

1 **A simple parameterization for segmented rating curves**

2 **T. O. Hodson<sup>1</sup>, T. A. Kenney<sup>2</sup>, T. M. Over<sup>1</sup>, and M. B. Yeheyis<sup>3</sup>**

3 <sup>1</sup>U.S. Geological Survey Central Midwest Water Science Center, Urbana, Illinois

4 <sup>2</sup>U.S. Geological Survey Water Resources Mission Area, West Valley City, Utah

5 <sup>3</sup>Environment and Climate Change Canada, Gatineau, Quebec

6 **Key Points:**

- 7
- Rating curves are widely used in hydrology to generate streamflow timeseries
  - However, most rating curves are still fit manually
  - This paper proposes a reference parameterization for automating that fitting
- 8
- 9

---

Corresponding author: Timothy Hodson, [thodson@usgs.gov](mailto:thodson@usgs.gov)

## Abstract

Streamflow is one of the most important variables in hydrology but is difficult to measure continuously. As a result, nearly all streamflow time series are estimated from rating curves that define a mathematical relationship between streamflow and some easy-to-measure surrogate like water-surface elevation (stage). Most ratings are still fit manually, which is time-consuming and subjective. To improve that process, the U.S. Geological Survey (USGS), among others, is evaluating algorithms to automate that fitting. Several automated methods already exist, and each parameterizes the rating curve slightly differently. Because of the nonconvex nature of the problem, those differences can greatly affect performance. After some trial and error, we settled on reparameterizing the classic segmented power law somewhat like a Bayesian physics-informed neural network. Being physics-informed and Bayesian, the algorithm requires minimal data and also estimates uncertainty. Its implementation is open source and easily modified so that others can contribute to improving the quality of USGS streamflow data.

## 1 Introduction

Streamflow time series are widely used in hydrologic research, water resource management, engineering design, and flood forecasting, but they are difficult to measure directly. In nearly all time-series applications, streamflow is estimated from rating curves or “ratings” that describe the relation between streamflow and an easy-to-measure surrogate like stage. The shape of the rating is specific to each streamgauge and is governed by channel conditions at or downstream from the gage, referred to as controls. Section controls, like natural riffles or artificial weirs, occur downstream from the gage, whereas channel controls, like the geometry of the banks, represent conditions along the stream reach (the upstream and downstream vicinity of the gage). Regardless of the type, the behavior of each control is often well-approximated with standard hydraulic equations that take the general form of a power law with an offset parameter

$$q = C(h - h_0)^b \quad (1)$$

where  $q$  is the discharge (streamflow);  $h$  is the height of the water above some datum (stage);  $h_0$  is the stage of zero flow (the offset parameter);  $(h - h_0)$  is the hydraulic head;  $b$  is the slope of the rating curve when plotted in log-log space; and  $C$  is a scale factor equal to the discharge when the head is equal to one (ISO 18320:2020, 2020). When multiple controls are present, the rating curve is divided into segments with one power law corresponding to each control resulting in a multi-segment or compound rating.

Although automated methods exist, most ratings are still fit manually using a graphical method of plotting stage and discharge in log-log space. With the appropriate location parameter, each control can be fit to a straight-line segment in log space (Kennedy, 1984; ISO 18320:2020, 2020). Variants of this method have been used for decades, first with pencil and log paper and now with computer-aided software; the fitting is still done by manually adjusting parameters until an acceptable fit is achieved.

Single-segment ratings are relatively easy to fit by automated methods (Venetis, 1970), but compound ratings are more challenging because their solution is nonconvex or multimodal (Reitan & Petersen-Øverleir, 2006). As a result, optimization algorithms can become stuck in local optima and fail to converge to the global optimum. General function approximators, such as natural splines (Fenton, 2018) or neural networks, can be easier to fit but their generality comes at a cost. The form of the power law matches that of the hydraulic equations governing uniform open-channel flow, like the Manning equation (Manning, 1891). Due to that physical basis, power laws are potentially more robust than other generic curve-fitting functions: requiring less data to achieve the same fit and being less prone to overfitting.

This paper describes a basic algorithm for fitting a segmented power law with offset parameters—the classic rating-curve form— and compares its performance against a natural cubic spline; “natural” indicates the second derivatives are set to zero at the endpoints. Several algorithms for fitting segmented power laws already exist. Some are more physical, meaning their structure corresponds to the governing hydraulic equations (Reitan & Petersen-Øverleir, 2008; Le Coz et al., 2014); some are more data-driven with more flexible structures like spline (Fenton, 2018) or local regression (Coxon et al., 2015); and some are a hybrid of the two (Hrafnkelsson et al., 2021). Each approach has different tradeoffs. More physical approaches require less data but may be nonconvex, which makes them challenging to fit, whereas data-driven approaches are easier to fit but require more data to constrain their greater flexibility.

