Development and Implementation of a Bayesian Model for Sediment Transport in Fluvial Systems

Mark Schmelter
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports

Part of the Statistics and Probability Commons

Recommended Citation
DEVELOPMENT AND IMPLEMENTATION OF A BAYESIAN MODEL FOR SEDIMENT TRANSPORT IN FLUVIAL SYSTEMS

by

Mark Schmelter

A report submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Statistics

UTAH STATE UNIVERSITY
Logan, Utah
2011
ABSTRACT

Development and Implementation of a Bayesian Model for Sediment Transport in Fluvial Systems

by

Mark Schmelter, Master of Science
Utah State University, 2011

Major Professor: Dr. Mevin Hooten
Department: Mathematics and Statistics

Recent studies in the field of fluvial sediment transport underscore the difficulty in reliably estimating transport model parameters, collecting accurate observations, and making predictions due to measurement error and conceptual model uncertainty. There is a pressing need to develop models that can account for measurement error, conceptual model uncertainty, and natural variability while providing probability-based predictions as well as a means for conceptual model discrimination. The model presented in this research employs an excess shear sediment transport equation for a uni-size sediment bed developed in a Bayesian statistical framework. This statistical model provides a means to rigorously estimate distributions of model parameters, such as critical shear, given observations of sediment transport. The model provides transport predictions in the form of a posterior predictive distribution from which credible intervals of sediment transport can be designated. The proposed framework utilizes the Deviance Information Criterion to quantify model fit with model parsimony. This approach relies upon the incorporation of expert judgment in the form of prior distributions for model parameters of interest. The uni-size sediment transport
model developed in this research was tested against simulated observations for which the ‘true’ model parameters were known. Results of the simulation studies indicate that such a modeling approach is valid and presents opportunities for expert judgment to bolster parameter inference through the incorporation of prior knowledge. The proposed model was also tested against laboratory flume data for validity; results indicate that this framework is promising as it allows modelers to evaluate competing conceptual models; provides credible intervals of parameters and predictions; and weighs prior expert knowledge with information contained in new observations.
ACKNOWLEDGMENTS

Deciding to pursue a degree in statistics while concurrently working on my Ph.D. was a crazy idea put in my mind by my sister-in-law. After some time thinking about it, I spoke with my wife, and she was fully supportive of the idea (big surprise) and, in fact, was more eager for me to do it than I probably was. Looking back on the decision to get a degree in statistics I truly feel that, had it not been for family encouraging me, I would not have received this additional training and education. As it turns out, my goals and interests in engineering are facilitated by the analytical tools I have learned in my statistics curriculum. Thank you Kady for putting that idea in my head; thank you Keely for letting me languish a little longer in graduate school while I acquire tools and skills that will allow me to pursue those interests.

Working with my M.S. advisor, Mevin Hooten, has been a real pleasure. I not only appreciate his instruction during various courses over the past several years, but also his help and direction in the completion of this research. Jürgen Symanzik has been the lucky individual appointed to try to teach me mathematical statistics—a challenge for both of us. I appreciate his eagerness to help me understand statistics as well as his availability to answer my questions. Lastly, I thank my Ph.D. advisor, David Stevens, for allowing me to pursue my own research interests, letting me sidetrack for a moment into statistics, and for the financial support he has been so generous with over the past few years. Thank you all.

Mark Schmelter
To my children—the very source of randomness, variability, and uncertainty in my life.
# Table of Contents

3.2.2 Model Results .......................................................... 72

4 DISCUSSION ........................................................................ 79
  4.1 Simulation Studies ......................................................... 79
    4.1.1 Data Scenario 1 .................................................... 84
    4.1.2 Data Scenario 2 .................................................... 88
    4.1.3 Data Scenario 3 .................................................... 89
    4.1.4 Data Scenario 4 .................................................... 91
    4.1.5 General Comments on Simulated Observations ........... 91
    4.1.6 Process Model Selection ......................................... 95
  4.2 Smart (1984) Flume Data ............................................... 96
    4.2.1 General Discussion on Flume Data ......................... 97

5 CONCLUSIONS ................................................................. 105
  5.1 Summary of Research .................................................... 105
  5.2 Addressing Needs from the Literature ............................ 107
  5.3 Future Research .......................................................... 109

REFERENCES ...................................................................... 111

APPENDICES ...................................................................... 118
  APPENDIX A R CODE .......................................................... 119
  APPENDIX B MODEL FIT SUMMARY PLOTS ....................... 141
LIST OF TABLES

Table | Page
--- | ---
1 | Sample values for fixed parameters. 24
2 | Parameterization of common variables for simulated observations. 47
3 | \( \tau_c \) posterior sensitivity table for data scenario 1. 52
4 | \( \tau_c \) posterior sensitivity table for data scenario 2. 53
5 | \( \tau_c \) posterior sensitivity table for data scenario 3. 54
6 | \( \tau_c \) posterior sensitivity table for data scenario 4. 55
7 | MSPE values for all data scenarios and priors. 60
8 | Comparison of priors for data scenario 4. 60
9 | \( \sigma^2 \) posterior sensitivity table for data scenario 1. 65
10 | \( \sigma^2 \) posterior sensitivity table for data scenario 2. 66
11 | \( \sigma^2 \) posterior sensitivity table for data scenario 3. 67
12 | \( \sigma^2 \) posterior sensitivity table for data scenario 4. 68
13 | Deviance Information Criterion results for competing simulation models. 69
14 | Summary of 10.5 mm flume observations from Smart (1984). 71
15 | Comparison of process model performance on the Smart (1984) 10.5 mm flume observations—\( \tau_c \). Asterisks denote dimensionless values. 73
16 | Comparison of process model performance on the Smart (1984) 10.5 mm flume observations—\( \sigma^2 \). 73
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Two possible priors for a hypothetical parameter $\theta$. Both priors have</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>the same mean value, though the black prior is much more precise than the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>red prior.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Data simulation algorithm</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>Univariate trace plot with dispersed initial values. Squares represent the</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>dispersed initial values used for iteration 0.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Bivariate trace plot with dispersed initial values for both: (a) immature,</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>(b) mature Markov chains. Squares represent the dispersed initial values</td>
<td></td>
</tr>
<tr>
<td></td>
<td>used for iteration 0.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Flexibility of the truncated normal distribution. The specifications of</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>each prior follows: $T.N.(\mu, \sigma, a, b)$.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Summary of simulated observation scenarios. Each data scenario is a</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>distinct combination of high/low variance and many/few observations.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(a) $N = 21$, $\sigma^2 = 0.05$, (b) $N = 6$, $\sigma^2 = 0.05$, (c) $N$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>21, $\sigma^2 = 1.10$, (d) $N = 6$, $\sigma^2 = 1.10$.</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Four prior scenarios. Red line denotes true $\tau_c$ in simulated data</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>scenarios. Parameterization follows: $T.N.(\mu_{\tau_c}, \sigma^2_{\tau_c}, \tilde{a}, b)$.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Priors for $\sigma^2$: (a) data scenarios 1 and 2, (b) data scenarios 3</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>and 4. Vertical lines denote the expected value.</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>$\tau_c$ posterior sensitivity plot for data scenario 1. (a) prior 1, (b)</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>prior 2, (c) prior 3, (d) prior 4.</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>$\tau_c$ posterior sensitivity plot for data scenario 2. (a) prior 1, (b)</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>prior 2, (c) prior 3, (d) prior 4.</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>$\tau_c$ posterior sensitivity plot for data scenario 3. (a) prior 1, (b)</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>prior 2, (c) prior 3, (d) prior 4.</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>$\tau_c$ posterior sensitivity plot for data scenario 4. (a) prior 1, (b)</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>prior 2, (c) prior 3, (d) prior 4.</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Cross-validation plots for data scenario 1. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.</td>
<td>56</td>
</tr>
<tr>
<td>14</td>
<td>Cross-validation plots for data scenario 2. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.</td>
<td>57</td>
</tr>
<tr>
<td>15</td>
<td>Cross-validation plots for data scenario 3. (a) Prior 1, (b) prior 2, (c) prior 3, (d) prior 4.</td>
<td>58</td>
</tr>
<tr>
<td>16</td>
<td>Cross-validation plots for data scenario 4. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.</td>
<td>59</td>
</tr>
<tr>
<td>17</td>
<td>Informed prior for data scenario 4. $\tau_c \sim T.N.(70, 30, 0, 6.92)$.</td>
<td>61</td>
</tr>
<tr>
<td>18</td>
<td>Summary plot for informed prior on data scenario 4.</td>
<td>62</td>
</tr>
<tr>
<td>19</td>
<td>MSPE plot using informed prior on data scenario 4.</td>
<td>63</td>
</tr>
<tr>
<td>20</td>
<td>Two prior scenarios for $\sigma^2$. Parameterization follows: $I.G.(\mu_0^2, \sigma_{\sigma^2}^2)$.</td>
<td>64</td>
</tr>
<tr>
<td>21</td>
<td>$\sigma^2$ posterior sensitivity plot for data scenario 1.</td>
<td>65</td>
</tr>
<tr>
<td>22</td>
<td>$\sigma^2$ posterior sensitivity plot for data scenario 2.</td>
<td>66</td>
</tr>
<tr>
<td>23</td>
<td>$\sigma^2$ posterior sensitivity plot for data scenario 3.</td>
<td>67</td>
</tr>
<tr>
<td>24</td>
<td>$\sigma^2$ posterior sensitivity plot for data scenario 4.</td>
<td>68</td>
</tr>
<tr>
<td>25</td>
<td>Smart (1984) 10.5 mm transport data.</td>
<td>70</td>
</tr>
<tr>
<td>26</td>
<td>Prior distributions for Smart (1984) observations.</td>
<td>72</td>
</tr>
<tr>
<td>27</td>
<td>Summary plot of Smart process model for 10.5 mm flume observations.</td>
<td>74</td>
</tr>
<tr>
<td>28</td>
<td>Summary plot of Meyer-Peter Müller process model for 10.5 mm flume observations.</td>
<td>75</td>
</tr>
<tr>
<td>29</td>
<td>Summary plot of Wong-Parker process model for 10.5 mm flume observations.</td>
<td>76</td>
</tr>
<tr>
<td>30</td>
<td>Posterior predictive distribution of Smart process model for 10.5 mm flume observations.</td>
<td>77</td>
</tr>
<tr>
<td>31</td>
<td>Posterior predictive distribution of Meyer-Peter Müller process model for 10.5 mm flume observations.</td>
<td>77</td>
</tr>
</tbody>
</table>
32 Posterior predictive distribution of Wong-Parker process model for 10.5 mm flume observations. ................................................................. 78

33 Annotated posterior predictive distribution of Smart process model for 10.5 mm flume observations. Circled points represent subsequent observations where the observation encircled in green has a lower grain shear than the observation encircled in orange. ................................. 101

34 Summary plot for data scenario 1, \( \tau_c \) Prior 1. ................................. 141

35 Summary plot for data scenario 1, \( \tau_c \) Prior 2. ................................. 142

36 Summary plot for data scenario 1, \( \tau_c \) Prior 3. ................................. 143

37 Summary plot for data scenario 1, \( \tau_c \) Prior 4. ................................. 144

38 Summary plot for data scenario 2, \( \tau_c \) Prior 1. ................................. 145

39 Summary plot for data scenario 2, \( \tau_c \) Prior 2. ................................. 146

40 Summary plot for data scenario 2, \( \tau_c \) Prior 3. ................................. 147

41 Summary plot for data scenario 2, \( \tau_c \) Prior 4. ................................. 148

42 Summary plot for data scenario 3, \( \tau_c \) Prior 1. ................................. 149

43 Summary plot for data scenario 3, \( \tau_c \) Prior 2. ................................. 150

44 Summary plot for data scenario 3, \( \tau_c \) Prior 3. ................................. 151

45 Summary plot for data scenario 3, \( \tau_c \) Prior 4. ................................. 152

46 Summary plot for data scenario 4, \( \tau_c \) Prior 1. ................................. 153

47 Summary plot for data scenario 4, \( \tau_c \) Prior 2. ................................. 154

48 Summary plot for data scenario 4, \( \tau_c \) Prior 3. ................................. 155

49 Summary plot for data scenario 4, \( \tau_c \) Prior 4. ................................. 156

50 Summary plot for data scenario 1, \( \sigma^2 \) Prior 1. ................................. 157

51 Summary plot for data scenario 1, \( \sigma^2 \) Prior 2. ................................. 158

52 Summary plot for data scenario 2, \( \sigma^2 \) Prior 1. ................................. 159
<table>
<thead>
<tr>
<th>Page</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>Summary plot for data scenario 2, $\sigma^2$ Prior 2.</td>
</tr>
<tr>
<td>54</td>
<td>Summary plot for data scenario 3, $\sigma^2$ Prior 1.</td>
</tr>
<tr>
<td>55</td>
<td>Summary plot for data scenario 3, $\sigma^2$ Prior 2.</td>
</tr>
<tr>
<td>56</td>
<td>Summary plot for data scenario 4, $\sigma^2$ Prior 1.</td>
</tr>
<tr>
<td>57</td>
<td>Summary plot for data scenario 4, $\sigma^2$ Prior 2.</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

1.1 Background and Motivation

1.1.1 Water in Human Civilization

A brief survey of human history underscores the importance of natural resources, and to a large degree water, in the cultural and political dealings of our ancestors. When one considers the vitality of a society, often it is not necessary to look further than the security of their water resources. Indeed, the statement that “the history of human civilization is entwined with the history of the ways we have learned to manipulate water resources” (Gleick, 2001) provides context for virtually all water development projects in the world, past and present.

Many water-related relics from past eras are still in existence today. For example, Graf (1971) discusses evidence from Mesopotamia that dates the construction of canals at around 4000 B.C.; canals extending from the Tigris to the Euphrates supplied Babylon with the water resources it required for its population. Additionally, under the rule of King Menes (3000 B.C.), dams began to be built in Egypt, and later (from 1000 to 700 B.C.) the Marib dam, one of the wonders of the ancient world, was constructed. As early as 1200 B.C. water tunnels were used to supply the city of Jerusalem with water. Graf (1971) also mentions water development projects in ancient China. In the environs of 2278 B.C. Chinese rivers were being tamed. So impressed at the accomplishments of his dam builder that the then Emperor Yau ceded the Empire not to his kin, but to Yu, the man who was able to bridle nine Chinese rivers. As we consider the intersection of human history and water history, we see
that the two are linked, albeit in a one-way relationship (we need water, but water does not need us). Empires grew and flourished when water was plentiful—Egypt being ‘the gift of The Nile’—and withered in its absence, as did the Angkor during the 15th Century (Stone and Clark, 2009).

While the study and manipulation of rivers and water began as long ago as pre-deluge societies, as far as we know these studies and projects were empirically-based—it was not until the Renaissance that formal studies of flow mechanics and transport began to persist in the nascent body of scientific literature. The earliest research is credited to Leonardo da Vinci (1452-1519), Galileo Galilei (1564-1642), and subsequent students who researched and wrote treatises on fluid mechanics and river hydraulics (Graf, 1971). During his studies, Leonardo da Vinci observed sediment transport and remarked: “Where the water has least movement, the bottom will be of the finest mud or sand, where the water has a stronger current the shingle is larger” (Simons and Senturk, 1992).

Later, French, German, and English researchers would build upon previous work to establish a rich literature. Pierre duBuat (1734-1809) has often been cited as the father of French hydraulics and produced a series of books—published in 1786—which are comprehensive in both theory and experiment entitled “Principes d’Hydraulique”. In this publication, duBuat states: “The way running water attacks the bed which cannot resist any more, and how the sand is shoved is fascinating indeed, and deserves description” (Dubuat-Nançay, 1786). DuBuat then continues to describe the mechanics of bed forms as well as bed armoring, and ultimately the shear-resistance concept, which some regard as his most influential contribution (Graf, 1971).

While researchers were making progress in theoretical aspects of sediment transport, observations were limited. It was not until Grove K. Gilbert, and his 1914 United States Geological Survey Professional Paper (Gilbert, 1914), that a compre-
hensive set of observations was published. Gilbert recorded transport observations coupled with measurements of depth, slope, bedforms, and velocity from extensive flume studies. The two-hundred-plus page report provided scientific linkages between water discharge and sediment transport. In more modern research, Luna Leopold taught: “Rivers are both the means and the routes by which the products of continental erosion are carried to the oceans of the world” (Leopold, 1994). Thus we see that rivers are composed of sediment as much as they are of water and in natural systems, where water flows, so will sediment.

After six-thousand years of human alteration of river systems (Nienhuis and Leuven, 2001), human societies are still pursuing water. Some manifestations of this pursuit include the Itaipu dam on the Parana River on the border of Paraguay and Brazil or the Three Gorges Dam on the Yangtze in China. The dams on the Colorado River in the American West are also iconic of the shaping power of water, both physical and political. Written between the lines of water resource development is politics—both intra- and inter-national. The allocation and disbursal of Colorado river water has caused friction not only among states, but also between the U.S. and downstream resident, Mexico (Reisner, 1993). Current events on the Asian continent indicate a similar scenario developing with Chinese interests proposing dams on rivers that ultimately join the sea in neighboring countries—these neighboring economies being historically tied to the water supplied by these rivers (Economist, 2010). The drive behind this insatiable thirst are reminders of the past—such as America’s Dust Bowl of the previous century; the present—water shortages in regions such as Africa and Australia and the desire of growing powers to industrialize; and the future—climate change in the face of population dynamics.

Given the extensive and on-going human alterations to rivers, very few systems remain unmodified. For example, a 2007 study of Australian rivers revealed that
only 14% of the continent’s rivers could be classified as “largely unmodified” (Norris et al., 2007). Further, Nilsson et al. (2005) report that only half (48%) of the world’s large river systems remain unfragmented by dams in the main channel. Because of these widespread alterations to rivers, recent years have seen a rapid rise in river restoration efforts in the U.S. and abroad. For instance, Bernhardt et al. (2005) provide a summary of over 37,000 river restoration projects contained in the National River Restoration Science Synthesis (NRRSS) database up to the year 2004. Of these projects, only 58% had associated financial records, by which an estimate of $9.1 billion total project costs were detailed with $7.5 billion being spent after 1990. These ecologically-focused projects recognize that a complete restoration to the systems’ original state is practically infeasible and so a more attainable goal is to rehabilitate rivers by reinitializing some of a system’s original ecological functions and values (Gore and Shields, 1995; Wohl et al., 2005). Further, Wohl et al. (2005) emphasize that restoration science is handicapped by fundamental uncertainties in process knowledge and that restoration decisions are necessarily dependent on expert judgment since they are made in the face of significant uncertainty. Wohl et al. (2005) advocate quantifying project outcomes in terms of probabilities that reflect the explicitly recognized uncertainties and complexities.

1.1.2 Decision-making under Uncertainty

At present, when dealing with issues related to water, we find ourselves operating in an environment of conflicting interests and uncertainty. In this arena, science needs to be relied upon as a rational, mediating influence, informing us what is known and how well it is known. Of course, the extent to which this is realized varies, but this notwithstanding, work must be done to provide scientifically-based conclusions in the presence of uncertainty.
Decades of research into the psychology of decision making in the face of uncertainty, however, provides little confidence in individuals’ intuitive ability to do so. Tversky and Kahneman (1981, 1986) cite reversals of preference to equifinal scenarios based solely on the framing of the scenario and the representation of probability. The fact that, when presented with uncertain situations, the ultimate decision can be (knowingly or unknowingly) influenced one way or the other suggests a need for scientists to not only carefully account for uncertainty in their professional endeavors, but that they pay careful attention to the way they present their results as well. Indeed, Tversky and Kahneman (1981) state: “The dependence of preferences on the formulation of decision problems is a significant concern for the theory of rational choice.” From this we see that even well-posed scenarios of quantified uncertainties can lead to contradictory outcomes based on one’s frame of reference. In addition to the tendency for preference reversal, humans are also prone to err when assigning probabilities to outcomes. Intuitive judgments of probabilities often (in laypeople as well as in experts) do not comply with simple and fundamental laws of probability and our natural assessments of probability are biased through perceptions of similarity, representativeness, attributions of causality, and associations (Tversky and Kahneman, 1983).

In their experiments, Tversky and Kahneman created environments where decisions needed to be made based on some disclosed information. In one paper (Tversky and Kahneman, 1981) subjects were presented with a set of contingencies in which they stood to win a pre-defined amount of money. In many instances, Tversky and Kahneman were able to influence subjects towards a specified decision based solely on the way the problem was presented even though the ‘target’ decision was not optimal. In looking at each of the problems specified, the best decision is almost trivially-clear when evaluated using a fundamental statistical concept called the expected value.
The application of expected values in scenarios posed by Tversky and Kahneman illustrates the utility of statistical methods in decision-making. Statistics is generally concerned with making decisions in uncertain situations and provide a way for us to negate the shortcomings of human intuition in uncertain circumstances.

1.1.3 Bayesian Modeling and Sediment Transport

Given human difficulties in dealing *intuitively* with uncertainty, as well as the complex phenomenon of sediment transport, a robust framework in which sediment transport and its associated uncertainties (both measurement and process) can be accommodated offers an interesting scientific problem and practical applications. Process uncertainties address the conceptual and deterministic models we use to calibrate and predict transport. In the literature, numerous sediment transport relations—specifically bed load—are provided, thus there is a need to select one relationship to another, and therefore some degree of uncertainty and judgment are associated with this selection. The selection of one model over another requires the intervention of an expert to choose the most appropriate model given some contextual knowledge. Next, measurement uncertainty describes the precision of the measured boundary conditions and collected data. Since sediment transport is a non-linear process, changes to the process model or to the measurements may result in radically altered output. The aim of this research is to propose and demonstrate a framework in which the uncertainties associated with sediment transport can be accommodated using established statistical modeling methods.

The fields of ecology, hydrology, and atmospheric and environmental science have seen increasing numbers of research projects that employ a Bayesian framework to model complex phenomena—for example, distributed rainfall-runoff models, species invasion dynamics, and uncertainty estimation in climate models (*Cressie et al.*, 2009;
These methods are also amenable to the process of sediment transport and provide significant benefits over a purely deterministic approach.

First, sediment transport modeling involves the estimation of model parameters—values which cannot be observed directly and which must be inferred from observational data. Traditional approaches to modeling transport entails the point-estimation of model parameters, that is, each parameter has one, and only one value. Philosophically, this means that it is treated as fixed and unknown. In the context of sediment transport, one important model parameter is the “critical shear”, $\tau_c$ or, the “critical Shields number”, $\tau_c^*$. A traditional deterministic approach to this problem requires the specification of a single critical shear value. The selection of a single critical shear value, however, is problematic. Buffington and Montgomery (1997) compiled data from eight decades of incipient motion studies and created lists of reported critical shear values. They note that the early researchers in the field (Shields, Grass, Gessler, and Paintal) acknowledged that the threshold at which sediments move is “inherently a statistical problem” and assert that a frequency distribution of critical shear values for any given grain size is more appropriate than any single value. In conclusion they write:

Our analysis indicates that less emphasis should be placed on choosing a universal $\tau_{c50}^*$ value, while more emphasis should be placed on choosing defendable values for particular applications, given the observed methodological biases, uses of each approach, and systematic influences of sources.

---

1Buffington and Montgomery refer to mixed grain size distributions when they use $\tau_{c50}^*$. The research explained in this report, however, was developed for a uni-size sediment model and is not appropriate for a mixed grain size distribution of sediments. For a mixed grain size model, the median grain size $D_{50}$ has a critical shear of $\tau_{c50}^*$. In this research, $\tau_{c50}^* = \tau_c^*$ since it is assuming uniformly sized sediment.
of uncertainty associated with different methods and investigative conditions [italics mine].

Traditional statistical frameworks in which model parameters—especially critical shear—are treated as single, fixed values and not probability distributions of values are inconsistent with Buffington and Montgomery's conclusion. A Bayesian statistical framework for sediment transport, however, allows variables to be treated, not as single values, but as random variables arising from distributions, thereby matching theoretical formulations more completely.

Second, Bayesian models incorporate expert knowledge that may have been collected, at some cost, beforehand. Science has always been a process in which we observe, theorize, and update our knowledge as we make subsequent observations. Bayesian models make it possible to stand on the foundation of previous research and update knowledge based on new observations.

