*. Monthly Notices of the Royal Astronomical Society: Letters, vol. 377, Nr 1, 2007, L74–L78(1).
- [11] Beaumont M.A., Rannala B., *The Bayesian revolution in genetics*. Nature Reviews Genetics 5, 2004, 251–261.
- [12] George E.I., McCulloch R.E., *Fast Bayes variable selection*. CSS technical report, University of Texas, Austin, 1994.
- [13] Zhang G., Leong P.H.W., Ho C.H., Tsoi K.H., Cheung C.C.C., Lee D., Cheung R.C.C., Luk W., *Reconfigurable acceleration for Monte Carlo based financial simulation*. In Proc. IEEE International Conference on Field-Programmable Technology (FPT), Singapore, 2005.
- [14] Gelfand A.E., Paliwal P., Abraham L., Barlow W., Elmore J., *Chronological Event Modeling For Screening Mammography*. Statistics in Medicine, 24, 2005.
- [15] Diebolt J., Robert C., *Estimation of finite mixture distributions through Bayesian sampling*. J.R. Statist. Soc. B, 56, 1994, 363–376.
- [16] Gelfand A.E., Discussion to: *Statistical Research: Some Advice for Beginners*. The American Statistician, 58, 2004, 197–199.
- [17] Pettit L.I., Young K.D.S., *Measuring the effect of observations on Bayes factors*. Biometrika, 77, 1990, 455–466.
- [18] Neyman J., Scott E.L., *Consistent estimates based on partially consistent observations*. Econometrica, 16, 1948, 1–32.
- [19] Cox D.R., *Tests of separate families of hypotheses*. In Proceedings of the 4th Berkeley Symposium, vol 1, Berkeley, University of California Press, 1961, 105–123.
- [20] D’Agostino R.B., Stephens M.A., *Goodness of Fit Techniques*. New York, Marcel Dekker, 1986.
- [21] McCullagh P., Nelder J. A., *Generalized Linear Models*. London, Chapman k. Hall, 1989.
- [22] Whittaker J., *Graphical Models in Applied Multivariate Statistics*. Chichester, Wiley, 1990.
- [23] Gelfand A.E., Sahu S., *Discussion to “On Model Expansion, Model Contraction. Identifiability and prior Information: Two Illustrative Scenarios Involving Mismeasured Variables*, Statistical Science, 20, 2005, 130–131.
- [24] Bedard M., *Weak convergence of Metropolis algorithms for non-IID target distributions*. Submitted to Ann. Appl. Prob., 2006.
- [25] Romanowski A., Grudzień K., *Modelowanie czasowe w systemach przepływu pneumatycznego z zastosowaniem łańcuchów Markowa i Metody Monte Carlo*. Seminarium „Przetwarzanie i analiza sygnałów w systemach wizji i sterowania” Słok k/Bełchatowa, Automatyka (półrocznik AGH), 8/3, 2004, 409–414.
- [26] Neal P., Roberts G.O., Yuen J., *Optimal Scaling of Random Walk Metropolis algorithms with Discontinuous target densities*. MIMS EPrint: 2007.91, University of Manchester, 2007.

- [27] Breyer L., Roberts G.O., *From Metropolis to diffusions: Gibbs states and optimal scaling*. Stochastic Process. Appl., 90, 2000, 181–206.
- [28] Neal P. J., Roberts G.O., *Optimal Scaling for partially updating MCMC algorithms*. Ann. Appl. Prob., 16, 2006, 475–515.
- [29] Roberts G.O., Rosenthal J.S., *Optimal scaling for various Metropolis-Hastings algorithms*. Statist. Science, 16, 2001, 351–367.
- [30] Roberts G.O., Gelman A., Gilks W.R., *Weak convergence and optimal scaling of Random walk Metropolis algorithms*. Ann. Appl. Prob., 7, 1997, 110–120.
- [31] Jennison C., *Discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo methods*. J.R. Statist. Soc. B, 55, 1993, 54–56.
- [32] Carlin B.P., Louis T.A., *Bayes and Empirical Bayes Methods for Data Analysis*. 2000, 419.
- [33] Priestley M.B., *Spectral Analysis and Time Series*. London, Academic Press, 1981.
- [34] Besag J., Green P.J., *Spatial statistics and Bayesian computation*. J. Roy. Statist. Soc. Ser. B 55, 1993, 25–37.
- [35] Fosdick L.D., *Monte Carlo calculations on the Ising lattice*. Meth. Comput. Phys., 1, 1963, 245–280.
- [36] Zellner A., Min C., *Gibbs sampler convergence criteria*. Technical report, Graduate School of Business, University of Chicago, 1994.
- [37] Cui L., Tanner M.A., Sinha D., Hall W.J., *Monitoring convergence of the Gibbs sampler: further experience with the Gibbs stopper*. Statist. Sci., 7, 1992, 483–486.
- [38] Cox D.R., Miller H.D., *The theory of Stochastic Processes*. London, Chapman & Hall, 1965.
- [39] Schwarz G., *Estimating the dimension of a model*. Ann. Statist., 6, 1978, 461–464.
- [40] Kiefer J., *Journal of the American Statistical Association*. 72, 1977, 789–827.
- [41] Kass R.E., Tierney L., Kadane J.B., *Asymptotics in Bayesian computation (with discussion)*. [w:] Bayesian Statistics 3, (eds J.M. Bernardo, M.H. DeGroot, D.V. Lindley, A.F.M. Smith), Oxford University Press, 1988, 261–278.