Like our algorithm, the more physical approaches tend to be Bayesian and use sampling (as opposed to optimization; Ma et al., 2019) and priors to help mitigate problems that occur in nonconvex settings. Examples of priors include constraining the exponent  $b$  to be around  $5/3$ , constraining the number of rating segments, or constraining the transitions between segments around a particular stage. Being Bayesian, these algorithms inherently estimate uncertainty in the fitted parameters and discharge, which is important for many applications. However, they differ in their exact parameterization, and because of their nonconvex nature, slight differences can greatly affect performance.

Our algorithm distinguishes itself in two ways: its simple implementation, which uses a community-developed open-source probabilistic programming library, and its robust parameterization. These two aspects are in a way interrelated. With a community-developed library, the underlying numerical code is maintained by a broader community of developers, so instead of developing that code, we could focus on testing different parameterizations for the rating curve.

## 2 Implementation

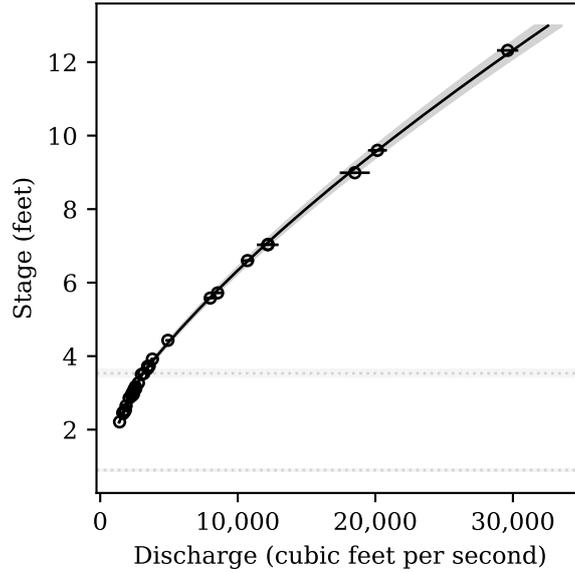
We implemented our curve-fitting algorithm, along with some additional plotting and evaluation tools, test datasets, and tutorials, as a Python package called `ratingcurve`. The algorithm uses PyMC (Salvatier et al., 2016), an open-source Python library for Bayesian statistical modeling and probabilistic machine learning. Using PyMC, the core algorithm can be expressed in several lines of code, making it easier to extend or modify, like changing the priors, the parameterization of the rating curve, or the inference algorithm to achieve different tradeoffs of speed and accuracy. This paper demonstrates two such algorithms: Automatic Differentiation Variational Inference (ADVI) (Kucukelbir et al., 2017) and Hamiltonian Monte Carlo with the No-U-Turn Sampler (NUTS) (Hoffman & Gelman, 2014). ADVI is an optimization algorithm, whereas NUTS is a sampling algorithm. In general, sampling is slower but better for nonconvex problems (Ma et al., 2019).

### 2.1 Usage

Refer to the *Open Research Section* for links to the source code repository and packaged versions of `ratingcurve`. Given observations of discharge ( $q$ ), stage ( $h$ ), and, optionally, the standard error of the discharge observations ( $e$ ), a two-segment rating is fit with

```
rating = PowerLawRating(q, h, e, 2)
trace = rating.fit()
rating.plot(trace)
```

and produces a plot like Figure 1.



**Figure 1.** Two-segment rating curve, with 95-percent prediction interval, for the Green River near Jensen, Utah (U.S. Geological Survey streamgage 09261000); generated by `rating.plot(trace)` and fit using ADVI. The circles with error bars show the observations and their uncertainty. Horizontal dotted lines show the segment breakpoints and their prediction intervals.

93 A rating curve can also be exported as a table for use by other applications, shown  
 94 in Table 1. In addition to the mean discharge for each stage, the table gives the median  
 95 and geometric standard error (GSE), which is multiplied and divided by the median to  
 estimate prediction intervals (Limpert et al., 2001).

**Table 1.** Rating table generated by `rating.table(trace)`. Units are feet (ft) and cubic feet per second ( $\text{ft}^3 \text{s}^{-1}$ ); geometric standard error (GSE) is a unitless factor.