Third, a Bayesian approach allows an expert to identify and allocate uncertainty associated with conceptual models and measured values used in the model. In statistical modeling, a hypothesis is formed, observations are made, and then observations are conditioned on to make inference on parameters. In traditional deterministic approaches (mixed-models aside), the observations are treated as absolute and so any inference that is made assumes that the observations are perfect. Bayesian approaches allow for the latent process and observations to each have quantifiable error or variability, thereby giving us an idea of how much variability comes from each component.

In natural river systems, sampling bed load transport is difficult and likely to have considerable error (Gaeman et al., 2009; Wilcock, 2001) so methods used for making inference from these observations should account for this uncertainty.

Bayesian models provide results in the form of a posterior distribution, that is, a distribution of values given the observations and associated uncertainties. Note
that in the quotation above, Buffington and Montgomery state that defensible values should be chosen *given* biases and uncertainties. As will be explained later, the posterior distribution is the distribution of our parameter, \( \tau_c^* \), for example, given our observations and their uncertainty.

Fourth, predictive probability distributions are often sought as a solution to a particular problem, and yet purely deterministic approaches only yield a line through a sometimes-substantial spread of observations. Since the solution to a Bayesian model is a distribution of values, it is simple to make predictions of, in this instance, sediment transport given the available observations. The result is a predictive probability distribution determined through established laws of probability.

### 1.2 Purpose and Objectives

The goal of this research is to incorporate and address many of the elements introduced in the previous section and develop them into a fluvial sediment transport model. While work on statistical sediment transport models has appeared in the literature (*Sun and Donahue*, 2000; *Wu and Chen*, 2009; *Einstein*, 1950), I am aware of no instances in which a Bayesian formulation to the sediment transport problem has been developed and implemented. As such, the objectives of my research are to:

1. Develop and implement a Bayesian sediment transport model that, when given transport observations, makes it possible to:

   (a) estimate a credible interval\(^2\) for the critical shear parameter, \( \tau_c \).

   (b) estimate a credible interval for the variance parameter, \( \sigma^2 \).

   (c) provide sediment transport predictions delineated in credible intervals.

---

\(^2\)In Bayesian statistics, a \( (1 - \alpha)100\% \) credible interval is a quantification of the probabilities in the posterior distribution. For example, a 95\% credible interval on some parameter is interpreted as: the posterior probability that our parameter lies in the credible interval is 95\%.
(d) compare different process models for fit via quantitative metrics.

2. Perform simulation studies to evaluate the proposed framework.

(a) Simulate synthetic data according to established transport relationships with multiplicative noise according to $\sigma^2$.

(b) Validate the model—verify that the model can recover the parameters that were specified when the synthetic data were generated.

(c) Explore the effect of various specifications for prior information on model inference.

3. Evaluate the model using observed transport data from uni-size flume studies.

(a) estimate model parameters, $\tau_c$ and $\sigma^2$.

(b) evaluate different process models.

(c) provide sediment rating curve in terms of credible intervals.

The objectives cited above contribute to the field of sediment transport by providing a robust model formulation that accommodates notions of randomness, variability, and uncertainty through established Bayesian statistical methods.
CHAPTER 2
METHODS

Since the focus of this research is to develop a Bayesian sediment transport model, the implementation can be naturally divided according to physical, theoretical, and practical considerations. As described in the Introduction, Bayesian models are able to accommodate notions of intrinsic variability and as well as epistemic and measurement uncertainty. To clearly explain and illustrate how such a model is developed I have divided the formulation of my research into the following subtopics: (1) governing sediment transport relations, which describe the mathematical constructs researchers use to model sediment transport from a physical and deterministic reference; (2) the specification of the Bayesian model, with considerations to basic Bayesian modeling concepts as well as likelihood and prior distribution selection, and error structure; (3) data simulation, which discloses the process for creating synthetic data from the governing sediment transport relations and known parameters; (4) computational methods which provide a means to integrate the multi-dimensional relationships that naturally result from a Bayesian model; and lastly (5) a means for model evaluation whereby the predictive capabilities of the sediment transport model can be assessed. These subtopics will be described in what follows.

2.1 Governing Sediment Transport Relations

While there are several options for the physical description of sediment transport, this research will use an excess shear model. This formulation is intuitive in its basic, non-dimensional state but quickly becomes less-so when brought into real space. The basic principle of an excess shear model is that the amount of sediment transport is related, non-linearly, to the difference between the force induced on the grains due to
the overlying movement of water and the amount of force required to move sediment
of an arbitrary size (Wilcock et al., 2009). Mathematically, this relationship is:

\[ q^* = a(\tau^* - \tau_c^*)^b, \tag{1} \]

where \( q^* \) is the “Einstein transport parameter”—a non-dimensional quantification of
sediment transport—\( \tau^* \), widely known as the Shields number, quantifies the shear
stress experienced by the grain, and \( \tau_c^* \) is the critical Shields number, or the critical
shear at which the particle will start to move. The parameters \( a \) and \( b \) are empir­
ical coefficients and can take on a range of values reflecting the numerous different
proposed transport relations. While the individual components of (1) are entirely
non-dimensional, these parameters can be expressed in dimensional terms:

\[ q^* = \frac{q_s}{\sqrt{(s - 1)gD^3}}, \tag{2} \]

where \( q_s \) is the sediment transport rate, \( s \) is the specific gravity of the sediment, \( g \) is
the acceleration due to gravity, and \( D \) is the particle diameter. The Shields number,
\( \tau^* \), (as described above) is represented by:

\[ \tau^* = \frac{\tau}{(s - 1)\rho g D}, \tag{3} \]

where \( \tau \) is the grain shear stress in real space, and \( \rho \) is the fluid density. The critical
Shields number quantifies how much shear is required to move an arbitrary particle
and is defined as:

\[ \tau_c^* = \frac{\tau_c}{(s - 1)\rho g D}. \tag{4} \]
where $\tau_c$ is the critical shear stress in real space. One method to estimate shear stress experienced by the grains, $\tau$, is disclosed in Wilcock (2001):

$$\tau = 0.052\rho(gSD_{65})^{0.25}u^{1.5}, \quad (5)$$

where $u$ is the mean flow (depth-averaged) velocity, $S$ is the surface slope, and $D_{65}$ is the 65th percentile grain size. Since the model proposed here is for a bed of uni-size grains, $D_{65}$ simply reduces to $D$. With the components of (1) defined, a fully-dimensional relationship can be derived in which $q_s$ can be isolated:

$$q_s = a\sqrt{(s-1)gD^3} \left( \frac{0.052\rho(gSD)^{0.25}u^{1.5}}{(s-1)gD} - \frac{\tau_c}{(s-1)\rho gD} \right)^b, \quad (6)$$

or in alternate forms:

$$q_s = a\sqrt{(s-1)gD^3} \left( \frac{(0.052\rho(gSD)^{0.25}u^{1.5} - \tau_c)}{(s-1)\rho gD} \right)^b, \quad (7)$$

$$q_s = q^*\sqrt{(s-1)gD^3}. \quad (8)$$

2.2 Parameter Estimation through Bayesian Models

The task of estimating model parameters is very common in scientific and engineering inquiries, and as a result statistical methods have been developed over the last century that allow scientists to observe a process and subsequently estimate the unknown parameter values. The key here is the word 'estimate,' since in these processes the parameters must be inferred from observation.

Recall, from Freshman Physics, the example of friction. The total amount of static friction, $F_f$ for an arbitrary object can be calculated as the product of the coefficient of friction, $\mu_s$, multiplied by the normal force, $N$, as follows: $F_f = \mu_sN$. Since
there exists no way to determine $\mu_s$ theoretically, one can experiment by measuring how much force can be applied to the object without moving it. Once the object has moved, the applied force has exceeded the frictional force. The goal, then, is to apply the maximum force possible without moving the object. When this maximum applied force has been determined, it is a trivial calculation to back-solve for the coefficient of friction.

In order to be more rigorous, we would repeat this process numerous times until we could be confident about the inferred value of $\mu_s$ given our observations. Invariably, and depending on our personal skill measuring force and weight, there will be spread in the data. This variability is indicative of several realities.

First, our ability to measure the applied force will vary with experimenter and equipment. Experimenter ‘A’ could be inexperienced with the laboratory procedures; or alternately Experimenter ‘B’ may be using antiquated equipment.

Second, the perception of when the object moved might vary. For example, what counts as movement—1 cm, 1 mm, 1 nm—and how well can we measure to the prescribed resolution?

Third, uncontrollable (or possibly unknown or unknowable) environmental factors may be affecting the experimental conditions adding variability to the results. For example, a surface may, to the naked eye, look uniform, but as we look closer (perhaps through a hand loupe or microscope) we will begin to see variability in the surface roughness. For each subsequent experiment are we placing the object on the same area as before—is the surface getting smoother with each experiment? These and potentially other conditions combine to add variability to the measurements that may have significant effects on the inferred values of $\mu_s$.

While these considerations with a coefficient of friction may seem somewhat contrived (since accurate measurement of the applied force and object weight are
easily obtained resulting in little variability) the extension to more complex physical processes, or ones where competing conceptual models have not been verified or where large amounts of measurement uncertainty are the best we can do—such as sediment transport—the considerations listed above become more meaningful.

Classical statistical methods, such as method of moments, or maximum likelihood (see Rice (1995) or Casella and Berger (2002) for a description of these methods), treat the model parameters as unknown, and fixed—that is, only one true value in nature describes it completely (e.g., 3.14159...). Cressie et al. (2009) points out that traditional curve fitting approaches to parameter estimation fail to discriminate errors due to measurement and those due to inaccurate conceptual models (model misspecification). In areas of research where uncertainty becomes problematic or worrisome, adopting a Bayesian approach to parameter estimation can benefit the analysis in several ways: (1) the parameters being estimated are assumed to be random variables and not fixed values; (2) uncertainty can be allocated and accounted for in numerous ways, such as lumped (gross residual error) or separated into, process and measurement error, through different modeling techniques; and (3) one property of Bayesian models is that the posterior distribution is a weighted combination of the observations and prior knowledge dictated by rules of conditional probability. This means that for processes which are difficult or expensive to measure—thereby resulting in few observations—a Bayesian treatment can leverage expert knowledge summarized in prior distributions of the random variables in a formal way. For a larger perspective on Bayesian models, Cressie et al. (2009) and Gelman et al. (2004) provide a more complete description of model properties and possibilities.

The model described in this research is an extension of basic principles of probability. The fundamental premise of Bayesian models is conditional probability—that is, to make decisions based on what we know (or what we do not know) and update
our knowledge with what we have just observed. This conditional thinking can be summarized, mathematically, as ‘A|B’ which represents the variable or event of A, given B. For example, given the observed forces required for movement in our static coefficient of friction example, what can we say about $\mu_s$, or, written mathematically: $\mu_s$|observed forces. The end-result of a Bayesian model is an updated view of the world that considers prior knowledge and new observations. The following sections summarize the foundations of probability for these types of models but are not intended to be complete proofs or descriptions.

2.2.1 Elementary Bayesian Concepts

Bayesian statistical modeling is derived from Bayes’ rule of conditional probability, which is:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$  
(9)

The events $B_1,...B_k$, are partitions of the sample space. The posterior probability of an arbitrary event, $B_j$, is conditional on some event $A$. The expression found in (9) is often referred to as Bayes’ Theorem. Over a continuous joint distribution, Bayes’ Theorem becomes:

$$[x|y] = \frac{[y|x][x]}{\int[y|x][x]dx} = \frac{[y|x][x]}{[y]},$$  
(10)

where the denominator is a normalizing constant. Alternately, (10) can be written:

$$[x|y] = c[y|x][x],$$  
(11)

where $c = (\int[y|x][x]dx)^{-1}$. The normalizing constant, $c$, can be dropped altogether since $c = \frac{1}{[y]}$ which, if $y$ is known, is fixed and the relationship can be expressed as a proportionality:

$$[x|y] \propto [y|x][x].$$  
(12)
To illustrate the utility in the context of applied problems, let us assume process $z$ is a function of some parameter $\theta$. A Bayesian model can be written as:

\[
\frac{[\theta|z]}{\propto} \frac{[z|\theta]}{[\theta]}.
\]  

(13)

Were we to observe, or collect data, on the process $z$, the model in (13) tells us that it would be controlled by some underlying parameter $\theta$. The posterior distribution, $[\theta|z]$, describes the distribution of values that $\theta$ can assume, given the observations of $z$. It is the posterior distribution that allows us to make inference on the model parameters and thereby establish credible intervals of values for $\theta$. The right-hand side of (13) is comprised of two parts: the likelihood and the prior. The likelihood, $[z|\theta]$—also referred to as the data model—describes the distribution of the observations, given $\theta$. The likelihood should describe the structure of the process, $z$, given the fact that it is controlled by a parameter $\theta$. Lastly, the prior distribution, $[\theta]$, encompasses what is known about $\theta$ before considering our new observations—it can be the summary of all previous research in the literature on $\theta$, it can be an educated guess, or it can simply limit possible ranges of values for $\theta$. For example, if we know that the parameter $\theta$ cannot take on values less than or equal to zero or greater than some value, $\theta_{lim}$, then we can impose these constraints in our prior. Alternately, if the literature indicates that $\theta$ can take on any value (positive or negative), then perhaps we give it a normal distribution centered around what we think the mean should be. In specifying a prior distribution, an expert has the ability to inform the model—indeed, independent of the newly gathered observations—with some degree of specificity. If an expert feels quite certain that $\theta$ should be centered at the value 5 and that it varies slightly, then the expert might specify $\theta \sim N(5, \sigma^2)$ where $\sigma^2$ is a small value thereby creating an informative distribution centered at 5. If the literature exhibits a larger spread in
estimated values of $\theta$, then it is possible for the prior to reflect this by increasing $\sigma_0^2$.

Figure 1 shows two such possibilities for $[\theta]$.

\[ \text{Possible Priors for Theta} \]

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{possible_priors.png}
\caption{Two possible priors for a hypothetical parameter $\theta$. Both priors have the same mean value, though the black prior is much more precise than the red prior.}
\end{figure}

This basic example can be extended to include a vector or parameters, $\theta$, such that:

\[ [\theta | z] \propto [z | \theta][\theta], \quad (14) \]

provided that either an appropriate multivariate prior, or individual prior distributions are specified for each element in $\theta$. 


2.2.2 Bayesian Sediment Transport Model

A Bayesian model for sediment transport can be written as:

\[
[\tau_c, \sigma^2 | \log(q_{s, \text{obs}})] \propto \left( \prod_{i=1}^{n} [\log(q_{s, \text{obs}, i}) | \tau_c, \sigma^2] \right) [\tau_c][\sigma^2],
\]

where, \( q_{s, \text{obs}} = (q_{s, \text{obs}, 1}, \ldots, q_{s, \text{obs}, n})' \). The model in (15) is identifiable by its three parts (though only the likelihood and prior require being specified directly):

\[
[\tau_c, \sigma^2 | \log(q_{s, \text{obs}})] \propto \left( \prod_{i=1}^{n} [\log(q_{s, \text{obs}, i}) | \tau_c, \sigma^2] \right) [\tau_c][\sigma^2].
\]

The model specified in (15) makes inference on multiple parameters, \( \tau_c \) and \( \sigma^2 \), given a set of observations, \( q_{s, \text{obs}} \), of sediment transport. The construction of the above model assumes independence between \( \tau_c \) and \( \sigma^2 \), thereby allowing us to take the product of the two distributions as their joint density.

While it is theoretically possible to determine analytically the form of the posterior distributions of Bayesian models, this determination is impractical in all but the most simple model formulations. In the case of this model, an analytical derivation of the posterior distribution was not attempted and the posterior distribution was determined using the sampling methods explained in the Computational Methods section of this report.

Irrespective of whether one chooses to derive the posterior analytically or not, the likelihood and prior distributions must be specified \textit{a priori}. The likelihood in this model is specified as a function of the governing equations, including the parameter \( \tau_c \), and additive noise (dictated by \( \sigma^2 \)) in log-space. Specifically, the likelihood is denoted as:

\[
\log(q_{s, \text{obs}, i}) | \tau_c, \sigma^2 \sim N(\log(q_{s,i}), \sigma^2),
\]
where each observation, \( i \), has a mean value of:

\[
q_{s,i} = q_i^* \sqrt{(s - 1)gD^3},
\]

and variance \( \sigma^2 \). Alternately, this can be written:

\[
\log(q_{s,\text{obs},i}) = \log(q_{s,i}) + \epsilon_i,
\]

where each observation is independent and identically distributed (i.i.d.):

\[
\epsilon_i \sim \text{i.i.d. } N(0, \sigma^2).
\]

The model in (17) says that the sediment transport process generally follows the governing relationship specified in (18), but with some variance. For this simple model, the variance combines natural variation of the process, measurement error, and model misspecification into one term.

The prior distributions for \( \tau_c \) and \( \sigma^2 \) must accommodate the physical realities of the parameters they represent. For example, \( \sigma^2 \) must have positive real support (negative values of variance are not reasonable), thus an inverse gamma distribution would be an appropriate selection as a prior, where the inverse gamma density is:

\[
P(\sigma^2|r, q) = \frac{1}{r^q \Gamma(q)} (\sigma^2)^{-(q+1)} \exp \left(-\frac{1}{r \sigma^2}\right),
\]

where \( \Gamma(\cdot) \) is the gamma function and \( 0 \leq \sigma^2 < \infty \). The prior specification for \( \sigma^2 \) in the sediment transport model would be:

\[
\sigma^2 \sim I.G.(r, q),
\]
where \( r \) and \( q \) are hyperpriors related to the mean and variance of an inverse gamma distribution by:

\[
E(\sigma^2) = \frac{1}{r(q-1)},
\]

\[
Var(\sigma^2) = \frac{1}{r^2(q-1)^2(q-2)},
\]

where \( r > 0 \) and \( q > 2 \). Given a prior mean \((\mu_{\sigma^2})\) and variance \((\sigma_{\sigma^2}^2)\) for the prior distribution of \(\sigma^2\) in (15), the corresponding values for \( r \) and \( q \) can be determined by:

\[
r = \frac{\sigma_{\sigma^2}^2}{\mu_{\sigma^2}(\mu_{\sigma^2}^2 + \sigma_{\sigma^2}^2)},
\]

and

\[
q = \frac{1}{\mu_{\sigma^2}r} + 1.
\]

The prior distribution for \(\tau_c\), like that of \(\sigma^2\), also has physical constraints that need to be captured in the specification. First, \(\tau_c\) must be positive; and second, if observations of sediment transport are made, then \(\tau_c\) is necessarily less than the minimum shear at which transport was observed. Intuitively, the lowest flow at which sediments are moving obviously induces grain shear greater than critical, otherwise the grains would not move. Given these constraints, a truncated normal distribution was selected for the prior on \(\tau_c\). The density of the truncated normal is:

\[
P(\tau_c|\mu_{\tau_c}, \sigma_{\tau_c}, \tilde{a}, \tilde{b}) = \frac{1}{\sigma_{\tau_c}} \phi \left( \frac{\tau_c - \mu_{\tau_c}}{\sigma_{\tau_c}} \right) \Phi \left( \frac{\tilde{b} - \mu_{\tau_c}}{\sigma_{\tau_c}} \right) - \Phi \left( \frac{\tilde{a} - \mu_{\tau_c}}{\sigma_{\tau_c}} \right),
\]

where \(\mu_{\tau_c}\) and \(\sigma_{\tau_c}\) are location and shape parameters, \(\tilde{a}\) is the lower-bound, \(\tilde{b}\) is the upper-bound, and \(\phi()\) and \(\Phi()\) are respectively the probability density function and cumulative density function of the standard normal distribution. The prior for \(\tau_c\) is
therefore:

\[ \tau_c \sim \text{N.}(\mu_{\tau_c}, \sigma_{\tau_c})^\frac{1}{2}, \]

where the lower- and upper-bounds for the support can be set appropriately.

### 2.3 Data Simulation

In order for a model to be useful, it must be verified and validated. Models are verified when the code has been checked for errors and the mathematics and logic that characterize the model are executed correctly. A validated model means that it adequately characterizes the phenomenon of interest. For the current research, a validated Bayesian model will be able to accurately infer the true distribution of the modeled parameters, \( \tau_c \) and \( \sigma^2 \). Since it is impossible to measure these parameters directly, we cannot know for certain if the model accurately infers these values using measured observations. In addressing this problem, The National Science Foundation \((Oden et al., 2006)\) discusses the importance of using simulation studies in model validation. Simulated 'observations' make it possible to judge model validity by generating synthetic data using the governing equations and predetermined parameter values. Since the simulated data originate from a known parameter set, estimated distributions from the Bayesian model's posterior distribution can be compared to the known values to check for accuracy and precision. In the case of this model, the parameters listed in (7) and (20) (in particular, \( \tau_c \) and \( \sigma^2 \)) are assigned values and a set of simulated observations can be generated.

#### 2.3.1 Data Simulation Algorithm

Although the process of simulating data simply involves assigning parameter values to use in the governing equations, care should be taken to ensure that plausible values are selected. Contained in (7) is a complete formulation of the governing
equations—(7) is reproduced here for reference:

\[ q_s = a \sqrt{(s - 1)gD^3} \left( \frac{(0.052\rho(gSD)^{0.25}u^{1.5} - \tau_c)}{(s - 1)\rho g D} \right)^b. \]

Given all the parameters listed in the governing equation above, we must pick reasonable values in order to generate synthetic data. The parameters \( a \) and \( b \) are model coefficients used in excess shear models, and one formulation in particular—Meyer-Peter Müller (1948)—defines \( a = 8 \) and \( b = \frac{3}{2} \). The specific gravity \( s = 2.65 \) is sensible for quartz density sediments. The water surface slope, \( S \) can take on a range of values, with a slope of 0.002 being a modest value. Water temperature varies in fluvial systems with season, and a realistic assumption of 5°C results in a convenient fluid density, \( \rho \), of 1000 kg/m\(^3\). The effect of gravity, \( g \), is assumed to be constant at 9.81 m/s\(^2\). The designation of grain diameter, \( D \), should be appropriate for a gravel-bed river or flume study since the value of \( D \) will directly influence the velocities, \( u \), required to produce transport.

The governing equations indicate that increased velocity results in increased shear, and increased shear amplifies transport. Smaller grain diameters are, therefore, transported at lower velocities. The selection of any grain size in particular is arbitrary, though bed load transport in nature is largely relevant to gravel-bed rivers with grain diameters greater than approximately 8 mm. A set of \( n \) velocities at which we wish to simulate transport are specified iteratively with attention to the magnitude of transport rates. The goal of simulation is to produce a sensible set of transport ‘observations’ from justifiable model parameters and not to produce extreme or unrealistic transport rates. With this in mind, the designated velocities should be adjusted accordingly to produce physically-plausible transport rates.
Lastly, in order to simulate data we must provide realizations for our ‘true’ parameters, \( \tau_c \) and \( \sigma^2 \). These are the parameters the Bayesian transport model will estimate. Numerous deterministic approaches to estimating critical shear from measurable conditions have been published in the literature, with the first being Shields (1936) followed by approximations to Shields’ data in Brownlie (1981) in which the following relation was proposed for the critical Shields Number:

\[
\tau_c^* = 0.22R_{ep}^{-0.6} + 0.06\exp(-17.77R_{ep}^{-0.6}),
\]

where \( R_{ep} \) is the particle Reynolds Number:

\[
R_{ep} = \frac{\sqrt{g}RU^2}{\nu},
\]

where \( R \) is the submerged specific gravity \((R = s - 1)\) and \( \nu \) is the kinematic viscosity of water.

The relation in (29) provides us with a physically-based method to specify a plausible \( \tau_c \) using the previously-designated parameter values proposed in Table 1.

Prescribing values for \( \sigma^2 \) is more direct. Values of \( \sigma^2 \) ranging from 0 to approximately 2 provide realistic-looking scatterplots of shear versus sediment discharge with varying degrees of data scatter.

Table 1 lists the required parameter set and suggests some potential values. While some values—such as \( a \) and \( b \)—are specified directly, others—like density and kinematic viscosity, \( \rho \) and \( \nu \)—are functions of temperature, \( T \). Although \( T \) is not

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>s</th>
<th>S</th>
<th>T, °C</th>
<th>( \rho ), kg/m³</th>
<th>( \nu ), ( 10^6 ) m²/s</th>
<th>g, m/s²</th>
<th>D, mm</th>
<th>u, m/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>3/2</td>
<td>2.65</td>
<td>0.002</td>
<td>5</td>
<td>1000</td>
<td>1.519</td>
<td>9.81</td>
<td>8.0</td>
<td>{u : 0.5...3.0}</td>
</tr>
</tbody>
</table>

Table 1: Sample values for fixed parameters.
present in (7) a reasonable value must still be specified due to other parameters' dependence.

