Further reading

Quick links

- Stan website
- The Stan Forums (get help from Stan developers and other users)
- Stan documentation (links to various kinds of documentation for Stan)
- Contributed talks and materials from the 2017 Stan conference (mostly about interesting applications of Stan), including slides & code (link to repository)
- Andrew Gelman's blog

R packages from the Stan development team

- rstan (R interface to Stan)
- rstanarm (provides a traditional R formula interface for fitting common applied regression models with Stan, without having to write the Stan code yourself)
- bayesplot (plotting)
- shinystan (interactive tables and visualizations)
- loo (efficient approximate leave-one-out cross-validation for Bayesian models)

Hamiltonian Monte Carlo (HMC) and related background

I highly recommend my Stan colleague Michael Betancourt's intro to HMC paper. Michael has a lot of very technical papers about HMC but this one is primarily focused on providing intuition (e.g., he has a whole section on the connection between HMC and the physics of planetary motion that I showed in a slide):

- A Conceptual Introduction to Hamiltonian Monte Carlo (paper)

This one is aimed at ecologists, but the HMC explanation is well written so it's a good read even if not an ecologist:

- Faster Estimation of Bayesian Models in Ecology using Hamiltonian Monte Carlo (paper)

This case study from my colleague Bob Carpenter uses simple simulations to demonstrate how things get strange (and challenging) very quickly as the number of dimensions grows:

- Typical Sets and the Curse of Dimensionality (case study)

Diagnostics, reparameterizations, priors

- Diagnosing Biased Inference with Divergences (case study)
- How the Shape of a Weakly Informative Prior Affects Inferences (case study)
- The Impact of Reparameterization on Point Estimates (case study)
- The prior can generally only be understood in the context of the likelihood (paper)
- The QR Decomposition for Regression Models (case study)

Visualization and graphical model checking

- Visualization in Bayesian Workflow (paper)
- **bayesplot** package tutorials (online vignettes)

Time series & spatial models

- Chapter 10 in the Stan Manual v2.16.0
- Spatial Models in Stan: Intrinsic Auto-Regressive Models for Areal Data (case study)
- Stan tutorial: Modern Bayesian Tools for Time Series Analysis contributed by Stan users Thomas P. Harte and R. Michael Weylandt.
- You can also find tons of examples of simple and complicated time series modeling in Stan just by Googling

Measurement error & missing data

- Missing data: chapter 11 in the Stan Manual v2.16.0
- Measurement error: chapter 14 in the Stan Manual v2.16.0

Survival (duration) analysis

Some Stan users have written Python and R libraries to help fit certain survival models using Stan:

- Library of Stan Models for Survival Analysis from Jacki Novik and HammerLab
- survHE R package for fitting survival models via RStan from Gianluca Baio
- Chapters 11 through 15 in the Stan Manual v2.16.0 all have content that relates in some way to survival models.
- Paper and Stan code for survival analysis with shrinkage priors from Aki Vehtari (link). (Note: this is a few

years old so the Stan code may use some deprecated syntax)

Model comparison, predictive performance, variable selection

Note: some of these papers have been published in various journals but I'm including links to the free arXiv preprint versions.

- Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC (arXiv, R package)
- Understanding predictive information criteria for Bayesian models (arXiv)
- Projection predictive variable selection using Stan+R (arXiv, R package)
- Using stacking to average Bayesian predictive distributions (arXiv)
- Comparison of Bayesian predictive methods for model selection (arXiv)

Item response theory

Note: the Stan programs in these case studies were written using some old syntax that is now deprecated but still works (e.g., assignment with "<-" instead of "=").

- Two-Parameter Logistic Item Response Model (case study)
- Rasch and Two-Parameter Logistic Item Response Models with Latent Regression (case study)
- Partial Credit and Generalized Partial Credit Models with Latent Regression (case study)
- Rating Scale and Generalized Rating Scale Models with Latent Regression (case study)
- Hierarchical Two-Parameter Logistic Item Response Model (case study)
- Chapter 9, section 11 in the Stan Manual v2.16.0
- Fitting Bayesian item response models in Stata and Stan (arXiv)

Mixture models

- Identifying Bayesian Mixture Models (case study)
- Chapter 13 in the Stan Manual v2.16.0

Gaussian processes

We didn't talk about Gaussian processes but I get asked about them a lot so here are some links just in case anyone is interested:

- Chapter 18 in the Stan Manual v2.16.0
- Hierarchical Gaussian Processes in Stan (Rob Trangucci's talk from StanCon 2017)

- Modeling the Rate of Public Mass Shootings with Gaussian Processes (Nathan Sanders' talk from StanCon 2017)
- GP example code recently updated by Rob Trangucci (example models repository)

Economics-related textbooks

This book is pretty good but it was written before Stan (everything in the book can be done in Stan though):

- An Introduction to Modern Bayesian Econometrics by Tony Lancaster

A forthcoming textbook that should be excellent but won't be published until 2018:

- Bayesian Econometrics with Stan by Jim Savage et al.

The author of that forthcoming book, Jim Savage, has a blog that sometimes has good economics-related Stan content:

- Jim Savage blog