Stage	Mean Discharge	Median Discharge	GSE
ft	$\text{ft}^3 \text{s}^{-1}$	$\text{ft}^3 \text{s}^{-1}$	-
2.20	1376.14	1376.16	1.0107
2.21	1388.27	1388.27	1.0107
2.22	1400.41	1400.40	1.0107
2.23	1412.57	1412.55	1.0106
2.24	1424.74	1424.73	1.0106
...	...	...	...

97

## 2.2 Parameterization

98

Our rating curve algorithm is similar to the segmented power law used in the manual method (Kennedy, 1984; ISO 18320:2020, 2020), as well as in automated approaches (Reitan & Petersen-Øverleir, 2008; Le Coz et al., 2014), but differs in its parameterization. Conceptually, the Reitan and Petersen-Øverleir (2008) parameterization slices the channel cross section horizontally to form each segment: segments stack one on top of the other. Once the stage rises beyond the range of a particular control, that control is "drowned out" and flow through that segment ceases to increase with stage. Our parameterization slices the channel cross-section vertically, so controls never drown out. The Le Coz et al. (2014) parameterization can slice in either direction but differs in that the segments are summed after transforming them back to their original scale; whereas, Reitan and Petersen-Øverleir (2008) sum the segments in log.

99

100

101

102

103

104

105

106

107

108

109

110

111

112

After testing several parameterizations, one seemed especially reliable and simple: slicing the channel cross section vertically into control segments and summing them in log, which is somewhat like a ReLU (rectified linear unit) neural network with hydraulic controls as neurons

$$X = \ln(\max(h - h_s, 0) + h_o) \quad (2)$$

$$\ln(q) = a + b^T X + \epsilon + \epsilon_o \quad (3)$$

113

114

115

116

117

118

where  $h_s$  are the unknown segment breakpoints;  $h_o$  is a vector of offsets, the first is 0 and the rest are 1;  $\max$  is the element-wise maximum, which returns a vector of size  $h_s$ ;  $a$  is a bias parameter equal to  $\log(C)$ , the scale factor;  $b$  are the slopes of each log-transformed segment;  $\epsilon$  is the residual error; and  $\epsilon_o$  is the uncertainty in the discharge observations (optional). The offset vector  $h_o$  ensures that  $X \geq 0$ , so additional segments never subtract discharge.

119

120

121

122

123

124

The default priors and settings are documented in the package; in general, they do not need to be modified. Besides selecting the number of segments, the user can specify a prior distribution on the breakpoints. The default assumes the breakpoints are monotonically ordered and uniformly distributed across the range of the data,  $h_{s1} < \min(h) < h_{s2} < \dots < h_{sn} < \max(h)$ . Alternatively, the user can specify approximate locations for each breakpoint and their uncertainty as normal distributions.

125

126

127

128

129

130

131

Uncertainty in the discharge observations is typically reported as relative standard error (RSE). For convenience, we convert that relative error to a geometric error as  $\epsilon_o \sim N(0, \ln(1+\text{RSE}/q)^2)$ . For small uncertainties, the difference is negligible, and for large uncertainties, it is not known which is better. Like Reitan and Petersen-Øverleir (2008), we assume  $\epsilon$  is normally distributed with mean zero and variance  $\sigma^2$ ,  $\epsilon \sim N(0, \sigma^2)$ . That simplification can create unaccounted heteroscedasticity (Petersen-Øverleir, 2004) but generally yields a reasonable estimate for the rating and its uncertainty.