Figure 2 presents the data simulation algorithm graphically to summarize the process. The final step in data simulation is to simply verify that the simulated values represent plausible outcomes—this is performed by plotting the data in dimensionless space and evaluating the results.

2.4 Computational Methods

The objective of Bayesian modeling is to update our knowledge regarding the selected process by evaluating the posterior distribution and marginal posterior distributions. To calculate these distributions, however, requires the knowledge of the normalizing constant expressed in (11). For the sediment transport model, the constant would be determined by solving:

$$c = \left( \int \int \int [\log(q_{s,obs})|\tau_c, \sigma^2]d\tau_c d\sigma^2 \right)^{-1}. \tag{31}$$

For all but the simplest models, the determination of this normalizing constant is intractable and another method must be used. Markov Chain Monte Carlo (MCMC) methods provide a means to simulate realizations from a target (the posterior) distribution, thereby circumventing the need to analytically solve for the normalizing constant. The origins of this numerical method derive from two papers in particular: Metropolis et al. (1953) and Hastings (1970). Markov Chain Monte Carlo has been adopted by Bayesian statisticians because it allows sampling from otherwise intractable mathematical expressions for the posterior distribution. For more detailed introductions to MCMC, see Gelman et al. (2004) and Robert (2007).
Select values for fixed parameters

\[ a, b, s, g, D, S, T, U \]

Calculate intermediate values

\[ \rho = f(T) \]
\[ \nu = f(T) \]
\[ R = s - 1 \]

Calculate Critical Shields Number

\[ Re_p = \frac{\sqrt{gRD}}{\nu} \]
\[ \tau^*_c = 0.22 Re_p^{0.6} + 0.6 \times exp(-17.77 Re_p^{0.6}) \]

Calculate Shields Number

\[ \tau = 0.052 \rho / (gSD_{av})^{0.25} U^{1.5} \]
\[ \tau^* = \frac{\tau}{(s-1)gD} \]

Calculate transport rates

\[ q^*_t = s(\tau^* - \tau^*_c)^6 \]
\[ q_s = \sqrt{(s-1)gD^3} q^*_t \]

Add noise to simulate observation

\[ \epsilon \sim N(0, \sigma^2) \]
\[ \log(q_{x,\text{sim}}) = \log(q_x) + \epsilon \]
\[ q_{x,\text{sim}} = q_x exp(\epsilon) \]
non-dimensionalize:
\[ q^*_{x,\text{sim}} = \frac{q_{x,\text{sim}}}{\sqrt{(s-1)gD^3}} \]

Return transport vectors

Fig. 2: Data simulation algorithm
2.4.1 Markov Chain Monte Carlo

The theory behind MCMC is to construct a Markov chain that has a stationary and ergodic distribution that coincides with the posterior (or marginal posterior) distribution. After an arbitrary number of iterations, the Markov chain will converge to the target distribution. Once a chain has converged on the target distribution, subsequent realizations from the chain coincide with the target distribution, and thus samples can be generated and collected to form the posteriors.

A Markov chain is a discrete time stochastic process, \( \{ x \} = \{ x_0, x_1, \ldots, x_k \} \) with the property that the value of any \( x_k \) is dependent on only the most recent value, \( x_{k-1} \). For the MCMC approach to work, several conditions must be satisfied. First, if \( x_0 \) is sampled from the target distribution, \( f(\cdot) \), then subsequent samples will also be from \( f(\cdot) \). Second, the target distribution must be stationary. And third, the chain must be ergodic, or insensitive to initial conditions.

Two common questions in the implementation of MCMC are: (1) how do we construct a Markov chain such that, once converged, its realizations coincide with the target distribution and that the conditions of stationarity and ergodicity are present and (2) how do we know when a Markov chain has converged? Fortunately, the theoretical aspects of these issues are, and have been, an active area of research from which we can enjoy practical benefits. Regarding the former, two notable methods have been developed to ensure convergence with the target distribution and satisfy the properties of stationarity and ergodicity: Metropolis-Hastings and Gibbs sampling.

Because a complete description of these methods is beyond the scope of this research, only the practical considerations relating to these methods will be presented. The following sections consist of a summary of research in MCMC methods and references to detailed explanations of these methods are provided appropriately. To start, Chib and Greenberg (1995) and Casella and George (1992) provide helpful
starting points for more-detailed descriptions of these algorithms. With reference to
the latter condition of convergence, methods for assessing convergence are discussed
in Gelman et al. (2004) and Robert (2007). These approaches will be summarized in
the following sections.

2.4.2 Metropolis-Hastings

The Metropolis-Hastings (M-H) algorithm was developed asynchronously by a
group of theoretical physicists and a statistician (Metropolis et al., 1953; Hastings,
1970). The algorithm facilitates sampling from a density, \( f(\cdot) \), under the assumption
that \( f(\cdot) \) is known up to a normalizing factor. Recall the generic construction of
a Bayesian model in that it is most-commonly presented as a proportionality, for
example:

\[
[\theta | z] \propto [z | \theta][\theta].
\]  

(32)

Note that (15) and (32) are expressed as proportionalities since the normalizing con­
stant has been dropped from the expression (see (10) through (12)). The algorithm
in this research performs a random walk through the parameter space and uses an
accept/reject rule in order to converge on the target density. The accept/reject com­
ponent of this algorithm relies on sampling proposed values for \( f(\cdot) \) from some easily-
sampled, non-symmetric, proposal or jump distribution, \( q(\cdot) \). If the proposed value
is likely from the target distribution, it is kept with some probability, otherwise it is
rejected. Formally, the algorithm can be summarized as follows:

1. Select a starting value, \( x_0 \).

2. For \( k = 1, 2, \ldots n \):

   (a) Given \( x_{k-1} \), generate a proposed value, \( x_* \) from the proposal distribution,

   \( q(\cdot) \), for example: \( x_* = x_{k-1} \), and add onto that a small perturbation, \( \epsilon_k \)
from a symmetric density (such as $\sim N(0, \sigma_k^2)$).

(b) Assign value for $x_k$:

$$x_k = \begin{cases} x_* & \text{with probability } \rho \\ x_{k-1} & \text{with probability } 1 - \rho, \end{cases}$$

where

$$\rho = \min \left( \frac{|x_*| q(x_{k-1}|x_*)}{|x_{k-1}| q(x_*|x_{k-1})}, 1 \right).$$

The logic presented above should be iterated $n$ times until a satisfactory number of samples are obtained for inference.

The variance of the random walk—$\sigma_k^2$ sometimes also called $\sigma_{\text{tune}}^2$—will influence how many samples are either accepted or rejected. Optimal sampling strategies described in Gelman et al. (2004) suggest that for this class of jump distributions, the desired acceptance rate is approximately $0.3/d$ where $d$ is the number of dimensions of the proposal distribution.

In effect, the process of sampling from a posterior distribution using the M-H algorithm arises by proposing potential values for the estimand, quantifying how likely that particular value is given the observations of the process and the prior information, and accepting the proposal with some probability.

### 2.4.3 Gibbs Sampling

While the M-H algorithm can be implemented for any valid Bayesian model, it relies on evaluating proposed values for $x$ from a proposal distribution. When the proposal density is much greater than the target density then the M-H algorithm will have a very low probability of acceptance even for valid values—distribution tails may be, therefore, difficult to resolve under certain proposal distributions.
et al., 2004). From this perspective, then, the Gibbs sampler provides an attractive alternative to the universal M-H algorithm, provided the full conditional distributions can be determined analytically. The Gibbs sampler is a special case of M-H where the accept/reject ratio reduces to 1, effectively accepting all proposed values for \( x \) (Gelman et al., 2004).

Suppose that for some dimension of parameters, \( p > 1 \), the random variable vector, \( \mathbf{\theta} \), can be partitioned into a set of elements, \( \{ \mathbf{\theta} : \theta_1, \theta_2, \theta_3, \ldots, \theta_p \} \). Further suppose that values can be simulated from each of the full conditional distributions:

\[
[\theta_j | \theta_{-j}^{k-1}, y],
\]  

(33)

where the index \( j \) denotes the \( j^{th} \) parameter, \(-j\) denotes the conditional absence of the \( j^{th} \) parameter, and \( k \) denotes the \( k^{th} \) sampling event of \( n \) total samples. For example:

\[
\begin{align*}
\theta_1^{k+1} & \sim [\theta_1 | \theta_2^k, \ldots, \theta_p^k, y] = [\theta_1 | .] \\
\theta_2^{k+1} & \sim [\theta_2 | \theta_1^{k+1}, \theta_3^k, \ldots, \theta_p^k, y] = [\theta_2 | .] \\
\theta_3^{k+1} & \sim [\theta_3 | \theta_1^{k+1}, \theta_2^{k+1}, \theta_4^k, \ldots, \theta_p^k, y] = [\theta_3 | .] \\
& \quad \vdots \\
\theta_p^{k+1} & \sim [\theta_p | \theta_1^{k+1}, \theta_2^{k+1}, \ldots, \theta_{p-1}^{k+1}, y] = [\theta_p | .]
\end{align*}
\]  

(34)

If these conditions are satisfied, then \( \{ \mathbf{\theta} \} \) is a Markov chain that coincides with the target distribution, \( \{ \theta \} \). The included material is not intended as a formal proof but merely a summary of previous research on MCMC and sampling methods. A more complete description is included in Casella and George (1992).

2.4.4 Sampling Algorithm for Transport Model

In order to fit the model presented in (15), we must decide on a computational
strategy, including which sampling methods to employ and where. There are two methods for sampling from the posterior: the Gibbs sampler, which requires the knowledge of full conditionals; and M-H, which requires a proposal distribution and tuning parameter $\sigma_k$. Because Gibbs sampling is generally more efficient than M-H, due to the acceptance ratio always being equal to 1, it is the preferred method. In order to implement a Gibbs sampler, however, the full conditional distribution, $[\theta|\cdot]$ must be calculated.

The model formulation presented in (15) consists of a normally-distributed likelihood, an inverse gamma prior for $\sigma^2$ and a truncated normal distribution for the prior on $\tau_c$. The concept of conjugate distributions is useful here in that the multiplication of the normal and inverse gamma distributions results in an inverse gamma distribution for $\sigma^2$ with updated parameters, $\tilde{r}$ and $\tilde{q}$. Since this is known, the full conditional distribution, $[\sigma^2|\cdot]$, can be determined analytically and sampled via Gibbs sampling. The case of critical shear, however, requires the use of M-H since conjugate distributions do not exist for the truncated normal.

The derivation of the full conditional distribution for $\sigma^2$ is as follows:

$$[\sigma^2|\cdot] \propto \prod_{i=1}^{n} [q_{s, obs,i}|\tau_c, \sigma^2][\sigma^2]$$
$$\propto \prod_{i=1}^{n} N(q_{s, obs,i}, \sigma^2) I.G.(\tilde{r}, \tilde{q})$$
$$\propto \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi \sigma^2}}e^{\frac{-1}{2}(\frac{q_{s, obs,i} - q_{s,i}}{\sigma^2})^2}\right)^{(\tilde{r} + q + 1)} e^{-\frac{1}{\sigma^2} \sum_{i=1}^{n} (q_{s, obs,i} - q_{s,i})^2 + \frac{1}{\tilde{r}}}(\sigma^2)^{-(q + 1)}$$

with the bottom line in (35) having the form of an inverse gamma kernel (compare to (21)), therefore:

$$[\sigma^2|\cdot] = I.G.(\tilde{r}, \tilde{q}),$$

(36)
where the updated parameters are calculated by:

\[
\hat{r} = \left( \frac{1}{\sum_{i=1}^{n} (q_{s,obs,i} - q_{s,i})^2 + \frac{1}{r}} \right)^{-1},
\]

and

\[
\hat{q} = \frac{n}{2} + q,
\]

where \( r \) and \( q \) are the hyperpriors from (25) and (26).

Because of the inability to analytically determine the full conditional for \( \tau_c \), the M-H algorithm will be adopted for the critical shear step in the MCMC sampling procedure as described previously. In order for the MCMC procedure to work, we must provide certain arguments, including the set of observations, \( q_{obs} \), the parameters for our prior distributions (known as hyperpriors): \( \mu_{\sigma^2} \) and \( \sigma_{\sigma^2}^2 \) (and their corresponding \( r \) and \( q \) parameterizations from (25) and (26)) for the variance parameter; and \( \mu_{\tau_c}, \sigma_{\tau_c}^2, \tilde{a}, \) and \( \tilde{b} \) for the critical shear parameter. The MCMC algorithm, then, is outlined as follows:

1. Select a starting value for critical shear, \( \tau_c(0) \).

2. For iterations, \( k = 1, 2, \ldots, N \):

   a) Calculate \( q_s | \tau_c(k-1) \) using (6).

   b) Calculate updated parameters, \( \hat{r} \) and \( \hat{q} \) using (37) and (38).

   c) Obtain \( k^{th} \) sample for \( \sigma^2 \sim I.G.(\hat{r}, \hat{q}) \).

   d) Sample \( \tau_{cs} \sim T.N.(\tau_c(k-1), \sigma_{\tau_c}^2 b)^{\frac{1}{\tilde{a}}} \).

   e) Calculate \( q_{ss} | \tau_{cs} \) using (6).

   f) Calculate the ratio, \( \rho \), via:
\[ \rho = \frac{[q_{ls}|q_{obs}, \sigma^2(k)][\tau_{cs}|\mu_{\tau_c}, \sigma^2]}{[q_{ls}|q_{obs}, \sigma^2(k)][\tau_{c}(k-1)|\mu_{\tau_c}, \sigma^2_{\tau_c}]} \]

(g) Accept/reject proposed value, \( \tau_{cs} \):

\[
\tau_c(k) = \begin{cases} 
\tau_{cs}, & \text{with probability } \rho \\
\tau_c(k-1), & \text{with probability } 1 - \rho.
\end{cases}
\]

This algorithm is repeated \( N \) times until an adequate number of samples are obtained from the target distribution. The total number of samples from the target distribution is equal to the number of samples, \( N \), less the number of iterations required for burn-in, \( N_{\text{burn-in}} \) (explained in following section). The samples from the target distribution can then be used to make inference on the parameters of interest: \( \tau_c \) and \( \sigma^2 \).

### 2.4.5 Considerations for Bayesian Simulation

Iterative numerical methods, while necessary in many modeling situations, present potential pitfalls if left unchecked. Assuming that a Markov chain has converged to the target distribution, parameter inference is based on the sample realizations from the simulated posterior—that is, by examining the marginal posterior distributions it is possible to establish upper- and lower-bounds for the parameters of interest. Because inference relies on samples from the posterior, as opposed to an analytic description, we must ensure that a sufficient number of samples have been collected after the MCMC algorithm has converged. To illustrate this, a simple example of Monte-Carlo integration is developed here. Let us assume that we wish to integrate the following function:

\[
f(x) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\}. \tag{39}\]
The function in (39) is the pdf of the standard normal distribution—that is, $X \sim N(0, 1)$. Now, in order to obtain the probability for an arbitrary range, we must integrate (39) over that range; for example $(0, 1)$. The analytical solution to (39) is intractable, and so a Monte-Carlo approach can be implemented as an approximation. A simple Monte-Carlo approach entails generating a set of uniformly-distributed samples from the limits of integration, $(0, 1)$, and evaluating (39) at those values. This will give us a set of densities which can be summed, and then averaged to provide an estimate of the area under the curve on that interval. Generally, the algorithm takes the following form:

1. let \( \{x_1, x_2, \ldots, x_n\} \) be uniformly distributed over limits of integration—$X \sim U(a, b)$.
2. obtain \( n \) samples from $X$.
3. evaluate $f(x_i)$ at $\{x_1, x_2, \ldots, x_n\}$.
4. compute average of density values $\frac{1}{n} \sum_{i=1}^{n} f(x_i)(b - a)$.

Returning to the example\(^1\) of integrating (39) from 0 to 1, we can write the integral as follows:

$$I(f) = \frac{1}{\sqrt{2\pi}} \int_0^1 \exp \left\{ -\frac{x^2}{2} \right\} dx,$$

$$\hat{I}(f) = \frac{1}{n\sqrt{2\pi}} \sum_{i=0}^{n} \exp \left\{ -\frac{x_i^2}{2} \right\},$$

where $I(f)$ is the true value and $\hat{I}(f)$ is its estimate. Now as the number of samples, \( n \), increases, so does the accuracy. As $n \to \infty$, $\hat{I}(f) \to I(f)$. Generally, Monte-Carlo integration is used to approximate moments of a posterior distribution.

\(^1\)adapted from (Rice, 1995)
Returning to MCMC sampling, an adequate number of samples may vary from one application to the next, though several thousand samples (after convergence) from a posterior is generally the guideline. One challenge, then, is to know when the chain has converged.

Because MCMC is an iterative method that is ‘seeded’ with a potentially arbitrary starting value, the initial value, as well as subsequent samples afterwards, may not be true samples from the target distribution, but an artifact of the initial value. As a Markov chain matures with each iteration, it moves progressively closer to the target distribution, regardless of where it started (the assumption of ergodicity). Some chains converge quickly, in only a few iterations, while others require a longer ‘burn-in’ period. One approach to assessing convergence is to provide the MCMC algorithm with multiple, dispersed starting values and monitor successive iterations to see if the chains converge on each other. Doing this will help inform two considerations: (1) it will provide insight to whether or not the constructed chain is ergodic, and (2) much like a second opinion, if each chain, though seeded at dispersed values, converges on the others, then one can look to that as confirmation that the target distribution has been located. Critical to these assessments are diagnostic trace plots which display the progress of the Markov chain with each iteration. Figures 3 and 4 are examples of trace plots.

Figure 3 illustrates the concept of dispersed chains converging on a single distribution in univariate space. On the far left, each square represents a different initial value. Six distinct chains were seeded and allowed to run. In Figure 3 we see quick convergence from the initial value to a common distribution. Figure 4 shows the same random walk but through the bi-variate parameter space. In this example, two parameters are being estimated and the MCMC algorithm searches the parameter space, evaluates the likelihood of various values for $\theta_1$ and $\theta_2$, and moves towards the
region with the highest likelihoods. Figure 4(a) shows the MCMC traces early on in the iterations. We see that the chains are immature as they have not yet converged on a common distribution. From this we can infer that any values that have been sampled up to this point should be ignored. Figure 4(b), however, shows a matured trace plot where the target distribution has been found by each of the chains. Recall from the previous discussion on Markov chains that once $\theta(k)$ is from the target distribution, subsequent samples, $\{\theta(k+1) \ldots \theta(N)\}$ will also be from the target distribution. So, this leaves for determination the point at which the sampled values derive from the target distribution. For the example presented in Figures 3 and 4, a reliable estimation for the burn-in period can conservatively be set at iteration 500.

Fig. 3: *Univariate trace plot with dispersed initial values.* Squares represent the dispersed initial values used for iteration 0.

In complex models, the time to burn-in may be substantially longer than illustrated in Figures 3 and 4. For these scenarios—and for those in which it is not
immediately obvious that the MCMC sequence has converged—it is helpful to have another measure by which convergence can be assessed. Gelman and Rubin (1992) present such a method that uses over-dispersed initial values (similar to what was done in Figures 3 and 4) and statistics from each individual chain as well as across all the chains. The ‘Gelman-Rubin’ statistic compares the within-chain variance to the across-chain variance. For a set of chains that have converged on the target distribution, the within-chain variance will equal the across-chain variance and, once the chains overlap in the trace plot and the respective variances are equal, then the chains are said to be converged. Due to the relatively simple model structure and fast convergence (assessed visually) for the Bayesian sediment transport model, the Gelman-Rubin statistic is not reported.

2.5 Model Evaluation

The previous sections have provided a basis on which a Bayesian model can be developed and solved. Following the implementation of such a model, it naturally follows that the modeler will want to see how well it actually worked. Characteristic
questions may include: ‘how well does this model fit the observations?’ or ‘which prior specification works better?’. These questions originate from two fundamental concepts: model checking and model comparison. As in any other modeling problem, Bayesian model evaluation is comprised of both quantitative metrics and expert judgment, as well be discussed in the following paragraphs.

Model checking addresses how well the designated model structure (including priors and likelihoods) makes predictions on the observed values. If we were to remove a single observation from the set of all observations, and subsequently make a prediction on the conditions for the removed observation, how close would the prediction be? What about using all collected observations and constructing a prediction interval for observations we have not yet made? These ideas are common to classical statistical methods, for example mean-squared error and cross-validation. Therefore, one goal in model evaluation is to determine how well the model fits the observed data.

Regarding model comparison, the quote “Essentially, all models are wrong, but some are useful”, attributed to George Box, provides insight into the type of reasoning that should be adopted when evaluating models. The notion that a complex system is understood to such a degree that a purely deterministic formulation can be worked out is unrealistic in many applied settings. Competing and mutually exclusive theories are not uncommon in practice because a set of theories seem to give valid results in varied scenarios. Because numerous models exist there is a need to objectively compare one model to another. In the context of what George Box wrote, while we cannot get any model perfect, we should try to identify which models are most helpful. For example, take the case of sediment transport. There exist numerous bed load relations that follow the form of an excess shear model: $q^* = a(\tau^* - \tau_c^*)^b$ though the values of $a$ and $b$ differ from one relation to the next. In this context,
having collected observations of sediment transport, we may want to know which bed load relation—that is, what values of $a$ and $b$—gives the best fit to our observations. In order to do so objectively, some standard for quantifying the model fit must be provided so that competing models can be compared.

The idea of model evaluation, and specifically model checking and model comparison, is developed in the following sections.

### 2.5.1 Model Checking

Several model checking principles are employed in this research, including sensitivity to the prior distribution, posterior predictive distribution checking, and evaluation of statistical inference on model parameters.

The first principle, sensitivity analysis, is common to many modeling disciplines in which model output is compared across any number of different model configurations to determine how sensitive the model output is to perturbations in the model inputs. This is often the approach to forward stochastic models (Monte-Carlo analysis) in which parameters are sampled from distributions (usually a normal distribution with some variance parameter) centered on the best guess for a point estimate, and are re-computed $n$ times, randomly sampling from the defined parameter distributions to give a distribution of solutions. In the context of a Bayesian model, model parameters are, by requirement, random and prior information in the form of a distribution must be specified. The model choice for the prior—as well as the choice of hyperpriors—will influence the posterior distribution to some degree. Quantifying the extent of this influence is the goal of sensitivity analysis.

Of course, prior distributions allow Bayesian models to still make inference in scenarios where few observations are collected and classical methods fail. As the number of observations increases, the posterior is progressively less-affected by prior
information to the point that the model may practically ignore the prior information altogether. This notwithstanding, one criticism of Bayesian models is that they are never truly objective due to the requirement of this prior distribution. This appraisal has, in part, given rise to the Objective-Bayesian branch of statistics in which the benefits of Bayesian analysis may still be enjoyed by employing prior distributions that do not—or at the very least, minimize—the influence of prior distributions.

The selection of the truncated normal prior distribution in the sediment transport model for $\tau_c$ is based on physical constraints of the process (for example, $\tau_c$ must be positive but less than the minimum shear at which transport was observed), but different parameterizations of the truncated normal distribution can yield quite different prior distributions. The degree to which the prior distributions can impact the posterior distribution should be evaluated. This is easily done using simulated data.

Since we know the true values used to generate the synthetic data, we can make sensible judgments on how the prior information might affect parameter inference. These considerations are particularly important when there are few observations to give the model—in such a scenario, prior information is weighed more heavily than if there are numerous observations. As part of this research, the posterior distribution was evaluated to assess its sensitivity to prior specifications of both $\tau_c$ and $\sigma^2$.

When selecting a prior distribution for a parameter, consideration must be given to the physical constraints of that parameter. The selection of the truncated normal distribution for the prior of $\tau_c$ was based on several physical reasons that were explained in the earlier in this report. Because of these constraints, other priors (such as a Gamma, for instance) will not be evaluated since it lacks the desired properties of having variable upper and lower bounds. The truncated normal distribution, as it turns out, is very flexible and can take on numerous different shapes, which re-
result from different hyperprior specifications. Figure 5 shows eight different shapes that are all truncated normal distributions. The truncated normal is specified as: 
\[
\tau_c \sim T.N.(\mu_{\tau_c}, \sigma_{\tau_c}^2 \tilde{b})
\]
The hyperpriors, \( \mu_{\tau_c}, \sigma_{\tau_c}^2, \tilde{b} \), and \( \tilde{a} \) control the shape of the probability density function as shown in Figure 5. The goal of a sensitivity study is to try different parameterizations of \( \tau_c \) and assess how they influence the posterior.

