Wind farm investment under spatial uncertainty: A Bayesian model with an option of waiting

Jostein Lillestøl & Johannes Mauritzen
Department of Business and Management Science
NHH Norwegian School of Economics
Bergen, Norway

June 15, 2015

Abstract

...

*We would like to thank Endre Bjørndal, Jonas Andersson, and Gunnar Eskeland for helpful comments and suggestions. Data has been generously provided by Kjeller Vindteknikk. Financial support has been received from the Research Council of Norway through the projects INTREPID and CenSES.

1 Introduction

Unlike traditional energy generation technologies, the efficiency and profitability of wind turbines and wind farms are highly dependent on their location. Even modest differences in long-term average wind speeds can have major consequences on the profitability of wind farms over time. However, the inherent appropriateness of a given location is subject to a high degree of uncertainty as wind speeds can vary substantially even within areas where overall wind resources are good. More so, average wind speeds show a high degree of variance from month-to-month and even year-to-year.

The high degree of spatial and temporal variation of wind speeds leads to large uncertainty when making wind farm investment decisions. In this article we propose a simple sequential Bayesian decision model that takes into account spatial uncertainty and allows for the option of waiting to collect more information on wind speeds. This method mimics actual behavior of wind power investors, who tend to make decisions on a sequential basis.

We first introduce a simple two period model with linear loss function that can be solved analytically, generating some stylized results that provide intuition and a frameworks for interpreting more complicated models. We then introduce a case study with data from a proposed wind farm in Northern Norway. In this model we introduce realistic features, such as a non-linear power-output function for a wind turbine.

We use the Bayesian posterior sampling and modeling language Stan [Stan Development Team, 2014b] to simulate sampling from the posterior. This approach allows the model to be easily extended to take into account, for example, price and regulatory uncertainty - though for the sake of brevity and clarity, we do not pursue these extensions here. However, we provide well documented code and data online `jmaurit.github.io#wind_invest_model`.

2 Literature

3 Analytic Frameworks

4 Case Study: The Andmyran Wind Farm in Northern Norway

We consider the case of the proposed Andmyran wind farm, located on an island in northern Norway. To establish the basis for the prior beliefs about the wind resources of the given location, we gather publicly available information on average wind speeds from three nearby weather stations: Andenes, Harstad and Sortland. This is obtained from the website of the Norwegian Meteorological Society's public weather page, `yr.no`.

Figure 1 shows a chart of the average monthly wind speeds for the three sites. We can note the significant variation in average wind speed between the sites, even though they are all within 60 kilometers of one another.

Wind speeds measured at shorter intervals are known to roughly follow a weibull distribution [Weibull, 1951]. [1].

$$\mu_{weibull} = \sigma * Gamma(1 + \frac{1}{\alpha}) \tag{1}$$

In turn the Weibull parameter values of α and σ are modeled as Gamma distributions as this provides a positive support and flexible shape.

[1] A weibull distribution is defined as having a probability distribution function of $f(x, \sigma, \alpha) = \frac{\alpha}{\sigma}\left(\frac{x}{\sigma}\right)^{\alpha-1} e^{-\left(\frac{x}{\sigma}\right)^{k}}$

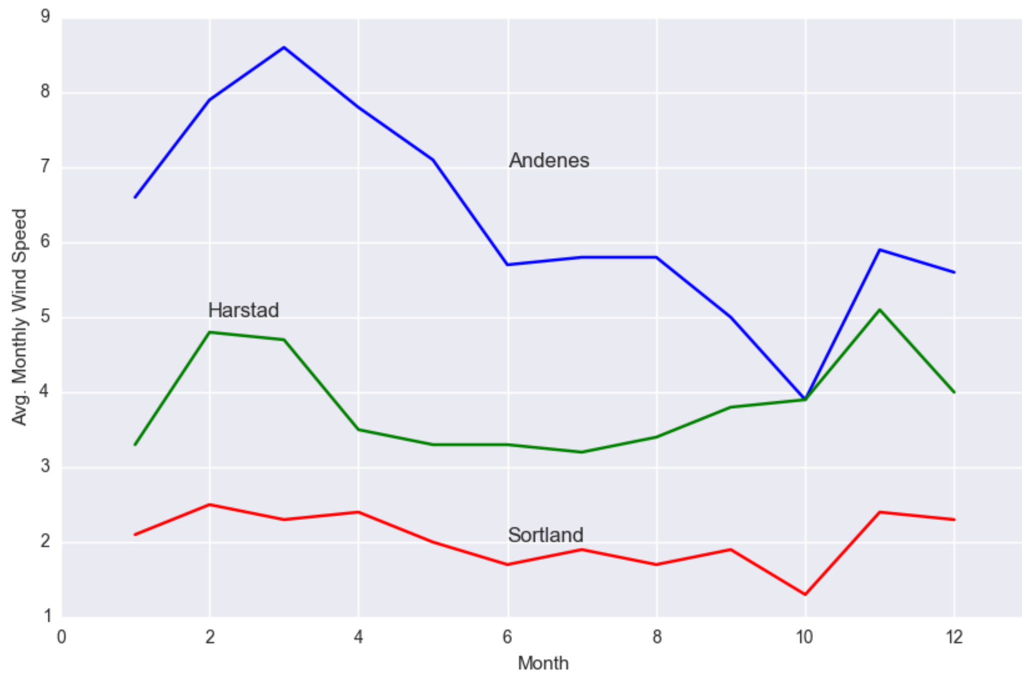

Figure 1: Average monthly wind speeds from the three closest weather stations to the proposed Andymren site.