132

## 3 Results

133

134

135

136

137

138

139

140

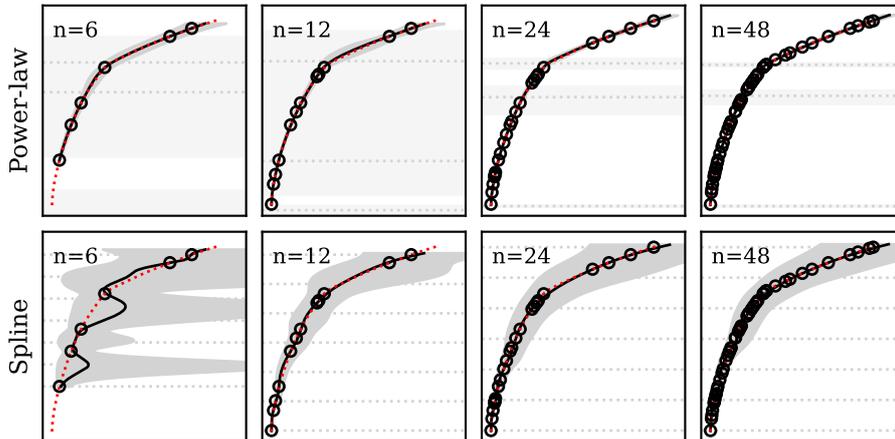
141

142

We compared the performance of the segment power law against a log-transformed natural spline on a simulated 3-segment rating curve. Both models use log transformations, which helps with heteroscedasticity. The power law is strictly increasing; otherwise, both approaches use log transformations, and both are flexible enough to approximate a wide variety of functions. Unlike the spline, the parameters in the power law have physical meaning in that they correspond to parameters in standard hydraulic equations for approximating open-channel flow. The segmented power law is notoriously difficult to calibrate (Reitan & Petersen-Øverleir, 2008), however, and its performance depends, in large part, on its parameterization—we tested several, some mathematically equivalent, some slicing the cross section vertically or horizontally—as well as its priors. If

143 the calibration challenges are overcome, the power law should yield better fits with fewer  
 144 observations (Reitan & Petersen-Overleir, 2008).

145 Figure 2 shows a side-by-side comparison of a spline and power law fit with 6, 12,  
 146 24, and 48 stage-discharge observations. For best accuracy, the curves were fit using NUTS.  
 147 We also specified that the power law had 3 segments and that the spline had 8 degrees  
 148 of freedom, the same as the power law (1 bias, 3 offsets, 3 slopes, and 1 uncertainty). Oth-  
 149 erwise, default settings were used for both.



**Figure 2.** Segmented-power law (top) and natural spline (bottom) fit with different numbers of observations ( $n$ ). The dashed red line is the true rating and the circles are the simulated observations. Horizontal dotted lines show segment breakpoints and knot locations for the power law and spline, respectively. The shaded regions depict 95 percent prediction intervals for the rating and breakpoints.

150 Although the natural spline was 5-20x faster, it yielded poorer fits, particularly when  
 151  $n = 6$ . Reducing the degrees of freedom might improve performance when  $n = 6$  but  
 152 also sacrifices flexibility when  $n = 48$ . By comparison, the power law yielded a good  
 153 fit with six observations—two fewer than the number of model parameters. Our intent  
 154 is not to disparage all splines—both parameterizations are technically splines. Rather,  
 155 we wanted to demonstrate a classic tradeoff between being easy to fit or being accurate,  
 156 which is a characteristic of data-driven and physical approaches.

157 In general, the accuracy of data-driven approaches is highly dependent on the ava-  
 158 ilability of data. For example, Coxon et al. (2015) recommends a minimum of twenty stage-  
 159 discharge measurements for their data-driven approach. Taken over the lifetime of a stream-  
 160 gage, twenty may be manageable. However, ratings shift through time from erosion, de-  
 161 position, vegetation growth, debris/ice jams, etc. (Herschy & Herschy, 2014; Mansanarez  
 162 et al., 2019), and it may be impracticable to collect twenty measurements between each  
 163 shift. In that case, a more physical approach like the power law may be a better choice,  
 164 because they require fewer observations.

165 This paper focuses on one parameterization of the classic multi-segment power law,  
 166 but undoubtedly more will emerge, which might achieve better tradeoffs of ease and ac-  
 167 curacy. For example, our comparison uses NUTS, which is accurate but slow. With 6  
 168 observations, NUTS fit the 3-segment power law in around 10 minutes. With 48 obser-  
 169 vations, NUTS completed in 1 minute; a 10x speedup. In general, stronger priors, more  
 170 observations, or fewer segments would reduce that time. By comparison, ADVI gener-  
 171 ally achieved a NUTS-like fit in several seconds, but it occasionally failed to converge

172 on the optimum solution. A better parameterization might yield better convergence with  
 173 a faster inference algorithm.

## 174 4 Conclusions

175 Despite the existence of automated methods, most stage-discharge rating curves  
 176 are still fit manually. Although the governing hydraulic equations are relatively simple  
 177 and well-understood, they are notoriously difficult to solve for multiple controls. Among  
 178 the automated methods, no parameterization has emerged as the standard, and func-  
 179 tionally equivalent parameterizations may vary greatly in performance.