The same argument applies for the prior of \( \sigma^2 \), though, the inverse gamma distribution is not nearly as flexible in terms of the types of shapes it can assume. The prior for \( \sigma^2 \) is parameterized as: 
\[
\sigma^2 \sim I.G.(r, q)
\]
though, instead of specifying \( r \) and \( q \) directly, we will specify hyperpriors \( \mu_{\sigma^2} \) and \( \sigma_{\sigma^2}^2 \) for the mean and variance. These hyperpriors are then used to calculate \( r \) and \( q \) through (25) and (26).

The evaluation of posterior sensitivity to prior specifications can be viewed in two ways. First, if used carelessly we can view the prior as a potential liability. A hasty and thoughtless specification of priors may, in certain scenarios, unduly drive the posterior away from the true distribution. These scenarios might include those where few observations were collected, highly variable observations were made, or both. Situations in which there are many high-quality observations will be less prone to poor selection of priors. The second light in which we might consider prior specification is that of informing the model. Given a few noisy observations, how can appropriately specified priors bolster inference? Traditional methods may fail under the setting of a few, noisy measurements, but a Bayesian model can borrow strength from prior knowledge. The goal of the sensitivity analysis is to evaluate these contingencies.

Secondly, posterior predictive distribution (PPD) checks make it possible for us to illustrate how well the model fits the observations. The posterior predictive distribution is—as its name suggests—a set of predictions based on the posterior distribution. Recall that the posterior distribution in the sediment transport model
Fig. 5: *Flexibility of the truncated normal distribution.* The specifications of each prior follows: \( T.N.(\mu, \sigma, a, b) \).
is: $[\tau_c, \sigma^2|\log(q_{s,obs})]$. Given that our model parameters are random variables arising from distributions, predictions based on the posterior will also be a random variable from the posterior predictive distribution. Formally, the posterior predictive distribution is defined as the posterior distribution integrated over the parameters. In a general formulation:

$$[\tilde{y}|y] = \int [\tilde{y}|\theta, y][\theta|y]d\theta,$$

(42)

where $\tilde{y}$ are unobserved, but observable outcomes. In a multi-parameter model—the sediment transport model, for instance—the posterior predictive distribution is:

$$[\log(\tilde{q}_{s,obs})|\log(q_{s,obs})] = \int \int [\log(\tilde{q}_{s,obs})|\log(q_{s,obs}), \tau_c, \sigma^2][\tau_c, \sigma^2|\log(q_{s,obs})]d\tau_c d\sigma^2.$$

(43)

The PPD can be used to either make predictions of sediment transport for future observables, or to evaluate how well the model fits the actual process observations. We can ask ourselves: 'do the observations look reasonable under the posterior predictive distribution?' The calculation of (43) is easily done through composition sampling. Since the posterior distribution is a collection of samples from the MCMC iterations, we can easily draw samples for $\tau_c$ and $\sigma^2$, and use them to simulate posterior predictive realizations. Using the PPD, we can make predictions under the same conditions as our observations and compare the predictions to the observations. As another check, we could remove one observation and use the remaining observations to build a new PPD and then predict for the observation we removed. For the current research, both approaches were employed: the PPD was compared to the observations, and a 'leave-one-out' cross-validation algorithm was implemented. Using the cross-validation approach, the mean square prediction error (MSPE) was calculated. This metric is a measure of how well the model can predict a missing observation. Since the observation was simply removed prior to running the model, the prediction of the
removed observation can be compared directly. The value of MSPE is calculated as follows:

$$MSPE = \frac{\sum_{i=1}^{n} (y_i - E(y_i|y_{-i}))^2}{n},$$

(44)

where the quantity $E(y_i|y_{-i})^2$ is the posterior predictive mean for $y_i$ from a model fitted using all points except $y_i$. From an algorithm perspective, the calculation of this quantity requires running the model $n$ times—once for each point removed.

Lastly, we must ask ourselves if the inferred model parameters are plausible or realistic. This step is implicit in any model—we need to make sure that the model results are sensible given the physics and prior knowledge of the process. If posterior inference points towards physically impossible or highly implausible results, then the model assumptions—such as choice of likelihoods and priors, parameterizations of priors, and selection of governing relations—should be reconsidered.

### 2.5.2 Model Comparison

In order to make a comparison between an arbitrary number of models, some objective and quantifiable criterion must be established. Traditional statistical methods have criteria, such as Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), to provide a measure of model fit and parsimony (the heuristic of Occam's razor applies here: the simplest explanation is usually the correct one). The concept of Deviance Information Criterion (DIC) was developed in Spiegelhalter et al. (1998). As they point out in their paper, AIC and BIC cannot be directly applied to complicated Bayesian models (especially hierarchical) since the parameters can potentially outnumber observations. Accordingly, DIC was developed as a general criterion which could be applied to Bayesian models. The principle behind DIC is to define some measure of model fit (deviance) to model complexity (the number of parameters). The models in which DIC is minimized are those that balance
model fit while avoiding over-parameterized models. The absolute magnitude of DIC is irrelevant—it is the relative magnitudes between models that makes it significant. Deviance is defined as:

\[ D(y, \theta) = -2\log[p(y|\theta)], \]  

(45)

and is a measure of model fit. The value for DIC is calculated by:

\[ DIC = 2D_{avg}(y) - D_{\hat{\theta}}, \]  

(46)

where,

\[
D_{avg}(y) = E(D(y, \theta)|y) \approx \frac{1}{N} \sum_{i=1}^{N} D(y, \theta^i),
\]

\[
D_{\hat{\theta}} = D(y, \hat{\theta}(y)),
\]

\[
\hat{\theta} = E(\theta|y).
\]

(47)

The Deviance Information Criterion is convenient since it can be easily calculated using MCMC samples and does not require any analytic solutions.

2.6 Coding the Model

Incorporating all of the elements presented in this chapter represent a significant computer programming effort. The statistical software \(R\) (\textit{R Development Core Team}, 2009) was selected as the environment of choice due to its computational efficiency and wide availability of common statistical algorithms. The resultant code that was written for this model is included as Appendix A.
3.1 Simulation Studies

As explained in the Methods section, simulation studies are central to validating and evaluating the Bayesian sediment transport model proposed in this research. The process of simulating data includes selecting values for the fixed parameters—those parameters identified in Table 1—and subsequently, selecting reasonable values for our parameters of interest, $\tau_c$ and $\sigma^2$. Once we have chosen a value for grain size, $D$, (29) results in a value for $\tau^*_c$, thus we avoid needing to specify $\tau^*_c$ directly—it is implicit through a prospective relationship between grain size and critical shear\(^1\). Relating to $\sigma^2$, we can choose any positive value, with the condition that it produces realistic transport results. It is conceivable that a practical upper limit exists on the variability of sediment transport measurements. Accordingly, for simulation, the value of $\sigma^2$ will be adjusted such that it results in sensible transport events.

Having simulated observations, several questions relating to the proposed sediment transport model can be addressed. First, we can relate model performance—its ability to correctly infer $\tau_c$ and $\sigma^2$ from observations—to characteristics of the simulated data. For example, the quantity and quality of sediment transport measurements are often limiting, that is, it is difficult and expensive to collect many high-quality observations. Naturally, we want to assess how well the model performs

\(^1\)In practice, the use of equations such as (29) serve not to provide a definitive value for critical shear—rather, such equations are proposed relationships that, in this case, suggest a general guideline for reasonable parameter values that will result in realistic transport values. This same caveat also applies to (5). Future models may incorporate uncertainty in these relationships as well, but for now, they are accepted as valid relationships since, it is not the absolute values of $\tau_c$ that are sought after in the simulation study—we merely wish to minimize the difference between the estimated value from the model and the true value specified during data simulation.
using combinations of high- and low-variance measurements, each comprising many and few observations. Accordingly, four datasets were simulated:

1. many observations of low-variance (large $N$, small $\sigma^2$)
2. few observations of low-variance (small $N$, small $\sigma^2$)
3. many observations of high-variance (large $N$, large $\sigma^2$)
4. few observations of high-variance (small $N$, large $\sigma^2$)

Second, using any or all of the four datasets listed above, we can experiment using different prior parameterizations to see how the selection of hyperpriors—the mean and variance of our priors—influences posterior inference.

### 3.1.1 Simulated Observations

The synthetic observations were generated according to the Meyer-Peter Müller (1948) bed load relation: $q^* = 8(\tau^* - \tau_c^*)^{3/2}$. This model is an excess shear formulation with designated parameters $a = 8$ and $b = 3/2$. The other fixed parameters are specified according to Table 2. As mentioned, critical shear is a function of grain size, so by specifying $D, \tau_c$ can be determined according to the Brownlie relation, (29). In order to generate realistic sets of simulated observations, we must select two values for $\sigma^2$ that represent both a low- and high-variance dataset. These values were identified iteratively: $\{\sigma^2 : 0.05, 1.10\}$. Figure 6 shows plots of the simulated observations.
Fig. 6: Summary of simulated observation scenarios. Each data scenario is a distinct combination of high/low variance and many/few observations. (a) $N = 21$, $\sigma^2 = 0.05$, (b) $N = 6$, $\sigma^2 = 0.05$, (c) $N = 21$, $\sigma^2 = 1.10$, (d) $N = 6$, $\sigma^2 = 1.10$. 
3.1.2 Model Results using Simulated Observations

A series of twenty-four model fits were completed in order to characterize model performance. Given each of the four data scenarios just described, four distinct prior parameterizations for $\tau_c$ and two for $\sigma^2$ were evaluated.

As presented in Figure 5, a variety of information can be expressed in the prior distribution for $\tau_c$. For the purpose of a sensitivity analysis, four cases are of particular interest (see Figure 7). Each of the four prior specifications were used in model fits for each data scenario, resulting in sixteen marginal posterior distributions for $\tau_c$. The intent of these priors is to examine how posterior inference is affected when using: (a) a diffuse prior that has a lower-bound at zero and an upper-bound at the minimum value of shear at which transport was observed, (b) a diffuse prior on compact support (lower- and upper-bounds are narrowed) and whose mean is inaccurate, (c) an informative prior with an inaccurate mean and low variance, and (d) an informative prior whose density increases with proximity to the lowest measured shear.

Each data scenario has a true variance of either 0.05 or 1.10. To isolate the effect of changing priors for $\tau_c$, the prior for $\sigma^2$ was set at the known value for each data scenario with a small variance ($\mu_{\sigma^2} = \{0.05, 1.10\}$, $\sigma_{\sigma^2}^2 = 0.15$). For data scenarios 1 and 2, the prior for $\sigma^2$ has a mean of 0.05. For data scenarios 3 and 4 the prior mean is 1.10. Thus, both the high- and low-variance data scenarios will have accounted for the true variance. Figure 8 shows the two prior distributions for the high- and low-variance data scenarios.

Figures 9 through 12 show the marginal posterior distributions for $\tau_c$ for each prior distribution specified in Figure 7 and are grouped by data scenario. Tables 3 through 6 are tabular summaries of those same model fits. Trace plots and MCMC summary plots of these model fits are provided in Appendix B—see Figures 34 through
Fig. 7: *Four prior scenarios.* Red line denotes true \( \tau_c \) in simulated data scenarios. Parameterization follows: \( T.N.(\mu_{\tau_c}, \sigma^2_{\tau_c}, \tilde{a}, \tilde{b}) \).
Fig. 8: *Priors for $\sigma^2$: (a) data scenarios 1 and 2, (b) data scenarios 3 and 4. Vertical lines denote the expected value.*

Leave-one-out cross-validation plots for each data scenario are included as Figures 13 through 16. Table 7 summarizes the $\text{MSPE}$ values for each data scenario and prior. Common y-axes were intentionally specified to provide easy comparison of each prior scenario show in Figures 13 through 16. This results in high-variance predictions sometimes plotting out of range.

Lastly, Figures 17 through 19 show the results of a single model fit for data scenario 4 using an informed prior. Figure 17 shows the prior for $\tau_c$ that was used in this model fit. Figure 18 shows the summary information from the MCMC iterations, and Figure 19 shows the leave-one-out cross-validation plot with the estimated MSPE value.
Fig. 9: $\tau_c$ posterior sensitivity plot for data scenario 1. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.

<table>
<thead>
<tr>
<th>Data Scenario</th>
<th>Prior</th>
<th>E($\tau_c$)</th>
<th>95% C.I.</th>
<th>E($\tau_c$)</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>6.72</td>
<td>6.75</td>
<td>(6.71, 6.78)</td>
<td>3.46</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>6.72</td>
<td>4.17</td>
<td>(3.38, 4.45)</td>
<td>3.00</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>6.72</td>
<td>6.74</td>
<td>(6.70, 6.78)</td>
<td>3.00</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>6.72</td>
<td>6.75</td>
<td>(6.71, 6.78)</td>
<td>5.25</td>
</tr>
</tbody>
</table>

Table 3: $\tau_c$ posterior sensitivity table for data scenario 1.
Fig. 10: $\tau_c$ posterior sensitivity plot for data scenario 2. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.

<table>
<thead>
<tr>
<th>True Value</th>
<th>Posterior $E(\tau_c)$</th>
<th>95% C.I.</th>
<th>Prior $E(\tau_c)$</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scenario 2 Prior 1</td>
<td>6.72</td>
<td>6.67 (6.61, 6.71)</td>
<td>3.46 (0.17, 6.75)</td>
<td></td>
</tr>
<tr>
<td>Data Scenario 2 Prior 2</td>
<td>6.72</td>
<td>3.76 (1.95, 4.44)</td>
<td>3.00 (1.62, 4.38)</td>
<td></td>
</tr>
<tr>
<td>Data Scenario 2 Prior 3</td>
<td>6.72</td>
<td>3.30 (2.30, 4.32)</td>
<td>3.00 (2.02, 3.98)</td>
<td></td>
</tr>
<tr>
<td>Data Scenario 2 Prior 4</td>
<td>6.72</td>
<td>6.66 (6.61, 6.71)</td>
<td>5.25 (1.36, 6.88)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: $\tau_c$ posterior sensitivity table for data scenario 2.
Fig. 11: \( \tau_c \) posterior sensitivity plot for data scenario 3. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.

<table>
<thead>
<tr>
<th>True Value</th>
<th>Posterior</th>
<th>Prior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( E(\tau_c) )</td>
<td>95% C.I.</td>
</tr>
<tr>
<td>Data Scenario 3 Prior 1</td>
<td>6.72</td>
<td>6.62 (6.33, 6.79)</td>
</tr>
<tr>
<td>Data Scenario 3 Prior 2</td>
<td>6.72</td>
<td>4.23 (3.58, 4.45)</td>
</tr>
<tr>
<td>Data Scenario 3 Prior 3</td>
<td>6.72</td>
<td>4.46 (3.02, 6.22)</td>
</tr>
<tr>
<td>Data Scenario 3 Prior 4</td>
<td>6.72</td>
<td>6.62 (6.34, 6.80)</td>
</tr>
</tbody>
</table>

Table 5: \( \tau_c \) posterior sensitivity table for data scenario 3.
Fig. 12: $\tau_c$ posterior sensitivity plot for data scenario 4. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.

<table>
<thead>
<tr>
<th>True Value</th>
<th>Posterior $E(\tau_c)$</th>
<th>95% C.I.</th>
<th>Prior $E(\tau_c)$</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scenario 4 Prior 1</td>
<td>6.72</td>
<td>6.51</td>
<td>(5.95, 6.81)</td>
<td>3.46</td>
</tr>
<tr>
<td>Data Scenario 4 Prior 2</td>
<td>6.72</td>
<td>3.98</td>
<td>(2.61, 4.44)</td>
<td>3.00</td>
</tr>
<tr>
<td>Data Scenario 4 Prior 3</td>
<td>6.72</td>
<td>3.56</td>
<td>(2.49, 4.72)</td>
<td>3.00</td>
</tr>
<tr>
<td>Data Scenario 4 Prior 4</td>
<td>6.72</td>
<td>6.56</td>
<td>(6.10, 6.82)</td>
<td>5.25</td>
</tr>
</tbody>
</table>

Table 6: $\tau_c$ posterior sensitivity table for data scenario 4.
Fig. 13: Cross-validation plots for data scenario 1. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.
Fig. 14: Cross-validation plots for data scenario 2. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.
Fig. 15: Cross-validation plots for data scenario 3. (a) Prior 1, (b) prior 2, (c) prior 3, (d) prior 4.
Fig. 16: Cross-validation plots for data scenario 4. (a) prior 1, (b) prior 2, (c) prior 3, (d) prior 4.
| Data Scenario 1 |  
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | Prior 1 | 0.00144 | Prior 2 | 0.30769 | Prior 3 | 0.00121 | Prior 4 | 0.00144 |
| Data Scenario 2 |  
|                 | Prior 1 | 0.00041 | Prior 2 | 2.17043 | Prior 3 | 1.98017 | Prior 4 | 0.00043 |
| Data Scenario 3 |  
|                 | Prior 1 | 0.00574 | Prior 2 | 0.10415 | Prior 3 | 0.11183 | Prior 4 | 0.00579 |
| Data Scenario 4 |  
|                 | Prior 1 | 0.00107 | Prior 2 | 0.03357 | Prior 3 | 0.05887 | Prior 4 | 0.00135 |

Table 7: MSPE values for all data scenarios and priors.

<table>
<thead>
<tr>
<th>Prior</th>
<th>True Value</th>
<th>Posterior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$E(\tau_c)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.51</td>
</tr>
<tr>
<td>Figure 7(a)</td>
<td>6.72</td>
<td>6.56</td>
</tr>
<tr>
<td>Figure 17</td>
<td>6.61</td>
<td>6.61</td>
</tr>
</tbody>
</table>

Table 8: Comparison of priors for data scenario 4.
Fig. 17: Informed prior for data scenario 4. $\tau_c \sim T.N.(70, 30, 0, 6.92)$. 
Fig. 18: Summary plot for informed prior on data scenario 4.
Fig. 19: MSPE plot using informed prior on data scenario 4.
Like critical shear, the variance is also a parameter of interest and is inferred in the posterior distribution. Because the prior distribution for $\sigma^2$ is not as flexible as the truncated normal distribution, only two prior scenarios were evaluated for $\sigma^2$. The true variances for data scenarios 1–2 and 3–4 are respectively 0.05 and 1.10. When specifying a prior for $\sigma^2$, the goal is to try to center the prior distribution over a reasonable estimate of $\sigma^2$, and then give the prior enough variance so that the posterior distribution can make the needed adjustments. Accordingly, Figure 20, shows the these priors with their parameterizations.

![Figure 20: Two prior scenarios for $\sigma^2$. Parameterization follows: I.G.$(\mu_{\sigma^2}, \sigma_{\sigma^2}^2)$.](image)

To isolate the effects of changing priors on $\sigma^2$, only one prior for $\tau_c$ was used in the model fits—the diffuse truncated normal prior (see Figure 7(a)). Figures 21 through 24 show the marginal posterior distributions for $\sigma^2$ for both prior distributions specified in Figure 20 and are grouped by data scenario.

Tables 9 through 12 are tabular summaries of those same model fits. Trace plots and MCMC summary plots of these model fits are provided in Appendix B—see Figures 50 through 57 for reference.
Number of Observations = 21
True Variance = 0.05

Fig. 21: $\sigma^2$ posterior sensitivity plot for data scenario 1.

<table>
<thead>
<tr>
<th></th>
<th>True Value</th>
<th>Posterior $E(\tau_c)$</th>
<th>95% C.I.</th>
<th>Prior $E(\tau_c)$</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scenario 1 Prior 1</td>
<td>0.05</td>
<td>0.07</td>
<td>(0.04, 0.12)</td>
<td>0.1</td>
<td>(0.02, 0.41)</td>
</tr>
<tr>
<td>Data Scenario 1 Prior 2</td>
<td>0.05</td>
<td>0.35</td>
<td>(0.20, 0.61)</td>
<td>1.5</td>
<td>(0.47, 4.42)</td>
</tr>
</tbody>
</table>

Table 9: $\sigma^2$ posterior sensitivity table for data scenario 1.
Prior Density
True Value
Posterior Mean

Number of Observations = 6
True Variance = 0.05

Fig. 22: $\sigma^2$ posterior sensitivity plot for data scenario 2.

<table>
<thead>
<tr>
<th></th>
<th>True Value</th>
<th>Posterior</th>
<th>Prior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$E(\tau_c)$</td>
<td>95% C.I.</td>
</tr>
<tr>
<td>Data Scenario 2</td>
<td>Prior 1</td>
<td>0.05</td>
<td>(0.02, 0.12)</td>
</tr>
<tr>
<td>Data Scenario 2</td>
<td>Prior 2</td>
<td>0.79</td>
<td>(0.33, 1.91)</td>
</tr>
</tbody>
</table>

Table 10: $\sigma^2$ posterior sensitivity table for data scenario 2.
Fig. 23: $\sigma^2$ posterior sensitivity plot for data scenario 3.

<table>
<thead>
<tr>
<th></th>
<th>True Value</th>
<th>Posterior $E(\tau_c)$</th>
<th>95% C.I.</th>
<th>Prior $E(\tau_c)$</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scenario 3 Prior 1</td>
<td>1.1</td>
<td>0.91</td>
<td>(0.50, 1.62)</td>
<td>0.1</td>
<td>(0.02, 0.42)</td>
</tr>
<tr>
<td>Data Scenario 3 Prior 2</td>
<td>1.1</td>
<td>1.09</td>
<td>(0.64, 1.91)</td>
<td>1.5</td>
<td>(0.47, 4.42)</td>
</tr>
</tbody>
</table>

Table 11: $\sigma^2$ posterior sensitivity table for data scenario 3.
Fig. 24: $\sigma^2$ posterior sensitivity plot for data scenario 4.

<table>
<thead>
<tr>
<th></th>
<th>True Value</th>
<th>Posterior</th>
<th>Prior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E(\tau_e)$</td>
<td>95% C.I.</td>
<td>$E(\tau_e)$</td>
</tr>
<tr>
<td>Data Scenario 4 Prior 1</td>
<td>1.1</td>
<td>0.30 (0.11, 0.81)</td>
<td>0.1</td>
</tr>
<tr>
<td>Data Scenario 4 Prior 2</td>
<td>1.1</td>
<td>0.98 (0.41, 2.23)</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 12: $\sigma^2$ posterior sensitivity table for data scenario 4.
3.1.3 Model Selection

To verify the utility of DIC to discriminate between competing conceptual models, a simulated set of data using the Meyer-Peter Müller (1948) governing equation—

\[ q^* = 8(\tau^* - \tau_c^*)^{3/2} \]

were modeled using the model presented in Wong and Parker (2006): \( q^* = 4.93(\tau^* - \tau_c^*)^{1.6} \). Table 13 includes these results

<table>
<thead>
<tr>
<th>Model</th>
<th>Num. Obs.</th>
<th>True ( \tau_c )</th>
<th>Posterior ( E(\tau_c) )</th>
<th>95% C.I.</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP-M</td>
<td>20</td>
<td>6.72</td>
<td>6.73</td>
<td>(6.70, 6.74)</td>
<td>44.46</td>
</tr>
<tr>
<td>W-P</td>
<td>20</td>
<td>6.72</td>
<td>6.72</td>
<td>(6.52, 6.73)</td>
<td>61.43</td>
</tr>
</tbody>
</table>

Table 13: Deviance Information Criterion results for competing simulation models.

3.2 Flume Data from Smart (1984)

In 1984, Graeme Smart published a series of flume experiments that were run at the Laboratory of Hydraulics, Hydrology, and Glaciology of the Swiss Federal Institute of Technology. These flume studies consisted of sediment discharge measurements in steep, uni-size beds. The results of his study are used in this research to validate the proposed Bayesian sediment transport model. Included in his measurements are the sediment discharge \( (q_s) \) and Shields number \( (\tau^*) \). Table 14 summarizes relevant measurements from a set of experiments run on 10.5 mm sediment at a variety of slopes. Because his parameter notation is different than that established in this report, it is helpful to re-define his measured parameters. The columns presented in Table 14 correspond to the original table disclosed in G. Smart’s publication. Column (2) is the corrected water discharge; column (3) is the slope; column (5) is the corrected flow depth; column (6) divided by a sediment density of 2680 \( kg/m^3 \) yields \( q_s \); column (9) is \( \tau^* \); column (10) is termed the flow resistance factor. The flume width was 0.2 m and the necessary parameters (such as, velocity, for example) were
calculated using measurements in Table 14.

