



---

# *Journal of Statistical Software*

October 2011, Volume 43, Book Review 4.

<http://www.jstatsoft.org/>

---

Reviewer: Juana Sanchez  
University of California, Los Angeles

---

## **Hidden Markov Models for Time Series – An Introduction Using R**

Walter Zucchini and Iain L. MacDonald  
Chapman & Hall/CRC, Boca Raton, FL, 2009.  
ISBN 978-1-58488-573-3. 288 pp. USD 82.95.  
<http://134.76.173.220/hmm-with-r/>

---

Hidden Markov models (HMMs) are mixture models in which the mixture component generating an observation is determined by the state of a hidden Markov process instead of a static mixing distribution. The marginal distribution of an HMM is a mixture distribution. The components of the mixture can be some of the known probability distributions, and the Markov process can be first or higher order.

The book by Walter Zucchini and Iain MacDonald is an introduction to the theory and computation of HMM (Part I) and their applications to a wide variety of time series data (Part II). Chapters 1–7 of Part I introduce the main ideas using a single running example, annual earthquake counts, and different versions of Poisson HMM; Chapter 8 generalizes to other HMM. Part II breaks that pattern and applies both Poisson and some of the HMM mentioned in Chapter 8 to very interesting data sets, namely epileptic episodes, eruptions of Old Faithful, *Drosophila* speed and change of direction, wind direction, financial series, births, homicides and suicides and animal behavior. This is not a book, as the authors say, written for those interested in developing the theory of HMM, but rather for applied statisticians in fields in which time series (or other dependent data) arise that are not amenable to analysis by ARMA models or models that assume independence or normality, respectively, especially categorical series and series of counts. Becoming acquainted with the theory of HMM with a single running example of earthquake data and one type of model, Poisson HMM, is very effective, even if the Poisson is not the model you need for your data. The study of the models is aided by the R code written by the authors and presented in Appendix A.

The book assumes knowledge of basic probability distributions, maximum likelihood estimation, basic time series and autocorrelation analysis, basic R and matrix algebra. Some familiarity with mixtures and Markov models will make the reading lighter, although Chapter 1 is devoted to introducing those concepts. The book would be a good text for a seminar or a course on HMM or for self-learning the topic. The applied, theoretical and computational exercises at the end of the chapters in Part I, the R code in Appendix A, the numerous references to alternative literature that may fill the gaps in the book, and the mention of alternative packages in R that do more general HMM, are all a good supplement for the two

learning scenarios. The reader can choose to get as deep into the theory, computation, and applications as he or she wants.

Part I of the book has eight chapters. Chapter 1 explains what a mixture model and a Markov chain is for those who do not already know that. Those already familiar with mixtures and Markov models may skip the chapter, but they would miss the introduction to the notation used throughout the book if they choose to do so.

Chapter 2 reiterates that an independent mixture model alone is not a good model to capture the dependence among observations typical of a time series. Dependence is introduced in the model by making the mixture component depend on the Markov chain. This is illustrated nicely in Figure 2.3. This chapter specifies the likelihood function for the Poisson HMM with complete and incomplete data. Chapter 3 discusses the issue of finding the maximum likelihood estimators (MLE) using optimization. Scaling the computations to prevent underflow and reparametrizing to avoid the computing problems caused by parameter constraints is recommended and the standard R numerical optimization functions `nlm` is used and illustrated with the R code of Appendix A and the earthquake data. The parametric bootstrap is recommended to estimate standard errors of the estimates. Chapter 4 describes how to find the MLE estimator with the EM algorithm. Chapter 5 discusses the important task of predicting the state and the value of the series. All the relevant conditional distributions are presented, and the earthquake forecast and state distributions resulting from three and four states Poisson HMM are illustrated nicely in Figures 1–4. Model checking by means of Akaike’s information criterion (AIC) and Bayesian information criterion (BIC) is covered in Chapter 6. In the context of HMM it is useful to compare the autocorrelation function (ACF) of the selected model and the ACF of the data, as well as the pseudoresiduals. The latter are the HMM version of residual checking in regression to determine how well the selected model fits the data and whether there are outliers. Why would the authors check for outliers after selecting the model is not explained in the book. Chapter 7 presents the estimation of HMM via Gibbs sampling, pointing out the label switching problem, and indicating the difficulties of using the more efficient reversible jump Markov chain Monte Carlo technique. Finally, the last chapter in Part I talks about other HMM specifications, such as HMM for categorical series, multinomial observations, multivariate series, series that depend on covariates, and other dependencies.

Part II of the book contains eight chapters of examples. In each chapter, estimated parameters of the mixture components and estimated transition probability matrix, as well as model diagnostics and ACF are discussed. The reader can replicate the analysis if the material in Part I has been understood and studied thoroughly. Different from Part I is the fact that the Markov models, in some cases, are of order higher than 1 (e.g., wind direction), the components of the mixture are normal (e.g., old Faithful duration model), the series are multivariate (e.g., *Drosophila* speed and change, financial time series) and the research questions vary according to the data studied. The reader will have no difficulty identifying applications where the same research questions may arise.

The R code in Appendix A is not given in a ready-form package, but rather as snippets of code to illustrate to the readers how to make their own codes for application-specific variations of the main models. The reader can replicate many of the analyses done in the book with this code.

Overall, this book is not an easy read for those without the necessary background mentioned

earlier in this review. However, it presents analysis of data that some of those readers may have seen analyzed ignoring the dependence in the data (for example, the Old Faithful data); this alone may prove useful to broaden the horizon for future analyses. Those who have the background necessary to use the R code and to replicate the results throughout the book, will find plenty of material in this book to extend what they learn to their own data. The book is written very pedagogically, albeit details of the computations need to be inferred by the reader at times, and all the data sets, errata sheet, R code, among other things, can be accessed at the web site <http://134.76.173.220/hmm-with-r/>.

**Reviewer:**

Juana Sanchez  
University of California, Los Angeles  
Department of Statistics  
8125 Math Sciences Building, Box 951554  
Los Angeles, CA 90095-1554, United States of America  
E-mail: [jsanchez@stat.ucla.edu](mailto:jsanchez@stat.ucla.edu)  
URL: <http://www.stat.ucla.edu/~jsanchez/>