180 Here, we implement a simple parameterization that works well with minimal data  
 181 and prior information. Notably, it does not address shifts in the rating curve through  
 182 time or hysteresis, and the curve is continuous but not smooth (twice differentiable). Such  
 183 limitations could be addressed, and any such effort will depend, in part, on building from  
 184 a good starting parameterization. Therefore, our simple-yet-reliable parameterization,  
 185 use of a community-developed probabilistic programming library, and packaging provide  
 186 a benchmark for operationalizing automated methods that could promote more widespread  
 187 use, testing, and refinement by the hydrologic community.

## 188 Open Research Section

189 The latest version of `ratingcurve` is available at <https://github.com/thodson-usgs/ratingcurve>, as well as <https://code.usgs.gov/wma/uncertainty/ratingcurve>.  
 190 Packaged versions are available via PyPI and conda-forge. A link to the official release  
 191 of the version used in this paper will appear here in due course.  
 192

## 193 Acknowledgments

194 This research was funded by the U.S. Geological Survey National Hydrologic Monitor-  
 195 ing Program. Any use of trade, firm, or product names is for descriptive purposes only  
 196 and does not imply endorsement by the U.S. Government.

## 197 References

- 198 Coxon, G., Freer, J., Westerberg, I. K., Wagener, T., Woods, R., & Smith, P. J.  
 199 (2015, jul). A novel framework for discharge uncertainty quantification ap-  
 200 plied to 500 UK gauging stations. *Water Resources Research*, 51(7), 5531–  
 201 5546. Retrieved from <https://doi.org/10.1002/2F2014wr016532> doi:  
 202 10.1002/2014wr016532
- 203 Fenton, J. D. (2018, sep). On the generation of stream rating curves. *Journal*  
 204 *of Hydrology*, 564, 748–757. Retrieved from [https://doi.org/10.1016/2Fj](https://doi.org/10.1016/2Fj.jhydro1.2018.07.025)  
 205 [.jhydro1.2018.07.025](https://doi.org/10.1016/2Fj.jhydro1.2018.07.025) doi: 10.1016/j.jhydro1.2018.07.025
- 206 Herschy, R. W., & Herschy, R. (2014). *Streamflow measurement* (Third ed.). CRC  
 207 Press. Retrieved from <https://doi.org/10.1201/2F9781482271485> doi: 10  
 208 [.1201/9781482271485](https://doi.org/10.1201/9781482271485)
- 209 Hoffman, M. D., & Gelman, A. (2014). The no-u-turn sampler: Adaptively setting  
 210 path lengths in hamiltonian monte carlo. *Journal of Machine Learning Re-*  
 211 *search*, 15(47), 1593–1623. Retrieved from [http://jmlr.org/papers/v15/](http://jmlr.org/papers/v15/hoffman14a.html)  
 212 [hoffman14a.html](http://jmlr.org/papers/v15/hoffman14a.html)
- 213 Hrafnkelsson, B., Sigurdarson, H., Rögnvaldsson, S., Örn Jansson, A., Vias, R. D.,  
 214 & Gardarsson, S. M. (2021, dec). Generalization of the power-law rating  
 215 curve using hydrodynamic theory and bayesian hierarchical modeling. *Environ-*  
 216 *metrics*, 33(2). Retrieved from <https://doi.org/10.1002/2Fenv.2711> doi:  
 217 10.1002/env.2711