Figure 25 is a plot of the 10.5 mm Shields number versus Einstein transport parameter

![Figure 25](image)

*Fig. 25: Smart (1984) 10.5 mm transport data.*

### 3.2.1 Specification of Priors

In order for the model to work, a prior for both $\tau_c$ and $\sigma^2$ must be specified. A diffuse truncated normal prior with parameterization $T.N.(7.79, 5000, 0, 16.96)$ was assumed for $\tau_c$, and an inverse gamma prior, with a mean of 0.1 and variance of
Table 14: Summary of 10.5 mm flume observations from Smart (1984).

<table>
<thead>
<tr>
<th>$q_w, m^2/s$</th>
<th>$S$</th>
<th>$H, m$</th>
<th>$G_m, kg/m/s$</th>
<th>$\tau^*$</th>
<th>$C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column 2</td>
<td>Column 3</td>
<td>Column 5</td>
<td>Column 6</td>
<td>Column 9</td>
<td>Column 10</td>
</tr>
<tr>
<td>0.0241</td>
<td>0.105</td>
<td>0.029</td>
<td>1.50</td>
<td>0.172</td>
<td>4.8</td>
</tr>
<tr>
<td>0.0237</td>
<td>0.150</td>
<td>0.019</td>
<td>4.05</td>
<td>0.162</td>
<td>7.5</td>
</tr>
<tr>
<td>0.0246</td>
<td>0.200</td>
<td>0.037</td>
<td>9.15</td>
<td>0.424</td>
<td>2.4</td>
</tr>
<tr>
<td>0.0466</td>
<td>0.070</td>
<td>0.045</td>
<td>1.90</td>
<td>0.178</td>
<td>5.9</td>
</tr>
<tr>
<td>0.0476</td>
<td>0.100</td>
<td>0.048</td>
<td>4.20</td>
<td>0.270</td>
<td>4.6</td>
</tr>
<tr>
<td>0.0481</td>
<td>0.150</td>
<td>0.046</td>
<td>11.50</td>
<td>0.393</td>
<td>4.0</td>
</tr>
<tr>
<td>0.0489</td>
<td>0.200</td>
<td>0.059</td>
<td>24.00</td>
<td>0.665</td>
<td>2.5</td>
</tr>
<tr>
<td>0.0635</td>
<td>0.034</td>
<td>0.051</td>
<td>0.35</td>
<td>0.098</td>
<td>9.6</td>
</tr>
<tr>
<td>0.0678</td>
<td>0.070</td>
<td>0.054</td>
<td>3.40</td>
<td>0.213</td>
<td>6.6</td>
</tr>
<tr>
<td>0.0703</td>
<td>0.100</td>
<td>0.059</td>
<td>8.00</td>
<td>0.335</td>
<td>4.9</td>
</tr>
<tr>
<td>0.0712</td>
<td>0.150</td>
<td>0.057</td>
<td>20.40</td>
<td>0.485</td>
<td>4.3</td>
</tr>
<tr>
<td>0.0725</td>
<td>0.200</td>
<td>0.066</td>
<td>44.20</td>
<td>0.745</td>
<td>3.1</td>
</tr>
<tr>
<td>0.0848</td>
<td>0.036</td>
<td>0.066</td>
<td>0.70</td>
<td>0.135</td>
<td>8.4</td>
</tr>
<tr>
<td>0.0874</td>
<td>0.050</td>
<td>0.066</td>
<td>2.04</td>
<td>0.186</td>
<td>7.4</td>
</tr>
<tr>
<td>0.0869</td>
<td>0.074</td>
<td>0.052</td>
<td>4.90</td>
<td>0.219</td>
<td>8.6</td>
</tr>
<tr>
<td>0.0895</td>
<td>0.100</td>
<td>0.054</td>
<td>11.50</td>
<td>0.304</td>
<td>7.3</td>
</tr>
<tr>
<td>0.0946</td>
<td>0.150</td>
<td>0.072</td>
<td>29.15</td>
<td>0.611</td>
<td>4.0</td>
</tr>
<tr>
<td>0.0956</td>
<td>0.200</td>
<td>0.073</td>
<td>51.50</td>
<td>0.824</td>
<td>3.5</td>
</tr>
<tr>
<td>0.1100</td>
<td>0.070</td>
<td>0.072</td>
<td>7.10</td>
<td>0.286</td>
<td>6.8</td>
</tr>
<tr>
<td>0.1131</td>
<td>0.100</td>
<td>0.072</td>
<td>14.25</td>
<td>0.410</td>
<td>5.9</td>
</tr>
<tr>
<td>0.1162</td>
<td>0.150</td>
<td>0.074</td>
<td>37.95</td>
<td>0.633</td>
<td>4.7</td>
</tr>
<tr>
<td>0.1179</td>
<td>0.200</td>
<td>0.075</td>
<td>70.70</td>
<td>0.856</td>
<td>4.1</td>
</tr>
<tr>
<td>0.1294</td>
<td>0.070</td>
<td>0.078</td>
<td>7.45</td>
<td>0.308</td>
<td>7.2</td>
</tr>
<tr>
<td>0.1329</td>
<td>0.100</td>
<td>0.075</td>
<td>17.65</td>
<td>0.427</td>
<td>6.5</td>
</tr>
<tr>
<td>0.1362</td>
<td>0.150</td>
<td>0.073</td>
<td>47.45</td>
<td>0.617</td>
<td>5.7</td>
</tr>
<tr>
<td>0.1398</td>
<td>0.200</td>
<td>0.079</td>
<td>74.55</td>
<td>0.898</td>
<td>4.5</td>
</tr>
</tbody>
</table>
10 was assumed for $\sigma^2$. Figure 26 shows these prior distributions as well as their specifications.

![Prior distributions for Smart (1984) observations.](image)

3.2.2 Model Results

Three different deterministic models were evaluated using the Smart (1984) flume observations. In his paper, G. Smart proposes a sediment transport relation, which is included as (48). The Meyer-Peter Müller equation is also evaluated as a potential process model and is included again for reference as (49). Lastly, Wong and Parker (2006) propose another model for sediment transport as shown in (50). Each model is used as a process model that defines the mean value, $q_{s,i}$, in (17).

\[ q^* = 4 \left[ \left( \frac{d_{90}}{d_{30}} \right)^2 S^{0.6} C_T \tau_s^{0.5} (\tau^* - \tau_c^*) \right] \]  
\[ q^* = 8(\tau^* - \tau_c^*)^{3/2} \]  
\[ q^* = 4.93(\tau^* - \tau_c^*)^{1.6} \]
Table 15: *Comparison of process model performance on the Smart (1984) 10.5 mm flume observations—\( \tau_c \)*. Asterisks denote dimensionless values.

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
<th>( \tau_c ) Posterior Mean</th>
<th>95% C.I.</th>
<th>( \tau_c ) Prior Mean</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart</td>
<td>-7.58</td>
<td>14.40 (13.47–15.22)</td>
<td>8.48</td>
<td>(0.43–16.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.084* (0.078*–0.088*)</td>
<td>0.049*</td>
<td>(0.002*–0.096*)</td>
<td></td>
</tr>
<tr>
<td>Meyer-Peter Müller</td>
<td>25.87</td>
<td>13.86 (12.53–14.85)</td>
<td>8.48</td>
<td>(0.43–16.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.080* (0.072*–0.086*)</td>
<td>0.049*</td>
<td>(0.002*–0.096*)</td>
<td></td>
</tr>
<tr>
<td>Wong-Parker</td>
<td>25.98</td>
<td>7.98 (5.37–10.50)</td>
<td>8.48</td>
<td>(0.43–16.53)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.046* (0.029*–0.060*)</td>
<td>0.049*</td>
<td>(0.002*–0.096*)</td>
<td></td>
</tr>
</tbody>
</table>

Table 16: *Comparison of process model performance on the Smart (1984) 10.5 mm flume observations—\( \sigma^2 \).*

<table>
<thead>
<tr>
<th>Model</th>
<th>DIC</th>
<th>( \sigma^2 ) Posterior Mean</th>
<th>95% C.I.</th>
<th>( \sigma^2 ) Prior Mean</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smart</td>
<td>-7.58</td>
<td>0.04 (0.03–0.07)</td>
<td>0.10</td>
<td>(0.02–0.41)</td>
<td></td>
</tr>
<tr>
<td>Meyer-Peter Müller</td>
<td>25.87</td>
<td>0.14 (0.08–0.24)</td>
<td>0.10</td>
<td>(0.02–0.41)</td>
<td></td>
</tr>
<tr>
<td>Wong-Parker</td>
<td>25.98</td>
<td>0.14 (0.08–0.23)</td>
<td>0.10</td>
<td>(0.02–0.41)</td>
<td></td>
</tr>
</tbody>
</table>

Figures 27 through 29 show summary plots for MCMC iterations of the flume observations using the Smart, Meyer-Peter Müller, and Wong and Parker process models, including trace plots, posterior marginal histograms, and lastly, mean and credible interval values for both the prior and marginal posterior distributions. Figures 30 through 32 show the posterior predictive intervals for each of the three process models. Each posterior predictive distribution plot contains two representations. The plot on the left shows the 68%, 90%, and 95% posterior predictive intervals represented in varying shades of grey. The plot on the right shows the total amount of variability in sediment transport due to the variation of the critical shear parameter and the variance due to \( \sigma^2 \). Tables 15 and 16 are tabular summaries comparing the inference of each process model for both \( \tau_c \) and \( \sigma^2 \).
Fig. 27: Summary plot of Smart process model for 10.5 mm flume observations.
Fig. 28: Summary plot of Meyer-Peter Müller process model for 10.5 mm flume observations.
Fig. 29: Summary plot of Wong-Parker process model for 10.5 mm flume observations.
Fig. 30: Posterior predictive distribution of Smart process model for 10.5 mm flume observations.

Fig. 31: Posterior predictive distribution of Meyer-Peter Müller process model for 10.5 mm flume observations.
Fig. 32: Posterior predictive distribution of Wong-Parker process model for 10.5 mm flume observations.
4.1 Simulation Studies

The results of the simulation studies provide a practical opportunity to evaluate the utility of Bayesian statistical modeling. The obvious benefit to simulation studies is that the ‘answers’ are known—that is, when we synthesized observations from the governing relationships, we specified (in one way or another) values for the parameters we are trying to estimate. Upon fitting the model we are able to compare the posterior distributions for $\tau_c$ and $\sigma^2$ to their true value and make some interpretation as to how the quantity and variance of the simulated observations may have affected inference. Similarly, several different priors were specified for both critical shear and variance on each of the four data scenarios whereby we can also evaluate the effect (positive or negative) of different prior specifications.

Generally, the four data scenarios represent a set of observations that range from many low-variance observations, to few high-variance observations. In terms of simulated data, the variance parameter that was specified at the outset can represent a number of different realities. In the context of simulated observations it could represent natural variability. Poff et al. (1997) discuss how natural variability is crucial in the proper functioning of ecosystems. Specifically, they cite the importance of natural variability of streamflow on river eco-system health. Clement and Piegay (2005) also refer to this concept:

Variability is not measurement imprecision or experimental error, but is a fundamental attribute of environmental data, which merits specific approaches for assessment and understanding.
Thus, variability is an intrinsically important characteristic and should be considered when modeling natural systems.

The variance parameter of the simulated data could also represent our ability to accurately measure sediment transport. In this scenario, the assumption is that the measurement techniques are not epistemically biased, but simply have random error associated with them that have a mean of zero and some variance.

In the context of Clement and Piegay (2005), as well as the realities associated with sediment transport measurement, the variance parameter for the simulated data is best interpreted as the combined effect of natural variability and measurement error. Due to the simple nature of the specified Bayesian sediment transport model, these two sources of variability were not separated, though more complex models in the future may be able to allocate the variance accordingly.

One important modeling issue relates to the use of different process models in the Bayesian framework. Process models in this research refer to those governing relationships that result from mapping conceptual models into physically-based deterministic equations. The process model that was used to generate the simulated observations was the Meyer-Peter Müller (1948) (MP-M) equation. The process model used in the parameter estimation and prediction model was also MP-M. This means that we should interpret the variance parameter as natural variability combined with measurement uncertainty. In settings where the process model is unknown, the variance parameter discussed in the previous paragraph will inflate when a less-plausible process model is being used to inform the likelihood. Since the simulated data have a known process model for sediment transport, model misspecification will contribute no variance to parameter inference. The same concept holds true for the expression of grain shear, as shown in (5).

The role of prior distributions in sediment transport modeling can be a great
help. As shown in the results section, the truncated normal distribution provides a very flexible distribution with which we can inform the sediment transport model depending upon expert knowledge. Generally speaking, the results show that a diffuse truncated normal prior is a very safe assumption for naïve modeling. A small amount of contextual information on the measurements, however, may provide valuable information that can be coded into a transport model.

Take, for instance, an example wherein researchers are making sediment transport observations during the rising limb of the hydrograph—the period of time just before spring runoff occurs in snow-melt dominated river systems. If these researchers take their first measurement and find that nothing was transported during their first pass across the river, they can infer that critical shear has not been reached. If they continue measurements and, on day three, finally start seeing movement, they can be reasonably certain that their first measurements of transport occurred very near the threshold where critical shear has been exceeded. With this basic knowledge, the prior distribution specified in Figure 5(h) would be most appropriate way of specifying a prior for \( \tau_c \).

Now, if these researchers share their data with colleagues without providing any contextual knowledge, it would be safest for the uninformed colleagues to use the prior specified in Figure 5(a), which is a diffuse truncated normal distribution.

As another example, suppose the research team arrives two days late on the river to find that significant transport has been occurring for the past several days. They could either use a diffuse prior (Figure 5(a)) or make an informed estimate using the prior specified in Figure 5(d). These priors have the shear of the research team’s first measured transport event for an upper-bound and zero for a lower-bound, so they are not limiting any physically-plausible value for critical shear. The majority of the prior density, however, can be centered wherever the team wishes, based on their
expert opinion. The use of user-bounded priors (see Figure 5(b), (c), (e), and (f)), however, is a less-desirable option and should generally be avoided for reasons that will be explained shortly.

The selection of the four priors for critical shear was done to maximize the disparity between the information contained in each prior specification. This was done in an effort to see how prior information affects parameter inference. Prior 1, in Figure 7(a), shows a diffuse truncated normal distribution for critical shear that has a lower-bound of zero, and an upper-bound of the lowest shear at which transport was observed. This prior gives equal weight to any value of critical shear within those bounds. Prior 2, Figure 7(b), is a diffuse truncated normal prior on compact support using user-defined upper- and lower-boundaries. Prior 2 is a purposeful effort to try and mislead or 'drive' the model thereby forcing a desired result in the posterior. The upper-bound was set such that it was well-below the known true value of critical shear set at the time the data scenarios were generated. This prior will not allow the MCMC algorithm to sample values outside the bounds set in the prior specification. Prior 3, 7(c), is an informative truncated normal prior with low-variance and is centered well-below the true critical shear value. This prior has the same upper- and lower-bounds as Priors (a) and (d), so while this prior is still an effort to mislead the model it is entirely possible that, given enough good observations, the prior information will be overwhelmed by observational evidence that the resulting parameter inference will still be valid. Prior 4, 7(d), is an informative prior that exhibits the shape of an exponential distribution mirrored about the y-axis. In this prior, the density of critical shear increases with proximity to the upper-bound of the truncated normal distribution, making it very useful for scenarios in which the first measured transport event is very close to the critical shear threshold.

The priors specified in Figure 20 reflect two possibilities of the inverse gamma
distribution for prior information on the variance parameter. The inverse gamma distribution was selected for $\sigma^2$ because it is conjugate with the likelihood—a normal distribution—thereby allowing the Gibbs sampler to be employed in the MCMC loop for $\sigma^2$. As was mentioned previously, Gibbs samplers are more efficient than M-H, and thus the preferred method. The inverse gamma distribution, however, is less flexible and gets 'pressed' up against the lower-bound of zero in low-variance data scenarios.

Priors for the variance parameter can be specified in a number of ways. Empirical Bayesian approaches advocate the use of the newly-collected observations to inform the prior, while a more 'traditional' Bayesian approach would be to evaluate the data collection methods, personnel, river system, and other information to inform the prior independent of the newly-obtained observations. For simulated observations, the latter approach is obviously infeasible so, after the data were generated, the transport events in Figure 6 were evaluated and an educated estimate for the prior mean was produced along with a prior variance that attempted to reflect a more-general uncertainty about the variance parameter's true value.

To summarize, a total of six different priors were evaluated on the four data scenarios disclosed in the Results section of this report. Four different priors were specified for critical shear and two different priors for the variance parameter. As was mentioned briefly in the Results section, when the four different priors were evaluated for critical shear, a prior for the variance parameter still needed to be specified. In an effort to isolate the effects of changing priors for critical shear and determining their effects on parameter inference, the prior $\sigma^2$ was set to have a mean that was equal to the true value of the simulated data scenario with a small variance. The same strategy was employed when varying priors for $\sigma^2$—a diffuse prior was used for $\tau_c$ in all model fits.
4.1.1 Data Scenario 1

Data scenario 1 represents a situation often sought but rarely achieved in most research projects—abundant low-variance observations. The variance of this data scenario, which is being interpreted as natural variability combined with measurement error \(^1\) is extremely low. The plot in Figure 6(a) shows a very well-behaved curve of observations that deviate very little from the deterministic relationship dictated by MP-M.

Figure 9 summarizes the marginal posterior distributions for \(T_c\) using the four prior distributions mentioned above. Each posterior histogram shows the true value of \(T_c\) used in the data simulations, as well as the posterior histogram, posterior mean, prior density, and prior mean. In the case of data scenario 1, all of the prior mean values are far enough away from the posterior mean that they are not visible.

Priors 1, 3, and 4 provide effectively identical inference on critical shear in data scenario 1. The diffuse prior in Figure 9(a) and the informative prior in (d) have the exact same posterior mean and 95% credible intervals (see Table 3). The informative prior in (c) was an intentional effort to mislead the model, but what we see is that there was enough observational evidence to overwhelm the misinformation given in the prior. The posterior mean and 95% credible interval are almost identical to the diffuse and informative priors of (a) and (d). Priors 1, 3, and 4 provide very accurate and precise estimates of critical shear.

Prior 2, however, limits the Bayesian model’s ability to infer the true distribution of \(T_c\). The diffuse prior on compact support imposes user-defined upper- and lower-bounds on the possible values for critical shear. The physics of sediment transport dictate that a valid value for critical shear must have positive real support. Thus, the \(^1\)recall, that model misspecification is not an issue with the simulated data since the MP-M relation was used for both data simulation and parameter estimation and prediction
lower-bound for critical shear is set at zero for priors 1, 3, and 4. Now, when observations of sediment transport are made, the shear induced by the river is obviously greater than the amount of shear the particles are able to resist, so there is transport. In priors 1, 3, and 4, the upper-bound for critical shear is set by the minimum flow at which transport has been observed. Critical shear is necessarily below this minimum observed shear because otherwise there would have been no transport. Since priors 1, 3, and 4 use a zero lower-bound and an upper-bound as just described, the Bayesian model is able to accurately infer, based on the evidence, the correct critical shear value. Even prior 3, though the majority of the density would tend to mislead the model, does not adversely affect inference because the upper- and lower-bounds for critical shear are rigorously-defined. Prior 2 does not have this same property. The lower-bound of prior 2 was set at 1.55 and the upper-bound at 4.45. The true value of critical shear is 6.72. Since prior 2’s upper-bound does not contain the true critical shear, the model will not be able to explore that parameter space and evaluate the likelihood of those values above 4.45. What we see in Figure 9(b) is that the model is pushing all the posterior mass against the upper-bound—the model would like to explore the data values above 4.45, but the prior is preventing that possibility. The result of prior 2 is an inaccurate estimate of critical shear.

Figure 13 shows the results of a leave-one-out cross-validation scheme where each point in a data scenario is removed, the Bayesian statistical model is fit, and a prediction is made at the removed point. This process is repeated as many times as there are points and the cross-validation plot can then be constructed. The red points show the removed observations and the green triangles the predicted values. It is immediately recognizable that the MSPE plots in Figure 13(a), (c), and (d) are much better than (b) because the predictions in (b) are far off from the observed values and the whiskers are lengthened due to an inflated variance in order to compensate for
the poor mean predictions. The obvious reason for the poor predictions in (b) is that the diffuse prior on compact support prevented accurate parameter inference. Since the upper-bound of the prior associated with (b) truncated at 4.45, the true value of 6.72 was never evaluated. The resulting posterior mean value for critical shear, 4.17 (Table 3) infers that it takes a much lower shear to move sediment than what it does in the simulated reality of data scenario 1. The net effect of this under-estimation of the shear required to move sediment is a set of over-predictions of transport, which is shown in Figure 13(b).

The fact that it is possible to get very poor results from a model is undeniable. In practice, however, we must be able to rely on the values given to us by our model, but we have just seen that in a certain situation (our diffuse prior on compact support) the model did very poorly on both parameter inference and prediction. A legitimate concern, then, knowing when we can rely upon model results.

In the context of this model, we should avoid setting user-defined upper- and lower-bounds on the truncated normal prior for $T_c$ unless physical justifications exist for compact support. We saw in Figure 13(c) that even though the prior was intentionally misleading, the model made correct inference because the strength of the observations was enough to pull the posterior up to the correct parameter values. This was possible because the physically-rigorous upper- and lower-bounds were specified. The results from data scenario 1 tell us that if we have enough good data, our prior does not matter unless we are overly-restrictive in its specification.

We saw in Figure 13(a), (c), and (d), that the predicted values were reasonable given the observations. One of the model validation techniques mentioned in the Methods section of this report was to evaluate model predictions in the context of the observations. In other words, we must decide whether the predicted values are reasonable given our observations. The predictions in Figure 13(a), (c), and (d) are
reasonable given our simulated observations. The predictions in Figure 13(b) are not. Thus by evaluating the predictions against the observations it is possible to determine when a model can be relied upon and when it cannot.

Another way of determining the reliability of a model is by evaluating the marginal posterior histograms. Included in the appendix are summary plots for each of the 24 simulated data model fits. Figures 34 through 37 are the associated summary plots for data scenario 1. Figure 35 is a summary plot for the diffuse prior on compact support that resulted in very poor inference and predictions. Relative to the other model fits shown in Figures 34, 36, and 37, the histogram for critical shear shows obvious differences. The inferred value of \( \sigma^2 \) for this prior scenario is two orders of magnitude larger than the true value (and prior mean). The reason the variance parameter is so inflated is because it is trying to compensate for the fact that \( \tau_c \) has an upper-bound that is much too low. In order to have observations that are even somewhat realistic, the variance parameter inflates until it can encompass these observations. Comparisons of summary plots from one prior specification to another may provide useful information regarding model reliability.

Figure 21 and Table 9 show the results for the different prior specifications for \( \sigma^2 \). Because the shape of the inverse gamma is less-flexible than the truncated normal, it is difficult to interpret diffuse priors in terms of their potential effect on the posterior. Figure 21 shows the two marginal posteriors for \( \sigma^2 \), and we see that when the prior distribution is put in the general neighborhood of the true variance, as shown in Figure 21(a), the posterior distribution gets close to the true value. In Figure 21(b), the prior grossly over-estimated the true variance and as a result, the posterior distribution was inflated relative to the true parameter value. Figures 50 and 51 show the corresponding summary plots for the MCMC iterations of these model fits. There are no obvious defects in the trace plots or posterior histograms that suggest prob-
lems with the algorithm or model fit. In fact, the posterior distribution for critical shear is centered at the true parameter value, though the 95% credible intervals are wider in Figure 51 than in Figure 50. The 95% credible interval for $\sigma^2$ in Figure 50 contains the true parameter value while the credible interval shown in Figure 51 does not contain the true parameter value. Thus we see that it is more difficult to infer the correct variance parameter if the prior is misspecified.

4.1.2 Data Scenario 2

The observations in data scenario 2 are just as ‘good’ as those in data scenario 1—there are simply fewer (six observations instead of twenty-one). Figure 10 and Table 4 present the marginal posterior results for $\tau_c$ using data scenario 2.