- [42] Hastings W.K., *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*. Biometrika, 57(1), 1970, 97–109.
- [43] Geman S., Geman D., *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*. IEEE Trans. Pattern Anal. Mach. Intel., 6, 1984, 721–741.
- [44] Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E., *Equations of state calculations by fast computing machines*. J. Chem. Phys., 21, 1953, 1087–1092.
- [45] Green P.J., *Discussion on representation of knowledge in complex systems* (by U. Grenander and M.I. Miller). J. R. Statist. Soc. B, 56, 1994a, 589–590.
- [46] Green P.J., *Reversible jump MCMC computation and Bayesian model determination*. Mathematics, Research Report, University of Bristol, 1994b.
- [47] Gilks W.R., Richardson S., Spiegelhalter D.J., *Markov Chain Monte Carlo in Practice*. London, Chapman & Hall, 1995, 75–88.
- [48] Liu C., Liu J., *Discussion on the meeting on the Gibbs sampler and other Markov chain Monte Carlo methods*. J.R. Statist. Soc. B, 55, 1993, 82–83.
- [49] Amit Y., *On rates of convergence of stochastic relaxation for Gaussian and non-Gaussian distributions*. J. Mult. Anal., 38, 1991, 82–99.
- [50] Tierney L., *Markov chains for exploring posterior distributions (with discussion)*. Ann. Statist., 22, 1994, 1701–1762.
- [51] Roberts G.O., *Markov chain concepts related to sampling algorithms*. [w:] Markov Chain Monte Carlo in Practice (eds W.R. Gilks, S. Richardson and D.J. Spiegelhalter). London, Chapman & Hall, 1995, 45–57.
- [52] Besag J.E., Green P., Higdon D., Mengersen K., *Bayesian computation and stochastic systems*. Statist. Science, 10, 1995, 3–66.

- [53] Zeger S.L., Karim M.R., *Generalized linear models with random effects: a Gibbs sampling approach*. J. Am. Statist. Ass., 86, 1991, 79–86.
- [54] Gelfand A.E., Smith A.F.M., *Sampling-based approaches to calculating marginal densities*. J. Am. Statist. Ass., 85, 1990, 398–409.
- [55] Gelman A., Rubin D.B., *A single series from the Gibbs sampler provides a false sense of security*. [w:] Bayesian Statistics (eds J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith), Oxford, Oxford University Press, 1992a, 625–631.
- [56] Gelman A., Rubin D.B., *Inference from iterative simulation using*. 1992b.
- [57] Geyer C.J., *Practical Markov chain Monte Carlo*. Statist. Sci., 7, 1992, 473–511.
- [58] Cowles M.K., Carlin B.P., *Markov chain Monte Carlo convergence diagnostics: a coraparative review*. Technical Report 94-008, Division of Biostatistics, School of Public Health, University of Minnesota, 1994.
- [59] Raftery A.E., Lewis S.M., *How many iterations in the Gibbs sampler?* [w:] Bayesian Statistics 4 (red. J.M. Bernardo, J.O. Berger, A.P. Dawid, A.F.M. Smith). Oxford, Oxford University Press, 1992a, 765–776.
- [60] Raftery A.E., Lewis S.M., *One long run with diagnostics: implementation strategies for Markov chain Monte Carlo*. Statist. Sci., 7, 1992b, 493–497.
- [61] Geweke J., *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. [w:] Bayesian Statistics 4 (red. J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith). Oxford, Oxford University Press, 1992, 169–193.
- [62] Ritter O., Tanner M.A., *Facilitating the Gibbs sampler: the Gibbs stopper and the griddy-Gibbs sampler*. J. Am. Statist. Ass., 87, 1992, 861–868.
- [63] Hills S.E., Smith A.F.M., *Parameterization issues in Bayesian inference*. [w:] Bayesian Statistics 4 (red. J.M. Bernardo, J. Berger, A.P. Dawid, A.F.M. Smith). Oxford, Oxford University Press, 1992, 227–246.
- [64] Muller P., *A generic approach to posterior integration and Gibbs sam-pling*. Technical Report, Institute of Statistics and Decision Sciences, Duke University, 1991.
- [65] Raftery A.E., Banfield J.D., *Stopping the Gibbs sampler, the use of morphology, and other issues in spatial statistics*. Ann. Inst. Statist. Math., 43, 1991, 32–43.
- [66] Grimmett G.R., Stirzaker D.R., *Probability and Random Processes*. 2nd edn. Oxford, Oxford University Press, 1992.
- [67] Liu J.S., *Monte Carlo Strategies in Scientific Computing*. Mathematics, 2001, 360.
- [68] Winkler G., *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods*. A Mathematical Introduction, Springer Series: Stochastic Modelling and Applied Probability, vol. 27, 2nd ed., XVI, 2006, 389.
- [69] Li S.Z., *Markov Random Field Modeling in Image Analysis*. Springer Series: Computer Science Workbench, 2nd ed., XIX, 2001, 323.
- [70] Romanowski A., *Algorytmy modelowania przestrzennego i czasowego dla potrzeb przetwarzania danych w systemach procesowej tomografii pojemnościowej*. Katedra Informatyki Stosowanej Politechniki Łódzkiej, 2008 (Praca doktorska).