- 218 ISO 18320:2020. (2020). *Hydrometry—measurement of liquid flow in open channels—*  
 219 *determination of the stage-discharge relationship*. International Organization  
 220 for Standardization.
- 221 Kennedy, E. J. (1984). *Discharge ratings at gaging stations, U.S. Geological Survey*  
 222 *techniques of water-resources investigations 09-a10*. Retrieved from [https://](https://doi.org/10.3133/twri03A10)  
 223 [doi.org/10.3133/twri03A10](https://doi.org/10.3133/twri03A10)
- 224 Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., & Blei, D. M. (2017). Au-  
 225 tomatic differentiation variational inference. *Journal of Machine Learning Re-*  
 226 *search*, 18(14), 1–45. Retrieved from [http://jmlr.org/papers/v18/16-107](http://jmlr.org/papers/v18/16-107.html)  
 227 [.html](http://jmlr.org/papers/v18/16-107.html)
- 228 Le Coz, J., Renard, B., Bonnifait, L., Branger, F., & Le Boursicaud, R. (2014,  
 229 Feb). Combining hydraulic knowledge and uncertain gaugings in the es-  
 230 timation of hydrometric rating curves: A bayesian approach. *Journal of*  
 231 *Hydrology*, 509, 573–587. Retrieved from [http://dx.doi.org/10.1016/](http://dx.doi.org/10.1016/j.jhydrol.2013.11.016)  
 232 [j.jhydrol.2013.11.016](http://dx.doi.org/10.1016/j.jhydrol.2013.11.016) doi: 10.1016/j.jhydrol.2013.11.016
- 233 Limpert, E., Stahel, W. A., & Abbt, M. (2001). Log-normal distributions across  
 234 the sciences: keys and clues. *BioScience*, 51(5), 341. Retrieved from [http://](http://dx.doi.org/10.1641/0006-3568(2001)051[0341:lndats]2.0.co;2)  
 235 [dx.doi.org/10.1641/0006-3568\(2001\)051\[0341:lndats\]2.0.co;2](http://dx.doi.org/10.1641/0006-3568(2001)051[0341:lndats]2.0.co;2) doi: 10  
 236 [.1641/0006-3568\(2001\)051\[0341:lndats\]2.0.co;2](http://dx.doi.org/10.1641/0006-3568(2001)051[0341:lndats]2.0.co;2)
- 237 Ma, Y.-A., Chen, Y., Jin, C., Flammarion, N., & Jordan, M. I. (2019, sep). Sam-  
 238 pling can be faster than optimization. *Proceedings of the National Academy of*  
 239 *Sciences*, 116(42), 20881–20885. Retrieved from [https://doi.org/10.1073/](https://doi.org/10.1073/pnas.1820003116)  
 240 [pnas.1820003116](https://doi.org/10.1073/pnas.1820003116) doi: 10.1073/pnas.1820003116
- 241 Manning, R. (1891). On the flow of water in open channels and pipes. *Transactions*  
 242 *of the Institution of Civil Engineers of Ireland*, 20, 179–207.
- 243 Mansanarez, V., Renard, B., Coz, J. L., Lang, M., & Darienzo, M. (2019, Apr). Shift  
 244 happens! adjusting stage-discharge rating curves to morphological changes at  
 245 known times. *Water Resources Research*, 55(4), 2876–2899. Retrieved from  
 246 <http://dx.doi.org/10.1029/2018wr023389> doi: 10.1029/2018wr023389
- 247 Petersen-Øverleir, A. (2004, jun). Accounting for heteroscedasticity in rating curve  
 248 estimates. *Journal of Hydrology*, 292(1-4), 173–181. Retrieved from [https://](https://doi.org/10.1016/j.jhydrol.2003.12.024)  
 249 [doi.org/10.1016/j.jhydrol.2003.12.024](https://doi.org/10.1016/j.jhydrol.2003.12.024) doi: 10.1016/j.jhydrol.2003.12  
 250 [.024](https://doi.org/10.1016/j.jhydrol.2003.12.024)
- 251 Reitan, T., & Petersen-Øverleir, A. (2006, jul). Existence of the frequentistic  
 252 estimate for power-law regression with a location parameter, with applica-  
 253 tions for making discharge rating curves. *Stochastic Environmental Research*  
 254 *and Risk Assessment*, 20(6), 445–453. Retrieved from [https://doi.org/](https://doi.org/10.1007/s00477-006-0037-6)  
 255 [10.1007/s00477-006-0037-6](https://doi.org/10.1007/s00477-006-0037-6) doi: 10.1007/s00477-006-0037-6
- 256 Reitan, T., & Petersen-Øverleir, A. (2008, Jun). Bayesian methods for estimat-  
 257 ing multi-segment discharge rating curves. *Stochastic Environmental Research*  
 258 *and Risk Assessment*, 23(5), 627–642. Retrieved from [http://dx.doi.org/10](http://dx.doi.org/10.1007/s00477-008-0248-0)  
 259 [.1007/s00477-008-0248-0](http://dx.doi.org/10.1007/s00477-008-0248-0) doi: 10.1007/s00477-008-0248-0
- 260 Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016, apr). Probabilistic program-  
 261 ming in python using PyMC3. *PeerJ Computer Science*, 2, e55. Retrieved  
 262 from <https://doi.org/10.7717/peerj-cs.55> doi: 10.7717/peerj-cs.55
- 263 Venetis, C. (1970, jun). A note on the estimation of the parameters in log-  
 264 arithmic stage-discharge relationships with estimates of their error. *Inter-*  
 265 *national Association of Scientific Hydrology. Bulletin*, 15(2), 105–111.  
 266 Retrieved from <https://doi.org/10.1080/02626667009493957> doi:  
 267 [10.1080/02626667009493957](https://doi.org/10.1080/02626667009493957)