Generally, the results from data scenario 2 are similar to those from data scenario 1, with a few exceptions. Figure 10(a) and (d) show the diffuse and informative prior distribution results. We see in the corresponding table entries that even with only a few observations, the use of these prior distributions result in adequate inference. Looking closely at the values in Table 4 we see that the marginal posterior mean values under-estimate the true value, and the corresponding credible intervals actually do not contain the true value (they fall short by 0.01). The diffuse prior on compact support shown in (b) still gives poor results due to the upper-bound constraint for the same reasons it did before. The informative prior in Figure 10(c) behaves very differently between data scenario 2 and data scenario 1, however. Previously, there were enough observations to overwhelm this misspecified prior, but with only six observations, their low variance notwithstanding, the marginal posterior distribution was unduly influenced by prior 3. The situation presented in data scenario 2 would benefit from a well-specified and informative prior. For example, a good prior for this data scenario would be that presented in Figure 5(h).
The cross-validation plots in Figure 14 show reasonable predictions in (a) and (d), but (b) and (c) grossly over-predict transport for a given shear stress. These over-predictions are due to the prior distributions under-estimating the force required to produce movement, therefore, a shear stress that would never produce transport under real conditions is thought to produce measurable transport due to this under-estimated critical shear. As before, it is very easy to tell by looking at the predictive plots in Figure 14 which models are reliable and which are not.

The behavior observed in Figure 22 and Table 10 for the $\sigma^2$ sensitivity analysis is similar to what we observed previously for data scenario 1. The posterior with the well-specified prior was able to correctly infer the true variance parameter though with a marginally wider credible interval than seen in data scenario 1 (compare 0.04–0.12 for data scenario 1 to 0.02–0.12 for data scenario 2.) The poorly-specified prior (prior 2) resulted in a marginal posterior distribution that over-estimated the variance parameter. In the previous data scenario, the abundance of observations was only modestly capable of overwhelming the misspecified prior resulting in a marginal posterior mean of 0.35. In the current data scenario, the reduction in observations resulted in a marginal posterior mean of 0.79—the true value in both of these cases was 0.05.

4.1.3 Data Scenario 3

Data scenario 3 returns us to a situation in which we have many observations, but the spread of these observations is greater than in data scenario 1. Figure 11 shows the marginal posterior distributions for $\tau_c$ for each of the four priors. Table 5 shows the results in a tabular format. As before, priors 1 and 4 result in almost identical inference—their marginal posterior means under-estimate the true value, but it is contained in the 95% confidence interval. Prior 2 still results in very poor inference,
and prior 3 is less influential on posterior inference than it was in data scenario 2, but it still significantly affects inference. Figure 15 shows the cross-validation plots for data scenario 3. The predictions in (a) and (d) are reasonable since the observations are contained in the bounds of the boxplots and the mean predictions follow the general shape of the observations, but the predictions in (b) and (c) are very poor. Once again, because of the undue influence imparted by the priors in (b) and (c), the posterior predictive capabilities are impaired as evidenced by the inflated variance (manifest in the wide whiskers of the boxplot) and consistent over-prediction of transport.

The effect of $\sigma^2$ priors is summarized in Figure 23 and Table 11. We see from the summary table that both prior parameterizations resulted in 95% credible intervals that contained the true mean. Recall that for data scenario 1, the misspecified prior (prior 2 of data scenario 1) resulted in a marginal posterior that did not contain the true variance parameter value. For data scenario 3, prior 2, the marginal posterior mean value is 1.09 where the true value is 1.10. In iterative simulations involving the specification of a prior for $\sigma^2$ it was observed that misspecified prior distributions on datasets that had low variance were much less successful in accurately predicting the true variance parameter than high-variance datasets with misspecified priors. Part of this may be due to the fact that each data simulation is only one realization of a dataset for a given set of parameters, but the consistency with which this occurred during numerous iterations of this research suggests that it is systematic. Another reason for this phenomenon may be due to the fact that an inverse gamma distribution with a mean value of 0.10 is pressed up against the y-axis at $x = 0$. Since the mean of a ‘low-variance’ prior must have the highest densities left of the mean (by virtue of the inverse gamma’s shape) the remaining mass in the tail is spread thinly over the support to the right. Contrast this with an inverse gamma distribution that has a
mean of 1.5—the distribution can take on a wider range of values and still maintain its shape.

4.1.4 Data Scenario 4

Data scenario 4 is the most challenging of the four data scenarios since the variance is high and the observations are few. Based on the results from the previous data scenarios, it is not difficult to imagine that data scenario 4 will also be most sensitive to the prior specifications.

Figure 12 and Table 6 show the marginal posterior distribution results for data scenario 4. The results generally mirror those of other data scenarios, but only with wider credible intervals. The cross-validation plot in Figure 16 displays a wide range of possible predictions given the few and noisy observations used in this model fit.

Figure 24 and Table 12 show the results for the specification of the variance parameter. We see, in Table 12, that a misspecified prior cannot be overcome by the six observations of this data scenario, but that the well-specified prior includes the true variance in its, admittedly wide, credible interval.

4.1.5 General Comments on Simulated Observations

Having reviewed all four data scenarios, we will now make some comparisons and look for trends. One unsurprising aspect of the model results is how the mean of the marginal posterior distribution for critical shear drifts to the left (towards the prior means) as we progress from data scenario 1 to data scenario 4. This is due to the observations carrying progressively less weight in comparison to the prior due to either increased variance, fewer observations, or in the case of data scenario 4, both. A general conclusion that can be drawn from this observation is that the data scenarios that can benefit the most from priors are also the most sensitive to prior
specification. On the other hand, it is in these scenarios where thoughtful priors can greatly aid parameter inference and transport prediction.

We have consistently seen the informative distribution (plot (c) in Figures 9 through 12) unduly influence parameter inference in all data scenarios except the first. From this we can see how a poorly specified prior can negatively affect parameter inference and prediction (the whole point of priors (b) and (c) in Figure 7 was to try to mislead model results and show how some sets of observations are unaffected by this). While it is important to recognize this possibility, it ignores a beneficial aspect of priors. An asset of Bayesian statistical modeling is that prior information can *lend strength* to parameter inference.

Recall the brief example of the researchers who are making measurements on the rising limb of the hydrograph. The first few tries to collect data were premature, and it was not until the third day when transport was finally observed. In modeling their observations, it was mentioned that an informative prior could be specified that increased in density with proximity to the first measured transport event. If these researchers were only able to collect a few noisy measurements (as in data scenario 4, for example) then their model fits would benefit from a thoughtful prior specification. Let us assume that the data these researchers collected comprise data scenario 4. Figure 17 illustrates an informative prior that would benefit these researchers’ analysis. Figure 18 is a summary plot of the MCMC iterations using this prior for data scenario 4 observations and Figure 19 shows the resulting cross-validation plot. Table 8 compares the three best-performing prior distributions for data scenario 4, including this newly specified prior in Figure 17. We see that the use of the informative prior presented in Figure 17 not only gives the most accurate posterior, but also the most precise.
Up to this point, the discussion has generally centered on the validity of parameter inference for $\tau_c$ and $\sigma^2$. Another aspect to Bayesian statistical modeling is prediction. For the simulated datasets, cross-validation plots were created and an estimate for MSPE was provided. The cross-validation plots provide predictions of observed transport events using all the observations (except the one being predicted) to construct a model from which a prediction can be made.

Looking generally at the estimated MSPE values for all of the data scenarios we can see a broad pattern. For data scenarios 2 through 4, the prior 1 and prior 4 MSPE values are generally lower by several orders of magnitude than the estimated MSPE values for priors 2 and 3. Also note that the MSPE values for priors 2 and 3 of data scenario 2 are the largest of any data scenario. Further, the MSPE value for prior 2 of data scenario 1 is larger than all of the other data scenario MSPE estimates for that prior. Data scenarios 1 and 2 were the low-variance priors, so the original data are very well described by the deterministic sediment transport equation with only a small amount of noise added in. Because of this well-defined relationship between the observations and the governing equation, any difference between the inferred value of critical shear and its true value will result in a large difference between the predicted and observed transport rates, which may explain why the low-variance data scenarios have the highest MSPE values for the priors that unduly influenced the posterior. It is expected that the MSPE values for data scenario 2 would be greater than those for data scenario 1 since the former has many fewer observations than the latter, and a cross-validation scheme which leaves out an observation and makes predictions based on the remaining observations will surely be more sensitive to datasets with few observations than datasets with many observations.

Another aspect of the MSPE estimates is that they are dependent on the number of observations that occur after the flattening of the sediment transport relation.
Take, for example, data scenario 1, prior 2. This model fit has an estimated MSPE value of 0.3077. Looking at Figure 13(b) we can see two distinct curves—one curve of the simulated observations, and another curve of the predictions. Starting with the smallest shear, we observe that the distance between the predicted values and the observed values decreases as the shear increases. After a value for $\log(\tau^*)$ of about -2.4, the distance between predicted values and simulated observations remains relatively constant. For data scenario 1, prior 2, we see a partition in the predictions and observations; at low shears, the difference between predicted and simulated observation is greater than the distance between the two at high shears.

Now, with this in mind, we look to Figure 14(b) and (c). Here we observe that four out of the six observations are below -2.4 on the x-axis. The elevated MSPE estimates then make sense because there are proportionally twice as many predictions and simulated observations below the -2.4 $\log(\tau^*)$ value where the distance between observed and predicted transport is much greater than at higher $\log(\tau^*)$ values. The MP-M relationship can be loosely broken up into a curve with two main tangents. There is an initial steep slope followed by a transition into a gentler slope (see Figure 6. For the cases of data scenarios 1 and 2, the inference provided by the model using a diffuse prior on compact support does not predict this general shape; it is more generally a single slope line. Because of this, any predictions after the flattening of the simulated observations will be closer to the simulated observations thereby resulting in a lower estimated MSPE value.

Because of this relationship between the estimated MSPE values and the location of the predicted and observed transport rates on the x-axis, MSPE values between data scenarios will not necessarily be comparable. On the other hand, estimated MSPE values within data scenarios (i.e., between different prior specification for the same observations) will be comparable.
4.1.6 Process Model Selection

In addition to the ability to make parameter inference and predictions while accounting for uncertainty and variability, Bayesian models make it possible to compare competing scientific theories. In the case of sediment transport, many bed load relations exist; each one having been developed because the extant models did too poor a job at explaining what was being observed. This reality was pointed out by Gomez and Church (1989) when they observed that we have more sediment transport relations than we do good datasets to test them.

The Methods section referred to a metric named the Deviance Information Criterion (DIC) that quantifies the balance between model parsimony and accuracy. For a given process model or governing equation, DIC can provide a comparison between competing theories to help scientists select the most appropriate relationship (the one that provides the best fit while minimizing the number of parameters). Using simulated observations that originate from the MP-M equation, two competing model fits were performed on the same simulated observations. The first model fit used the MP-M relationship as a governing equation for parameter estimation and the second used the W-P relation. This was done to illustrate and verify that the DIC value will be inflated for model fits where a less-appropriate process model is used in the likelihood compared to a lower DIC value for model fits where the ‘correct’ model is used in the likelihood.

Table 13 contains the results of this exercise. In this table we see that the MP-M model fit has a lower DIC value relative to the W-P model fit. To correctly use DIC values it is essential to know that it is the relative difference between two competing models that designates one model more appropriate than the other and not the absolute magnitude. In the context of this example we should expect the MP-M model fit to have a lower DIC value because we know that it is the correct
process model for the simulated observations. The information presented in Table 13 illustrates that this is case.

As we transition from discussing simulated observations—where process models and true parameter values are known—to laboratory flume data where nothing is known for certain, it is important to have earned confidence in the proposed framework and tools through demonstrations, as well as have experience in how to specify priors and evaluate competing theories so that a useful model can be implemented.

4.2 Smart (1984) Flume Data

The relative simplicity of the Bayesian sediment transport model disclosed in this report requires that, when using flume observations, several important considerations must be met.

First, the bed load model describes uni-size sediment transport—that is, only one grain size will comprise the flume bed. Of course, natural systems consist of a grain-size distribution in their beds. For now, a simple uni-size model will serve as a solid proof of concept prior to expanding the Bayesian transport model to accommodate a multi-fraction sediment transport problem.

Second, the bed should be planar. In many flume studies, the effects of bed forms—dunes, ripples, anti-dunes, and the like—are studied. Since the model formulation presented in this report does not explicitly account for bed forms, observations from a planar bed are most appropriate.

Third, since the overall aim of this research is to eventually progress towards a bed load discharge model for gravel bed rivers, this first simplified model should account for bed load transport (transport of large grains, saltating along the river bed) as opposed to suspended transport (sand and silt being transported above the bed; suspended by turbulent movements of the fluid). Bed load transport becomes
the dominating process when the grain size is greater than approximately 5 mm in diameter.

The Smart (1984) flume observations satisfy these three criteria and therefore provide a good opportunity to test the Bayesian transport model that has been developed in this research.

4.2.1 General Discussion on Flume Data

The laboratory flume observations reported by G. Smart in 1984 were motivated by poor prediction of the MP-M bed load relation in steep channels. Using steep flume channels of uni-size, planar beds, numerous experiments were run. One of the experiments used a 10.5 mm grain diameter gravel—the results of this experiment are included as Table 14. The flume observations are also plotted in Figure 25.

In order to fit the Bayesian model, we are required to provide prior information. For critical shear, we observed in data scenario 1, prior 1, that a diffuse prior results in valid inference for a dataset with many low-variance observations. We saw from data scenarios 2, 3, and 4 that a diffuse prior resulted in valid inference and that it was not until data scenario 4 (consisting of only a few noisy measurements) that an informative prior was helpful. Since there are twenty-six observations in Table 14, a diffuse truncated normal prior with parameterization \( T.N.(7.79, 5000, 0, 16.96) \) was specified. This prior is plotted in Figure 26.

To provide prior information for \( \sigma^2 \) we can only rely on vague contextual information. First, we know that the experiments were performed in a controlled laboratory experiment meant to minimize extraneous error (such as measurement error, for example) but the process is still somewhat stochastic. Second, we can evaluate the graph of sediment transport provided in Figure 14 to get a general idea of the magnitude of variance. Given that there are numerous observations, we know that the
marginal posterior will be influenced more by the observations than a vague prior, so the prior was specified such that it had a small mean with large variance. The prior for \( \sigma^2 \) is also plotted in Figure 26 and has a mean of 0.1 and a variance of 10.

One fundamental difference between the simulation studies discussed previously, and a laboratory flume study is that in the former, the process model was known. In the flume study the true process model is unknown and so it advisable to try multiple process models when relevant. Smart (1984) proposes a new bed load relation whose goal is to overcome the poor prediction of MP-M in steep channels. With this in mind, it will be helpful to try both the model proposed by Smart as well as the competing MP-M equation. Additionally, the W-P model is also a candidate process model since it is an updated version of the MP-M equation (Wong and Parker, 2006).

Each model fit has its own summary plot, included as Figures 27 through 29. Additionally, there are two tables that compare the marginal posterior inference for each of the estimated parameters using each process model (Tables 15 and 16.) Table 15 reports the inferred parameters for critical shear in two ways. The first is the dimensional result (in Pascals) and the second is a dimensionless quantity (denoted by an asterisk). Recall the dimensionless relations presented in (1) through (5) for these variables.

In the Smart and MP-M summary plots (Figures 27 and 28) we see better convergence on critical shear than for the W-P model (Figure 29). The Smart model has a very low estimated variance while the MP-M and W-P models have a larger variance. Also, the Smart and MP-M models seem to have converged on a different posterior distribution than the W-P. The 95% credible interval of critical shear inferred by the W-P model is (5.02–10.39) while the 95% credible intervals for the MP-M and Smart models are approximately (12.53–15.22). From this we see that process model selection will change the inferred values of critical shear—this concept
The DIC ranking for the three models are, in increasing order, the Smart model, MP-M, and W-P. The DIC value is a metric that allows us to compare the fits of different process models. Relying solely upon this criterion leads us to think that the Smart model is the best-suited model to describe these observations. The MP-M is only marginally better than the W-P model.

Figures 30 through 32 show the posterior predictive distribution for each of the model fits. When comparing the Smart model PPD to the MP-M and W-P model PPD's, it is immediately obvious why the DIC value and estimates for the variance parameter in the Smart model are so much smaller than those for the MP-M and W-P models; the Smart model is over-parameterized. In the Methods section, several properties of DIC were discussed, with one of them being that it balances the number of parameters with how well it fits the observations. This might cause us to wonder why, if the Smart model is over-parameterized, was the DIC value unable to resolve this. The key to understanding this apparent paradox is that in both process models, the Bayesian model is still only estimating \( \tau_c \) and \( \sigma^2 \). The Smart, MP-M, and W-P process models are included below (in order) as a reference.

\[
q^* = 4 \left[ \left( \frac{d_{90}}{d_{30}} \right)^2 S^{0.6} C^{1.5} \right],
\]

\[q^* = 8(\tau^* - \tau_c^*)^{3/2},\]

\[q^* = 4.93(\tau^* - \tau_c^*)^{1.6}.\]

One of the assumptions of this simple Bayesian model is that all of the other parameters in the deterministic relationships are well-known enough that we are treating them as fixed. Some of these parameters include channel slope, \( S \), grain size, \( D \), and
velocity, \( u \). The Smart model introduces a few new fixed parameters, including \( \frac{d_{90}}{d_{30}} \), and \( C \).

The quantity \( \frac{d_{90}}{d_{30}} \) is a measure of how close to uni-size the sediments are. Practically speaking, it is very difficult to get a large quantity of identically sized particles collected, so the uniformity of the sediment mixture is expressed as the ratio of the 90\(^{th}\) percentile grain diameter to the 30\(^{th}\) grain size diameter. For the 10.5 mm dataset reported in *Smart* (1984), this ratio is 1.34, with a perfectly uniform sediment mixture being equal to 1.

The \( C \) variable is termed the "flow resistance factor" and is a dimensionless quantity determined by the following relationship:

\[
C = \frac{u}{\sqrt{gHS}},
\]

where \( u \) is velocity, \( g \) is the acceleration due to gravity, \( H \) is the flow depth, and \( S \) is the channel slope. From this it is obvious that the variable \( C \) is simply a function of parameters that are assumed to be fixed. Now, we can take (51) and re-write it in the general excess shear model form: \( q^* = a(\tau^* - \tau_0^*)^b \) by defining the coefficient \( b = 1 \) and \( a \):

\[
a = 4 \left( \frac{d_{90}}{d_{30}} \right)^2 S^{0.6} C T^{-0.5}.
\]

Since all of the parameters in (55) are assumed to be fixed in terms of the Bayesian model, the DIC only sees two parameters as being estimated.

With this in mind, the DIC value for the Smart model is much lower than the other models because it fails to account for the over-parameterization occurring in the fixed parameters.

An annotated version the PPD for the Smart model (Figure 30) is included here:
Fig. 33: Annotated posterior predictive distribution of Smart process model for 10.5 mm flume observations. Circled points represent subsequent observations where the observation encircled in green has a lower grain shear than the observation encircled in orange.
The over-parameterized predictions do provide a better model fit, but not for any physically-based reasons; the Smart model is fitting to the variability of the observations. Take for example, the two points circled in green and orange in Figure 33. The general concept of an excess shear model is that as the amount of shear in excess of that required to move an object increases, the amount of transport will increase non-linearly. Looking at Figure 33 we see two observations highlighted. The point encircled in green shows an observation with a locally-high transport rate, while the observation encircled in orange shows a locally-low transport rate. To interpret the Smart model in the context of an excess shear model is paradoxical, as we will see. The interpretation of the Smart model is that, for these two observations, when you increase the amount of excess shear in this region (moving from the green observation to the orange observation), we can expect not an increase in transport, but a decrease. This interpretation of the Smart model contradicts the concept of an excess shear model, thus we should conclude that the Smart model is not appropriate despite having the lowest estimated variance, lowest DIC, and realistic parameter inference.

Now, the remaining models include the MP-M model and the W-P model. The inferred values for $\sigma^2$ from both models are very consistent (Table 16), but the inferred values for $\tau_c$ do not even overlap credible intervals. The fact that two different process models infer distinct parameter spaces using the same data is, at the very least, disheartening. On the other hand, when the process models are studied and it is recognized that the empirically-defined coefficients differ (by a magnitude of nearly 2 between MP-M and W-P), it becomes less-surprising. At this point, the determination of the most appropriate model must depend on which model makes the observations most reasonable under the PPD, and which model’s parameter inference makes the most sense. Because of the difference in the inferred values of critical shear between the MP-M and W-P models, it is helpful to understand what critical shear is.
The concept of critical shear is that, for a given bed in a river or flume, it is possible to define a value that quantifies how much stress is required to move the bed. The idea of critical shear has been the subject of criticism in the past, with a paper entitled “Do Critical Stresses for Incipient Motion and Erosion Really Exist?” by Lavelle and Mofjeld (1987). This paper questions the idea of whether or not a single threshold exists at all. The authors ask ‘what constitutes bed movement?’ and ‘for how long shall we watch for bed movement?’ and these considerations make the existence of a threshold semantic.

For example, should bed movement be defined as when the first particle moves? If yes, then how long should the bed be watched before we can conclude that bed movement is not occurring? If the first particle moving does not define bed movement, then what proportion of the bed must move to constitute transport? These questions are discussed in Lavelle and Mofjeld (1987) and as a conclusion they state that stochastic aspects of transport lead us to reason that the critical shear concept will not be sharply defined and that it is most appropriately described using statistical distributions. In the case of these the Smart (1984) flume observations and the MP-M and W-P process models, both are inferring statistical distributions, but they are practically distinct.

For gravel beds, typical critical shear values are generally estimated around 0.030 to 0.045 (Wilcock et al., 2009). In extremely steep channels it is understood that critical shear can increase over a low-slope channel by up to 2.5 times (Cao, 1986). Looking at the inferred values of the MP-M and W-P models with this in mind leads us to believe that the MP-M process model—whose dimensionless inferred critical shear values are 0.072–0.086—is more consistent with the literature than the W-P model whose inferred values are 0.029–0.060.

The PPD plots in Figures 31 and 32 also illustrate some differences in the shape
of the MP-M and W-P models. Figure 31 can be broken down into a model with
two distinct slopes, with a steeper slope at low shear values and a flatter slope at
high shear values. Because of this curve, it seems that the inferred model parameters
result in slight over-predictions of transport (many of the observed values are on the
lower-half of the MP-M PPD plot). The W-P model, on the other hand, displays
a flatter overall relationship in Figure 32 with a slight tendency to under-predict
transport. For both process models, however, the observations still look reasonable
under the PPD, with neither one really being much 'better' than the other. This
latter observation is confirmed by the fact that they have almost identical DIC values
(the MP-M model has a marginally lower DIC than the W-P model.)

Given the ensemble of evidence discussed in this section, the MP-M model is
given preference since it gives the most reasonable parameter inferences in addition
to providing reasonable predictions under the PPD.
CHAPTER 5
CONCLUSIONS

The research presented in this report proposes a framework in which sediment transport observations can be used for parameter inference, predictions, and process model discrimination while accommodating forms of uncertainty and variability. Developing useful sediment transport models presents numerous challenges; these challenges, however, are not singular to geomorphologists and engineers. Researchers from other disciplines, such as ecology, hydrology, and atmospheric and environmental science, face similar obstacles in measuring complex phenomena; accounting for uncertainty in measurements and conceptual models; representing outcomes in terms of risk and probability; and discriminating between competing scientific theories. The results presented in this report indicate that a Bayesian sediment transport model is a promising framework that provides solutions to modeling challenges.

5.1 Summary of Research

The Bayesian sediment transport model proposed in this research is based upon an excess shear sediment transport model which is coupled with a log-normal data model. This Bayesian model estimates two important parameters of sediment transport. The first parameter is critical shear and the second is variance. The estimation of these parameters is obtained through an application of Bayes’ theorem and Markov Chain Monte Carlo calculation methods. The resulting posterior distribution describes the range of values critical shear and variance can take on based on prior information and collected observations. The Bayesian model requires the specification of prior distributions in which expert judgment and opinion can inform the model. The prior distribution for critical shear is a truncated normal distribution on either
robustly-defined or user-defined support. A robust determination of prior support has a lower-bound of zero and an upper-bound equal to the minimum shear at which transport was observed. User-defined support can also be specified in instances where expert judgment warrants it. The variance parameter, which combines natural variability, model misspecification, and measurement error, was given an inverse gamma prior distribution. Model predictions were obtained through composition sampling of the posterior distribution, thereby resulting in a posterior predictive distribution. The PPD makes it possible to calculate credible intervals of sediment transport predictions. Discrimination of different process models was achieved through the implementation of the Deviance Information Criterion, which is similar in spirit to AIC and BIC in that it balances model fit to the number of parameters, but DIC is uniquely appropriate for Bayesian models.

The proposed model was tested against simulated observations for validation purposes. Four different data scenarios were evaluated, varying from many low-variance observations to few high-variance observations. Each data scenario was fit by the proposed model using four different priors for critical shear, and two different priors for variance. The sensitivity of posterior inference and prediction was evaluated for each data scenario and prior combination. Model results from simulated observation fits indicate that the proposed model performs well in most instances, except when intentional efforts to mislead or 'drive' model results were implemented. In model fits where diffuse or well-informed priors were specified, the inferred parameter values matched very closely to the true quantities. It was also shown that in data-poor situations (few observations with high-variance) valid inference can still be made through the use of thoughtful prior specification.

In addition to testing the model against simulated observations, where the process model and true parameter values were known, the model was tested against flume
observations from Smart (1984). These observations were made in steep flumes of uni-size sediment over planar beds. Parameter inference and predictions were obtained using three different process models with these observations. It was concluded that the Meyer-Peter Müller (1948) model performed the best of the three tested models.

5.2 Addressing Needs from the Literature

The scientific literature pertaining to river restoration and sediment transport contain numerous discussions about the weaknesses and shortcomings of available modeling methodologies. These discussions motivated the present research, and as shown in the Results and Discussion sections of this report, a Bayesian approach to sediment transport can address several of these shortcomings in a robust manner.

First, a Bayesian sediment transport model provides parameter estimates for critical shear that are congruent with philosophical interpretations of critical shear. As described in Buffington and Montgomery (1997), the specification of a single critical shear value is problematic, and the point at which sediments begin to move is a statistical problem. The recommendation of Buffington and Montgomery (1997) is that modelers should be more concerned with specifying defensible values for critical shear, given the methodological biases and uncertainties rather than a universal value. The model disclosed in this research allows us to make inference on critical shear as a random variable that arises from a probability distribution given the observations we have collected.

Second, environmental processes, and therefore our observations of these processes, are naturally variable. Clement and Piegay (2005) remark that variability associated with environmental data does not necessarily reflect measurement error or imprecision, rather, it is an intrinsic characteristic of natural phenomena that war-
rants the development of approaches for assessment and understanding. Wohl et al. (2005) also recognize the existence of natural variability in natural systems and Poff et al. (1997) emphasizes its importance. A Bayesian sediment transport model—inclusive of the model presented in this research—makes it possible for modelers to estimate the natural variance of a process, as well as measurement error and model appropriateness. The uni-size transport model from this report rolls natural variability, measurement error, and model misspecification into one variance term, thereby accounting for all three of these sources of variance. Future Bayesian models can be specified in a way such that these terms may be isolated into individual terms.

Third, a posterior intuitive assessment of model result uncertainty is subject to significant human error. The research performed by Tversky and Kahneman (1983, 1986) illustrate simple scenarios in which intuitive assessments of uncertainty and probability result in flawed decisions. By applying simple concepts of probability to the scenarios proposed in their research (the expected value, for example) the optimal decision is clear and unambiguous. In a like manner, an intuitive treatment of uncertainty in sediment transport measurements, natural variation, process model selection, and predictions should be avoided and replaced by a sound method based on laws of probability. A Bayesian sediment transport does not rely upon intuitive assessments of uncertainty; rather, Bayesian models are founded upon established laws of probability. The proposed Bayesian approach to sediment transport modeling is, therefore, not subject to the errors of human cognition in uncertain circumstances.

Fourth, Wohl et al. (2005) emphasize not only the need to explicitly recognize complexities and uncertainties in river restoration activities (of which, sediment transport plays a significant role) but that there is a ‘pressing need’ to use analytical techniques that can incorporate expert judgment and opinion. Bayesian models require expert knowledge in the form of prior distributions to inform the posterior distribu-
tion. As shown in the Results section of this report, the specification of a meaningful prior on a dataset with only a few, noisy measurements bolstered inference in a measurable way.

Fifth, Wilcock et al. (2003) discusses the importance of clearly communicating the uncertainty in model predictions so that management decisions can be better informed. The use of posterior predictive distributions from Bayesian models fulfills such a requirement. The cross-validation plots as well as the posterior predictive distribution plots from the Results section clearly define the credible intervals of model predictions. The ability of a Bayesian sediment transport model to define credible intervals of parameter inference and predictions can be a major asset to decision making.

5.3 Future Research

One of the reasons for developing a uni-size sediment transport model was to start simple and illustrate that a Bayesian sediment transport model is a valid tool. Further, it was important that the product of this research have the potential to address some of the needs communicated in the literature. The next step in the modeling of sediment transport is, therefore, to add complexity to the model commensurate with the realities of real river systems.

The model disclosed in this research is simple in that it is only valid for a planar, uni-size bed. This simplification obviously does not hold in rivers where the bed is composed of a distribution of sediment sizes. Thus, the next model should accommodate a sediment size distribution and utilize multi-fraction sediment transport observations for inference and prediction. Along with this concept of multi-fraction transport, appropriate hiding functions should be evaluated, as well as functions which serve to estimate the grain shear experienced on the sediments. Both of these
functions can be nested into the Bayesian transport model in such a way that individual variances can be associated with the proposed relationships. Further, the variance due to natural variability, model misspecification, and measurement error should be separated to provide helpful estimates of these quantities.

Given the success of the uni-size sediment transport model, as well as the needs that a Bayesian treatment fills, the development of a more realistic multi-fraction model will provide a much-needed analytical tool to sediment transport modelers and decision makers.
REFERENCES


# Set up model boundary conditions

thetas = list(a = 8, b = 3/2, s = 2.65, g = 9.81, D = 1.83e-3, S = 0.001, T = 5, rho = funcGetWaterDensity(5), U = c(seq(0,1.1,,25)), s2 = 0.15)

# Simulate data and plot

simobs = funcSimulateTransport(thetas)
plot(simobs$tstar/simobs$tcstar,simobs$qstar,log='xy', xlab='t*/t_c*', ylab='q*_s*', xlim=c(0.1,10)) plot(log(simobs$tstar),log(simobs$qstar))

# Specify prior parameters

priors = list(# Specify priors for tc
tc.mu0 = 0.045*(thetas$s-1)*thetas$rho*thetas$g*thetas$D, tc.var0 = 100000, tc.lower = 0, tc.upper = min(simobs$t),

# Specify priors for s2
s2.mu0 = .5, s2.var0 = 8
priors$tc.mu0.TN = priors$tc.mu0 + funcTNPrior(priors$tc.mu0, priors$tc.var0, priors$tc.lower, priors$tc.upper)

# Play around with potential priors

# Because our prior on tc is a Truncated Normal, it is a good idea to
# check out the prior distribution since it behaves weirdly when the mean
# value is close to either end of the distribution (when truncated)
x11()
funcPlotPrior(priors,simobs,thetas)
funcPlotPrior_s2(priors,simobs,thetas,breaks=2000)

### Adjust upper and lower bounds to match the 0.025 and 0.975 quantiles
### and then force a functionally uniform prior (or not!)
###
priors$tc.var0 = 0.025
tmp.s2 = 0.05
priors$tc.lower = qnorm(0.025, priors$tc.mu0,sqrt(tmp.s2))
# constrain lower bound to 0
if(priors$tc.lower<0){priors$tc.lower = 0}
priors$tc.upper = qnorm(0.975, priors$tc.mu0,sqrt(tmp.s2))
# constrain upper bound to the minimum tau at which transport
# was observed
if(priors$tc.upper>min(simobs$t)){priors$tc.upper = min(simobs$t)}
priors$tc.mu0.TN = priors$tc.mu0 + funcTNPrior(priors$tc.mu0, priors$tc.var0, priors$tc.lower, priors$tc.upper)

x11()
funcPlotPrior(priors,simobs)

###
### Alternately...
###
### Adjust tc.mu0 to force tc.mu0.TN to converge to tc.mu0 with
### lower bound = 0
###
x11()
# try guessing it un-aided
priors$tc.mu0 = 350
priors$tc.var0 = 100
priors$tc.mu0.TN = priors$tc.mu0 + funcTNPrior(priors$tc.mu0, priors$tc.var0, priors$tc.lower, priors$tc.upper)

funcPlotPrior(priors,simobs)
# Specify numerical/computational parameters

## MCMC

Posterior = funcMCMC(thetas, simobs, priors, mcmc, FALSE)

### DIC

DIC = funcCalcDIC(thetas, simobs, Posterior, mcmc)

### MSPE.hat

MSPE = funcCalcMSPEhat(thetas, simobs, priors, mcmc, units=FALSE)

### Posterior Predictive Interval

PPD = funcPostPred(thetas, simobs, Posterior, mcmc, units=FALSE)
### Summary Plots

```
x11()
funcCreateSummaryPlots(simobs,priors,Posterior,mcmc,DIC,thetas)
# Adjust default burn in
mcmc$n.burn = 200
funcSavePlot('Rfig_SummaryPlot')
```

### Posterior Predictive Interval Plots

```
x11()
funcCreatePPIPlots(PPD,simobs,mcmc)
funcSavePlot('Rfig_PPIPlot')
```

### MPSE Boxplots

```
x11()
funcCreateMSPEBoxPlots(PPD,simobs,MSPE,mcmc,FALSE)
funcCreateMSPEBoxPlots(PPD,simobs,MSPE,mcmc,TRUE,0.2)
funcSavePlot('Rfig_MSPEPlot')
```

### Posterior Predictive Interval Plots

```
x11()
funcCreatePPIPlots(PPD,simobs,mcmc)
funcSavePlot('Rfig_PPIPlot')
```

### MPSE Boxplots

```
x11()
funcCreateMSPEBoxPlots(PPD,simobs,MSPE,mcmc,FALSE)
funcCreateMSPEBoxPlots(PPD,simobs,MSPE,mcmc,TRUE,0.2)
funcSavePlot('Rfig_MSPEPlot')
```

```r
mcmc = function(thetas,obs,priors,mcmc,priorplots){
  a = thetas$a
  b = thetas$b
  s = thetas$s
  g = thetas$g
  D = thetas$D
  S = thetas$S
  T = thetas$T
  tc.mu0 = priors$tc.mu0
  tc.var0 = priors$tc.var0
  tc.lower = priors$tc.lower
  tc.upper = priors$tc.upper
  tc.mu0.TN = tc.mu0 + funcTNPrior(tc.mu0,tc.var0,tc.lower,tc.upper)
  s2.mu0 = priors$s2.mu0
  s2.var0 = priors$s2.var0
  n.samples = mcmc$n.samples
  tc.seed = mcmc$tc.seed
  s2.tune = mcmc$s2.tune
  n.obs = mcmc$n.obs
}
```
\[ [s2] \sim \text{I.G.}(rp, qp) \]
\[ rp = \frac{s2.var0}{(s2.mu0^2 + s2.var0)} \]
\[ qp = \frac{1}{s2.mu0*rp} + 1 \]

# create object for output of prior parameters
priorparams = list(tc.muO.TN = tc.muO.TN, tc.lower = tc.lower, tc.upper = tc.upper, rp = rp, qp = qp)

if(priorplots==TRUE){
  def = par(mfrow=c(1,2))

  hist(tmpx, xlab = 'tc', main='tc Prior', freq = FALSE)
  abline(v = tc.muO.TN, col='red', lwd=2)
  abline(v = tc.muO, col='blue', lwd=2)

  tmp = rgamma(1e3,qp,rp)^(-1)
  hist(tmp, xlab = 's2', main = 's2 prior', freq=FALSE)
  abline(v = obs$s2, col='red', lwd=2)
}

# Pre-allocate vectors
tc.save = rep(0, n.samples)
s2.save = rep(0, n.samples)
accept = rep(0, n.samples)

for(i in 1:n.samples){
  # Calculate qsi (deterministically) given our current value of tc
  qsi = funcCalcqs(a,b,s,g,D,S,obs$U,T,tc, RemoveZeros=FALSE)

  # Calculate updated parameters, ru and qu for [ s2 ]
  # [s2] \sim \text{I.G.}(ru, qu)
  ru = \frac{1}{\text{sum}((\log(obs$qs) - \log(qsi$qs))^2)/2 + 1/rp}
  qu = \text{n.obs}/2 + qp

  # Sample [ s2 ] from I.G.( ru, qu )
  s2.save[i] = rgamma(1,qu,ru)^(-1)
  s2 = s2.save[i]

  # Sample [ tc ] from N( tc(i-1) , s2.tune )
  # Normally, I'd call the variable tcstar (i.e., tc from the
  # proposal distribution),
  # but since tcstar is dimensionless shear, we'll go with tc.prop
  tc.prop = rtnorm(1,tc,sqrt(s2.tune),lower = tc.lower, upper = tc.upper)

  # Calculate the transport rates using tc.prop
qsi.prop = funcCalcqs(a,b,s,g,D,S,obs$U,T,tc.prop, RemoveZeros=FALSE)

# Calc our ratio to determine acceptance
# [ qsiob | tc, s2 ] ~ Norm( qsi, s2 )
numerator = sum(dnorm(log(obs$qs), log(qsi.prop$qs), sqrt(s2), log=TRUE)) + dtnorm(tc.prop, tc.mu0, sqrt(tc.var0), lower = tc.lower, upper = tc.upper,log=TRUE)
denominator = sum(dnorm(log(obs$qs), log(qsi$qs), sqrt(s2), log=TRUE)) + dtnorm(tc, tc.mu0, sqrt(tc.var0), lower = tc.lower, upper = tc.upper,log=TRUE)

ratio = exp(numerator-denominator)

if(numerator == 0 || denominator == 0){
cat("Error at iteration ", i,
numerator," Denominator:" ,denominator," Ratio:" ,ratio,
" Sep=""
})

if(ratio > runif(1)){
tc.save[i] = tc.prop
tc = tc.prop
accept[i] = TRUE
} else {
tc.save[i] = tc
accept[i] = FALSE
}

acceptancerate = round(length(accept[accept==1])/n.samples*100,2)

cat("\n","MCMC Sampling Complete... \n", length(accept[accept==0]), " of ", n.samples, " rejected \n", "Acceptance rate = ", acceptancerate,"% \n", "Optimal acceptance rate is approx. 30\% \n")
list(tc.save = tc.save, s2.save = s2.save, acceptancerate = acceptancerate, priorparams = priorparams)
}

funcPostPred = function(thetas,obs,post,mcmc,units){
###
### Posterior predictive interval (sample from likelihood using newly found parameters)
### I can use this to plot a line of fitted values vs. observed values (should plot along 1:1 line)
### lik ---> [ {log(qsi)} | tc, s2 ]
###

a = thetas$a
b = thetas$b
s = thetas$s
g = thetas$g
D = thetas$D
S = thetas$S
T = thetas$T
tc.save = post$tc.save
s2.save = post$s2.save
n.samples = mcmc$n.samples
n.burn = mcmc$n.burn

# we could parallelize this in a separate foreach loop, need to do:
# library(foreach)
# library(doMPI)
# install a version of MPI (e.g., Open MPI)
pred = matrix(0, n.samples, length(obs$U))
prednoise = matrix(0, n.samples, length(obs$U))

for(i in n.burn:n.samples){
pred[i,i] = t(funcCalcqs(a,b,s,g,D,S,obs$U,T,tc.save[i],
RemoveZeros=FALSE)$qs)
error = rnorm(length(obs$U),0,sqrt(s2.save[i]))
prednoise[i,i] = pred[i,i] * exp(error)
}

if(units==FALSE){
# then we want to return qstar and not qs
# take qs and convert back to qstar
pred = pred/sqrt((s-l)*g*D^3)
prednoise = prednoise/sqrt((s-l)*g*D^3)
}
list(pred = pred, prednoise = prednoise)
}

funcCalcDIC = function(thetas,obs,post,mcmc){
a = thetas$a
b = thetas$b
s = thetas$s
g = thetas$g
D = thetas$D
S = thetas$S
T = thetas$T

n.obs = mcmc$n.obs
n.samples = mcmc$n.samples
n.burn = mcmc$n.burn
\[ N = n.\text{samples} - n.\text{burn} \]

```r
###
### Calc \( D(y, \theta(i)) \)
###
D1 = matrix(0, N, 1)
j = 1

for(i in n.burn:n.samples){
# Get the ith tc from post
tc = post$tc.save[i]
# get the ith s2 from post
s2 = post$s2.save[i]

if(tc > obs$t[1]){
  cat('
 iteration: ', i,
  '\n tc: ', tc,
  '\n min observed shear: ', obs$t[1],
  '\n max tc in posterior: ', max(post$tc.save))
}
lik.mean = log(funcCalcqs(a, b, s, g, D, S, obs$U, T, tc, RemoveZeros=FALSE)$qs)
lik.var = s2
Dl[j] = -2*sum(dnorm(log(obs$qs), lik.mean, sqrt(lik.var), log=TRUE))
j = j + 1
}

Davg = sum(D1)/N

###
### Calc \( D_{\hat{\theta}(y)} \)
###
# Get \( E[\theta(i) \mid y] \)
# we're getting the average \( \theta(i) \) for both tc and s2 for all
# the samples after burn-in

tc.hat = mean(post$tc[n.burn:n.samples])
s2.hat = mean(post$s2[n.burn:n.samples])
# cat('\n Calc \( D_{\hat{\theta}(y)} \)\n',
# '\n true tc: ', simobs$tc,
# '\n tc.hat: ', tc.hat,
# '\n min observed tau: ', simobs$t[1])
lik.mean = log(funcCalcqs(a, b, s, g, D, S, obs$U, T, tc.hat, RemoveZeros=FALSE)$qs)
lik.var = s2.hat

Dth.hat = -2*sum(dnorm(log(obs$qs), lik.mean, sqrt(lik.var), log=TRUE))
DIC = 2*Davg-Dth.hat
cat('\n', DIC, '\n')
```
DIC
}

funcCalcMSPEhat = function(thetas, obs, priors, mcmc, units){
a = thetas$a
b = thetas$b
s = thetas$s
g = thetas$g
D = thetas$D
S = thetas$S
T = thetas$T
tc.muO = priors$tc.mu0
tc.varO = priors$tc.var0
tc.lower = priors$tc.lower
tc.upper = priors$tc.upper
s2.mu0 = priors$s2.mu0
s2.var0 = priors$s2.var0
n.samples = mcmc$n.samples
tc.seed = mcmc$tc.seed
s2.tune = mcmc$s2.tune
n.obs = mcmc$n.obs
n.burn = mcmc$n.burn
truevals = obs$qs
meanpreds = rep(0,length(obs$U))
pred = rep(0,n.samples)
prednoise = matrix(0,n.samples, length(obs$U))

# library(doSNOW)
# library(foreach)

for(i in 1:n.obs){
cat('
 Fitting model ' , i, ' of ' , length(obs$U), ' ...\n')
  # Remove a datapoint
tmpobs = list( U = obs$U[-i],
t = obs$t[-i],
tstar = obs$tstar[-i],
tc = obs$tc,
tcstar = obs$tcstar,
qs = obs$qs[-i],
qstar = obs$qstar[-i],
s2 = obs$s2 )
  # Fit the model using tmpobs
posti = funcMCMC(thetas,tmpobs,priors,mcmc,TRUE)
tc.save = posti$tc.save
s2.save = posti$s2.save

# Predict the ith observation (get mean value of postpred)
for(j in n.burn:n.samples){
# predict on missing value, i
pred[j] = funcCalcqs(a,b,s,g,D,S,obs$U[i],T,tc.save[j],
RemoveZeros=FALSE)$qs
error = rnorm(1,0,sqrt(s2.save[j]))
prednoise[j,i] = pred[j]*exp(error)
}

meanpreds[i] = mean(prednoise[,i])

cat("\n .. . Model ', i, ' complete \n")
}

if(units==FALSE){
# then we want to return qstar and not qs
# take qs and convert back to qstar
pred = pred/sqrt((s-l)*g*D^3)
prednoise = prednoise/sqrt((s-l)*g*D^3)
meanpreds = meanpreds/sqrt((s-l)*g*D^3)
truevals = obs$qstar
}

# Calculate MSPE.hat
MSPE.hat = sum((truevals - meanpreds)^2)/n.obs

list(preds = prednoise, truevals = truevals, meanpreds = meanpreds,
MSPE.hat=MSPE.hat)
}

###########################################################################
#
# Data functions
#
###########################################################################
funcTNMean = function(tc,s,rho,g,D,mu,s2,a,b){
    mu + funcTNPrior(mu,s2,a,b) - tc*(s-1)*rho*g*D
# mu + sqrt(s2)*(dnorm((a-mu)/sqrt(s2))-dnorm((b-mu)/sqrt(s2)))
# /(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/sqrt(s2))) - tc*(s-1)*rho*g*D
}

funcTNPrior = function(mu,s2,a,b){
    sqrt(s2)*(dnorm((a-mu)/sqrt(s2))-dnorm((b-mu)/sqrt(s2)))
    /(pnorm((b-mu)/sqrt(s2))-pnorm((a-mu)/sqrt(s2)))
}

funcCalcqs = function(a,b,s,g,D,S,U,T,tc,RemoveZeros){
    # Pre-allocate
t = rep(0,length(U))
tstar = rep(0,length(U))
qs = rep(0,length(U))
qstar = rep(0,length(U))

# Calculate the grain stress
rho = funcGetWaterDensity(T)
t = 0.052*rho*(g*S*D)-0.25*U-1.5

# Non-dimensionalize
tstar = t/(s-1)/rho/g/D
tcstar = tc/(s-1)/rho/g/D

# Calculate sediment discharge
# Basic exces shear (MPM)
qstar = a*(tstar-tcstar)^b

# Basic exces shear (MPM)
# qstar = 4.93*(tstar-tcstar)^1.6

# Convert qstar to qs
qs = sqrt((s-1)*g*D^3)*qstar

# Any row where tstar < tcstar qs = 0
# We won't see any transport when tstar is less than tcstar
tmp = tstar-tcstar
idx = which(tmp<0)
qstar[idx] = 0
qs[idx] = 0
nerr = 0

# Zeros should be removed when simulating data, but they should be kept for when likelihoods are calculated
if(RemoveZeros == TRUE && length(idx)!=0){
  return(list(U=U[-idx],t=t[-idx],tc=tc,tstar=tstar[-idx],
tcstar=tcstar,qstar=qstar[-idx],qs=qs[-idx]))
} else {
  if(length(idx)>0){
    cat('n Problem in funcCalcqs --- tcstar exceeded tstar n ',
    'tcstar:',tcstar,'n',
    'tstar:',tstar,'
')
  }
  return(list(U=U,tc=tc,tstar=tstar,tcstar=tcstar,qstar=qstar,qstar=qstar))
}

funcSimulateTransport= function(thetas){
  # unload parameters
  a = thetas$a
  b = thetas$b
  s = thetas$s
Given the grain size, $D$, let's find an approximate $\tau_{\text{cstar}}$ to use in the simulated data.

$\tau_{\text{cstar}} = \text{funcTaucStar}(D, g, s, T)$

Convert to dimensional space and give our random variable a variance $\rho = \text{funcGetWaterDensity}(T)$

$\tau_{\text{c}} = \tau_{\text{cstar}} (s-1) \rho g D$

Simulate the transport observations...

$\text{simvals} = \text{funcCalcqs}(a, b, s, g, D, S, U, T, \tau_{\text{c}}, \text{RemoveZeros}=\text{TRUE})$

Add Noise - Multiplicative Error Model

$\text{error} = \text{rnorm}(\text{length}(\text{simvals}\$U), 0, \text{sqrt}(s^2))$

$\text{simvals}\$qs = \text{simvals}\$qs \times \exp(\text{error})$

Let $q_{\text{star}}$ reflect the noise that we just added

$\text{simvals}\$q_{\text{star}} = \text{simvals}\$qs / \sqrt{(s-1) g D^3}$

list($U = \text{simvals}\$U,$t = \text{simvals}\$t,$t_{\text{star}} = \text{simvals}\$t_{\text{star}},$
tc = \text{simvals}\$tc,$tc_{\text{star}} = \text{simvals}\$tc_{\text{star}},$
qs = \text{simvals}\$qs,$q_{\text{star}} = \text{simvals}\$q_{\text{star}},$
s2 = s2$
)

---

Ancillary functions

funcGetWaterDensity = function(T){
  # Returns the density of water in kg/m$^3$ based on Temperature in C
  Temps = c(0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50)
  rhos = c(999.9, 1000.0, 999.7, 999.1, 998.2, 997.1, 995.7, 994.1, 992.2, 990.2, 988.1)
  rho = approx(Temps, rhos, xout=T, method='linear')$y
\[
\text{funcGetKinematicViscosity} = \text{function}(T)\{
\text{# Returns the kinematic viscosity of water in m}^2/\text{s based on Temp in C}
\}
\]

Temps = c(0,5,10,15,20,25,30,35,40,45,50)
Viscosities = c(1.792,1.519,1.308,1.141,1.007,0.897,0.804,0.727,0.661,0.661,0.605)
u = approx(Temps,Viscosities,xout=T,method='linear')$y*1e-6
nu
\]

\[
\text{funcRep} = \text{function}(D,g,R,T)\{
\text{# Returns the Particle Reynold's Number given:}
\text{# D = Diameter in m}
\text{# g = gravity in m/s}^2
\text{# R = (rho_s - rho)/rho}
\text{# T = Temperature in Degree C}
\text{# See Eqn 2-73b of ASCE STM No. 110}
\nu = \text{funcGetKinematicViscosity}(T)
\text{Rep} = \sqrt{g*R*D}/nu
\text{Rep}
\}
\]

\[
\text{funcTaucStar} = \text{function}(D,g,s,T)\{
\text{# Returns the Shields Number given:}
\text{# D = Diameter in m}
\text{# g = gravity in m/s}^2
\text{# s = rho_s/rho (around 2.65)}
\text{# T = Temperature in Degree C}
\text{# See Eqn 2-59a of ASCE STM No. 110}
\text{R} = s-1
\text{Rep} = \text{funcRep}(D,g,R,T)
\text{taucstar} = 0.22*Rep^{-0.6}+0.06*exp(-17.77*Rep^{-0.6})
\text{taucstar}
\}
\]

\[
\text{funcCalcIGDensity} = \text{function}(x,r,q)\{
\text{x}^{-(q+1)}/(r^q*lgamma(q))*exp(-1/(r*x))
\}
\]
funcCreateHistogram = function(x,n.burn){
  hist(x[-c(1:n.burn)],xlab="",ylab="",main="", freq = FALSE)
  abline(v = mean(x[-c(1:n.burn)]),col='red',lwd=2)
}

funcCreateTracePlot = function(x,n.burn,n.samples,log){
  plot(x,type='n',xlab="",ylab="",main="",log=log)
  points(x[c(1:n.burn)],col='purple',type='l',lwd='2')
  points(c(n.burn+1:(n.samples-n.burn)),x[-c(1:n.burn)],type='l',lwd='2')
}

funcCreateSummaryPlots = function(obs,priors,post,mcmc,DIC,
  thetas,sim=TRUE){
  n.samples = mcmc$n.samples
  n.burn = mcmc$n.burn
  n.obs = mcmc$n.obs

  tc.mu0 = priors$tc.mu0
  tc.var0 = priors$tc.var0
  tc.mu0.TN = priors$tc.mu0.TN
  s2.mu0 = priors$s2.mu0
  s2.var0 = priors$s2.var0

  tc.lower = post$priorparams$tc.lower
  tc.upper = post$priorparams$tc.upper

  qp = post$priorparams$qp
  rp = post$priorparams$rp

  library(RColorBrewer)
  cols = brewer.pal(6,"Set1")

  layout(matrix(c(1,2,3,4,5,6,7,8,9),3,3,byrow=TRUE),
  heights = c(1,2,2),byrow = TRUE), def=par(mar = c(1,1,1,1))

  plot(1,1,type='n',axes=FALSE,xlab='',ylab='')
  title(main='Trace Plot Legend')
  if(sim==TRUE){
    legend('center', c('Burn-in', 'MCMC Samples','True Value'),
    col=c('purple','black',cols[3]),lwd=2,bg='white')
  } else {
    legend('center', c('Burn-in', 'MCMC Samples'),
    col=c('purple','black'),lwd=2,bg='white')
  }

  plot(1,1,type='n',axes=FALSE,xlab='',ylab='')
  title(main='Histogram Legend')
  if(sim==TRUE){

legend('center', c('True Value', 'Posterior Mean', 'Posterior 95% Interval', 'Prior Density', 'Prior Mean', 'Prior 95% Interval'), col=c(cols[3], cols[1], cols[1], cols[2], cols[2], cols[2]), lwd=2, lty=c(1, 1, 2, 4, 1, 2), bg='white', cex=0.95)
}
else {
  legend('center', c('Posterior Mean', 'Posterior 95% Interval', 'Prior Density', 'Prior Mean', 'Prior 95% Interval'), col=c(cols[1], cols[1], cols[2], cols[2], cols[2]), lwd=2, lty=c(1, 2, 4, 1, 2), bg='white', cex=0.95)
}

# MCMC summary text pane
plot(c(0,1), c(0,1), type='n', axes=FALSE, xlab=' ', ylab='')
text(0.5, 1, adj = c(0.5, 1), labels=paste(
  "MCMC Samples: ", n.samples,
  "Burn-in: ", n.burn,
  "Total Observations (n.obs): ", n.obs,
  "Acceptance Rate: ", post$acceptancerate, ", ",
  "Deviance Information Criterion: ", round(DIC,2),
  sep=""), cex=1.0)

def = par(mar = c(5, 4, 4, 2) + 0.1)

# Trace plot for tc
funcCreateTracePlot(post$tc.save, n.burn, n.samples, log='')
title(main = 'Critical Shear Trace', ylab='tc', xlab='Iteration')
if(sim==TRUE){
  abline(h = obs$tc, col=cols[3], lwd = 2)
}

# Histogram for tc
priorquantiles = qtnorm(c(0.025, 0.975), tc.mu0, sqrt(tc.var0),
  lower = tc.lower, upper = tc.upper)
postquantiles = quantile(post$tc.save[-c(1:n.burn)], c(0.025, 0.975))
priormean = post$priorparams$tc.mu0tn
xmin = min(postquantiles[1]*0.95, obs$tc)
xmax = max(postquantiles[2]*1.05, obs$tc)

priordensity = dtnorm(seq(xmin, xmax, .200), tc.mu0, sqrt(tc.var0),
  lower = tc.lower, upper = tc.upper)
postmax = max(density(post$tc.save[-c(1:n.burn)])$y)
priormax = max(priordensity)

# Normalize the densities so one doesn’t dwarf the other
fac = funcNormalizeDensities(postmax, priormax)
hist(post$tc.save[-c(1:n.burn)], xlab='', ylab='', main='', freq = FALSE, xlim=c(xmin, xmax))
abline(v = mean(post$tc.save[-(1:n.burn)]), col = cols[1], lwd = 2, lty = 1)
abline(v = postquantiles, col = cols[1], lwd = 1, lty = 2)
points(seq(xmin, xmax, 200), fac*priordensity, type = 'l', lty = 4, lwd = 2, col = cols[2])
abline(v = tc.muO.TN, col = cols[2], lwd = 2, lty = 1)
abline(v = priorquantiles, col = cols[2], lwd = 1, lty = 2)
if(sim == TRUE){
  abline(v = obs$tc, col = cols[3], lwd = 2, lty = 1)}
title(main = 'Critical Shear (after burn-in)', ylab = 'Density', xlab = 'tc')

# Summary text pane for critical shear
plot(c(0,1), c(0,1), type = 'n', axes = FALSE, xlab = '', ylab = '')
title(main = 'Critical Shear Summary')

if(sim == TRUE){
  tempnum = 1/(thetas$s - 1)/thetas$rho/thetas$g/thetas$D
  text(0.5, 0.95, adj = c(0.5, 1), labels = paste("\nTrue Value: ", round(obs$tc, 2), " (", round(obs$tcstar, 3), " )", "\nPosterior Mean: ", round(mean(post$tc.save[-(1:n.burn)]), 2), " (", round(mean(post$tc.save[-(1:n.burn)])*tempnum, 3), " )", "\n95% Cred. Int. (", round(postquantiles[1], 2), " - ", round(postquantiles[2], 2), ",")", "\nPrior Mean: ", round(tc.muO.TN, 2), " (", round(tc.muO.TN*tempnum, 3), ",)", "\n95% Cred. Int. (", round(priorquantiles[1], 2), " - ", round(priorquantiles[2], 2), ",")", "\n\n95% Cred. Int. (", round(priorquantiles[1]*tempnum, 3), " - ", round(priorquantiles[2]*tempnum, 3), ",)", sep = "", cex = 1.0)
} else {
  tempnum = 1/(obs$s - 1)/obs$rho/obs$g/obs$D
  text(0.5, 0.95, adj = c(0.5, 1), labels = paste("\n\nPosterior Mean: ", round(mean(post$tc.save[-(1:n.burn)]), 2), " (", round(mean(post$tc.save[-(1:n.burn)])*tempnum, 3), ",")", "\n95% Cred. Int. (", round(postquantiles[1], 2), " - ", round(postquantiles[2], 2), ",")", "\nPrior Mean: ", round(tc.muO.TN, 2), " (", round(tc.muO.TN*tempnum, 3), ",)", "\n95% Cred. Int. (", round(priorquantiles[1], 2), " - ", round(priorquantiles[2], 2), ",")", "\n\n95% Cred. Int. (", round(priorquantiles[1]*tempnum, 3), " - ", round(priorquantiles[2]*tempnum, 3), ",)", sep = "", cex = 1.0)}
# Trace plot for s2
funcCreateTracePlot(post$s2.save,n.burn,n.samples,log='y')
title(main = 'Variance Trace',ylab='s2', xlab = 'Iteration')
if(sim==TRUE){
  abline(h = obs$s2,col=cols[3],lwd = 2)
}

# Histogram for s2
priorquantiles = qgamma(c(0.025,0.975),qp,,rp)^(-1)
#priorquantiles = quantile(rgamma(1e6,qp,,rp)^-1,c(0.025,0.975))
postquantiles = quantile(post$s2.save[-(1:n.burn)],c(0.025,0.975))

xmin = min(postquantiles[1]*0.75,obs$s2)
xmax = max(postquantiles[2]*1.25,obs$s2)

priordensity = funcCalcIGDensity(seq(xmin,xmax,,200),rp,qp)
postmax = max(density(post$tc.save[-c(1:n.burn)])$y)
priormax = max(priordensity)
# Normalize the densities so one doesn't dwarf the other
fac = funcNormalizeDensities(postmax,priormax)

# xmin = min(min(priorquantiles),min(postquantiles),obs$s2)
# xmax = max(max(priorquantiles),max(postquantiles),obs$s2)
hist(post$s2.save[-c(1:n.burn)],xlab='','',ylab='','',main=''
  freq = FALSE,xlim=c(xmin,xmax))
abline(v = mean(post$s2.save[-(1:n.burn)]),col=cols[1],lwd=2,lty=1)
abline(v = postquantiles, col=cols[1],lwd=1,lty=2)
points(seq(xmin,xmax,,200),fac*priordensity,lty=4,type='l'
  lwd=2,
  col = cols[2])
abline(v = s2.mu0, col=cols[2],lwd=2,lty=1)
abline(v = priorquantiles, col=cols[2],lwd=1,lty=2)
if(sim==TRUE){
  abline(v = obs$s2,col=cols[3],lwd=2,lty=1)
}
title(main = 's2 (after burn-in)',ylab='Density', xlab = 's2')

plot(c(0,1),c(0,1),type='n',axes=FALSE,xlab='','',ylab=''
  title(main='s2 Summary')
if(sim==TRUE){
text(0.5,0.75, adj = c(0.5, 1),labels=paste(  
"n True Value: ", round(obs$s2,2),  
"n \n Posterior Mean: ", round(mean(post$s2.save[-(1:n.burn)]),  
2),  
"n 95% Cred. Int. (",round(min(postquantiles),2)," - ",  
round(max(postquantiles),2),")",  
"n \n Prior Mean: ", round(s2.mu0,2),  
}
"\n \n 95% Cred. Int. (" , round(min(priorquantiles),2), ", ",
round(max(priorquantiles),2), "]", 
sep="" ), cex=1.0)
} else {
  text(0.5,0.75, adj = c(0.5, 1), labels=paste(
  "\n \n  Posterior Mean: ", round(mean(post$s2.save[-(1:n.burn)]),
2),
  "\n 95% Cred. Int. (" , round(min(postquantiles),2), ", ",
round(max(postquantiles),2), "]",
  "\n  Prior Mean: ", round(s2.mu0,2),
  "\n 95% Cred. Int. (" , round(min(priorquantiles),2), ", ",
round(max(priorquantiles),2), "]", 
sep="" ), cex=1.0)
}
}

funcNormalizeDensities = function(postmax, priormax){
  fac = 1
  if((postmax/priormax)>3){
    fac = (postmax/priormax)*0.5
  } else if((postmax/priormax)<1.1){
    fac = (postmax/priormax)*0.75
  } else {
    fac = 1
  }
  fac
}

funcCreateShadePlot = function(postpredn, obs, ylims=c(-10,2)){
  cols = brewer.pal(6, 'Greys')
  
  # add column to postpredn matrix to house the tstars
  postpredn = rbind(obs$tstar,postpredn)
  
  # sort postpredn according to the last column
  postpredn = postpredn[,order(postpredn[1,])]
  
  # reassign ordered shears
  x = log(postpredn[1,])
  
  # remove first row
  postpredn = postpredn[-1,]
  
  # Create polygon plots
  q = c(0.025,0.05,0.16,0.84,0.95,0.975)
  
  yq = apply(postpredn,2,quantile,q)
  yq = log(yq)
  
  maxvals = apply(postpredn,2,max)
  minvals = apply(postpredn,2,min)
matplot(c(x,rev(x)),log(c(maxvals,rev(minvals))),type='n',lty=1, xlab='log(t*)',ylab='log(q*)',ylim=ylims)
polygon(c(x,rev(x)),c(yq[1,],rev(yq[6,])),col=cols[3],border=NA)
polygon(c(x,rev(x)),c(yq[2,],rev(yq[5,])),col=cols[4],border=NA)
polygon(c(x,rev(x)),c(yq[3,],rev(yq[4,])),col=cols[5],border=NA)
points(log(obs$tstar),log(obs$qstar),col='red',pch=19,cex=0.75)
}

funcCreatePolygonPlot = function(postpred,obs,n.burn,n.samples, ylims=c(-10,2)){
  library(RColorBrewer)
cols = brewer.pal(6,'Paired')
  # add column to postpredn matrix to house the tstars
  y1 = rbind(obs$tstar,postpred$pred[n.burn:n.samples,])
y2 = rbind(obs$tstar,postpred$prednoise[n.burn:n.samples,])
  # sort postpredn according to the last column
  y1 = y1[,order(y1[1,])] 
y2 = y2[,order(y2[1,])]
  # reassign ordered shears
  x = log(y1[1,])
  # remove first row
  y1 = y1[-1,]
y2 = y2[-1,]
  qstc = apply(y1,2,quantile,c(0.025,0.975))
  qsall = apply(y2,2,quantile,c(0.025,0.975))

  matplot(x, log(t(postpred$prednoise)),type = 'n',lty=1, xlab='log(t*)',ylab='log(q*)',ylim=ylims)
polygon(c(x,rev(x)),log(c(qsall[2,],rev(qsall[1,]))),col=cols[1], border=NA)
polygon(c(x,rev(x)),log(c(qstc[2,],rev(qstc[1,]))),col=cols[2], border=NA)
points(log(obs$tstar),log(obs$qstar),col='red',pch=19,cex=0.75)
}

funcCreatePPIPlots = function(postpred,obs,mcmc,ylims=c(-10,2)){
  n.burn = mcmc$n.burn
  n.samples = mcmc$n.samples
  layout(matrix(c(1,2,3,4),2,2,byrow = TRUE),heights = c(1,5) )
def = par(mar = c(0,0,0,0))
plot(1,1,type='n',axes=FALSE,xlab='',ylab='')
cols = brewer.pal(6,'Greys')
legend('center', c('95% Posterior Predictive Interval ','}
'90% Posterior Predictive Interval ', '68% Posterior Predictive Interval '),
col=c(cols[3:5]),lwd=4,bg='white',cex=1)

cols = brewer.pal(6,'Paired')
plot(1,1,type='n',axes=FALSE,xlab='',ylab='')
legend('center', c('Total Variability(s2 and tc)',
'Variability from tc'),
col=c(cols[1:2]),lwd=4,bg='white',cex=1)
def = par(mar = c(5, 4, 4, 2)+0.1)
funcCreateShadePlot(postpred$prednoise[n.burn:n.samples],obs,ylims)
funcCreatePolygonPlot(postpred,obs,n.burn,n.samples,ylims)
}

funcCreateMSPEBoxPlots = function(postpred, obs, mspe, mcmc, animate,
t=0, leg=TRUE){
library(RColorBrewer)
cols = brewer.pal(6,"Set1")
n.burn = mcmc$n.burn
n.samples = mcmc$n.samples
n.obs = mcmc$n.obs

x = log(obs$tstar)
# PPD is plotted as type='n' just to setup the plot-space.
matplot(x, log(t(postpred$prednoise)),type='n',lty=1,
xlab='log(t*)',ylab='log(q*)')

if(leg==TRUE){
title(main= ' MSPE Cross-validation ')
legend('topleft', c('Predicted values (Boxplot)',
'Mean Predicted Value', 'Observed Value',paste('MSPE.hat: ',
round(mspe$MSPE.hat,4),sep='')),
col=c('black',cols[3],cols[1]),lwd=0,lty=c(1,-1,-1,-1),
pch=c(-1,17,20,-1),bg='white',cex=0.95)
} else {
title(main=paste('MSPE.hat = ', round(mspe$MSPE.hat,4),sep=''))
}

for(i in 1:n.obs){
boxplot(log(mspe$preds[n.burn:n.samples,i]),at = x[i],add=TRUE,
boxwex = 0.05, outline=FALSE)
points(x[i],log(mspe$mean preds[i]), col = cols[3],pch=17)
points(x[i],log(mspe$true vals[i]),col=cols[1],pch=19)
if(animate==TRUE){
Sys.sleep(t)
}
funcPlotPrior = function(priors, obs, thetas, breaks=50,
titletext='prior distribution for tc', leg=TRUE, legtext=''
, tyaxt='s', txaxt='s'){
library(RColorBrewer)
cols = brewer.pal(8,'Set1')

x = rtnorm(1e5, priors$tc.mu0, sqrt(priors$tc.var0),
lower=priors$tc.lower, upper=priors$tc.upper)

hist(x, breaks= breaks, main = titletext ,xlim = c(0, min(obs$t)),
freq=FALSE, yaxt=tyaxt, xaxt=txaxt)
if (txaxt=='n') {
axis(1, labels=FALSE)
}
if (leg==TRUE) {
c (O, min(obs$t), col=cols[2], lwd=4, lty=1)
abline(v = mean(x), col=cols[8], lwd=4, lty=3)
abline(v = min(obs$t), col=cols[1], lwd =4, lty=1)
legend('topleft', c(paste('tc.mu0:', round(priors$tc.mu0,2), sep=''),
'tc.mu0.TN ', round(priors$tc.mu0.TN,2),
'Min(observed tau): ', round(min(obs$t) ,2)),
col=c(cols[2], cols[8], cols[1]), lty=c(1,3,1), lwd=4, bg='white')
} else {
mtext(legtext, 3, adj=O, line=O)
}

}

funcPlotPrior_s2 = function(priors, obs, thetas, breaks=20,
titletext='prior distribution for s2', leg=TRUE, legtext=''
, tyaxt='s', txaxt='s'){
library(RColorBrewer)
cols = brewer.pal(8,'Set1')

# [ s2 ] ~ I.G. (rp,qp)
rp = priors$s2.var0/(priors$s2.mu0*(priors$s2.mu0~2+priors$s2.var0))
qp = 1/(priors$s2.mu0*rp)+1
x = rgamma(1e5, qp,,rp)^(-1)
xlims = c(0, quantile(x,0.975))

hist(x, breaks = breaks, main = titletext, xlim=xlims, freq=FALSE,
yaxt=tyaxt, xaxt=txaxt, cex=.8)
if(txaxt=='n'){
axis(1, labels=FALSE)
}

if(leg==TRUE){
}
else {
  mtext(legtext, 3, adj=0, line=0)
}

}
APPENDIX B
MODEL FIT SUMMARY PLOTS

Fig. 34: Summary plot for data scenario 1, τ_c Prior 1.
Fig. 35: Summary plot for data scenario 1, $\tau_c$ Prior 2.
Fig. 36: Summary plot for data scenario 1, $\tau_c$ Prior 3.
Trace Plot Legend

- Burn-in
- MCMC Samples
- True Value

Histogram Legend

- True Value
- Posterior Mean
- Posterior 95% Interval
- Prior Density
- Prior Mean
- Prior 95% Interval

MCMC Samples: 5000
Burn-in: 1500
Total Observations (n.obs): 21
Acceptance Rate: 30.92%
Deviance Information Criterion: 4.82

Critical Shear Trace

Critical Shear (after burn-in)

Critical Shear Summary

True Value: 6.72 (0.052)
Posterior Mean: 6.75 (0.052)
95% Cred. Int. (6.71 - 6.78)
(0.052 - 0.052)
Prior Mean: 5.25 (0.041)
95% Cred. Int. (1.36 - 6.88)
(0.011 - 0.053)

Variance Trace

s2 (after burn-in)

s2 Summary

True Value: 0.05
Posterior Mean: 0.06
95% Cred. Int. (0.04 - 0.11)
Prior Mean: 0.05
95% Cred. Int. (0.01 - 0.21)

Fig. 37: Summary plot for data scenario 1, $\tau_c$ Prior 4.
Fig. 38: Summary plot for data scenario 2, $\tau_c$ Prior 1.
Fig. 39: Summary plot for data scenario 2, $\tau_c$ Prior 2.
Fig. 40: Summary plot for data scenario 2, $\tau_c$ Prior 3.
Fig. 41: Summary plot for data scenario 2, $\tau_c$ Prior 4.
Fig. 42: Summary plot for data scenario 3, $\tau_c$ Prior 1.
Fig. 43: Summary plot for data scenario 3, \( \tau_c \) Prior 2.
Fig. 44: Summary plot for data scenario 3, $\tau_c$ Prior 3.
**Trace Plot Legend**

- **Burn-in**
- **MCMC Samples**
- **True Value**

**Histogram Legend**

- **True Value**
- **Posterior Mean**
- **Posterior 95% Interval**
- **Prior Density**
- **Prior Mean**
- **Prior 95% Interval**

**Critical Shear Trace**

**Critical Shear (after burn-in)**

**Critical Shear Summary**

- True Value: 6.72 (0.052)
- Posterior Mean: 6.62 (0.051)
- 95% Cred. Int.: (6.34 – 6.8) (0.049 – 0.052)
- Prior Mean: 5.25 (0.041)
- 95% Cred. Int.: (1.36 – 6.88) (0.011 – 0.053)

**Variance Trace**

**s2 (after burn-in)**

**s2 Summary**

- True Value: 1.1
- Posterior Mean: 1.04
- 95% Cred. Int.: (0.67 – 1.58)
- Prior Mean: 1.1
- 95% Cred. Int.: (0.58 – 2.06)

**Fig. 45:** Summary plot for data scenario 3, \( \tau_c \) Prior 4.
Fig. 46: Summary plot for data scenario 4, $\tau_c$ Prior 1.
Fig. 47: Summary plot for data scenario 4, $\tau_c$ Prior 2.
Fig. 48: Summary plot for data scenario 4, $\tau_c$ Prior 3.
Trace Plot Legend

Histogram Legend

Critical Shear Trace

Critical Shear (after burn-in)

Critical Shear Summary

Variance Trace

s2 (after burn-in)

s2 Summary

MCMC Samples: 5000
Burn-in: 1500
Total Observations (n.obs): 6
Acceptance Rate: 30.08%
Deviance Information Criterion: 14.33

Fig. 49: Summary plot for data scenario 4, $\tau_c$ Prior 4.
Fig. 50: Summary plot for data scenario 1, $\sigma^2$ Prior 1.
Fig. 51: Summary plot for data scenario 1, $\sigma^2$ Prior 2.
Fig. 52: Summary plot for data scenario 2, $\sigma^2$ Prior 1.
Fig. 53: Summary plot for data scenario 2, $\sigma^2$ Prior 2.
Fig. 54: Summary plot for data scenario 3, $\sigma^2$ Prior 1.
Fig. 55: Summary plot for data scenario 3, $\sigma^2$ Prior 2.
Figure 56: Summary plot for data scenario 4, \( \sigma^2 \) Prior 1.
Fig. 57: Summary plot for data scenario 4, $\sigma^2$ Prior 2.