To obtain semi-informative priors on the α and σ variables, vectors $\hat{\alpha_i}$ and $\hat{\sigma_i}$ are estimated from the monthly average wind speed data shown above for locations i. We use the formula for the mean of a weibull distribution, as shown in equation 1, where $Gamma$ represents the Gamma function. We then minimize squared errors as in equation 2 to obtain the vectors of estimated parameters.

$$min_{\alpha,\sigma} \sum_{t=1}^{n} (\mu_t - \sigma * Gamma(1 + 1/\alpha))^2 \tag{2}$$

Figure 2 show the prior gamma distributions on the α and σ weibull parameters that best fit the estimated vectors $\hat{\alpha_i}$ and $\hat{\sigma_i}$. The gamma distribution for the α parameter has a shape parameter of $\lambda_\alpha = 3.9$ and a scale parameter $\zeta_\alpha = .252$. The prior gamma distribution on the σ parameter is estimated to have a shape parameter of $\lambda_\sigma = 34.5$ and a scale parameter of $\zeta_\sigma = .085$.

The top panel of figure 3 shows the distribution of 1-year of hourly wind speeds drawn from a weibull distribution with fixed α and σ parameters.

To transform the simulated wind speed data from the weibull distribution to power output,

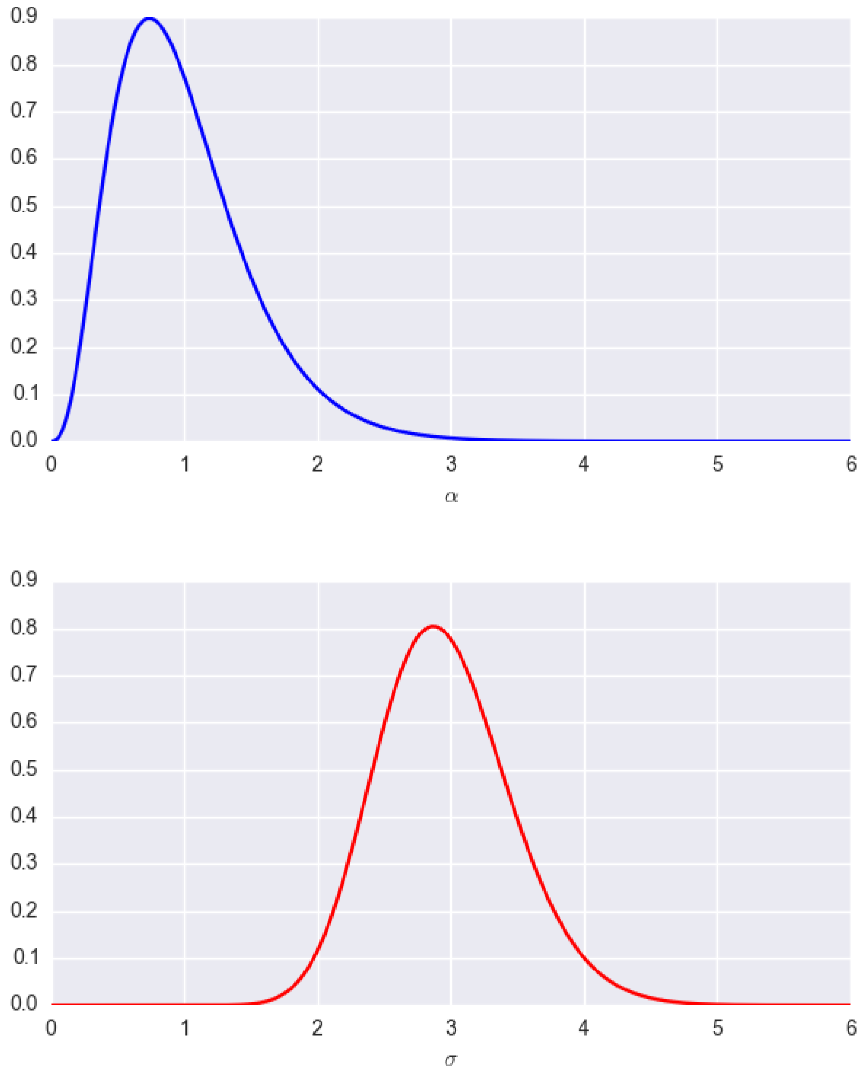

Figure 2: Gamma prior probability distributions on the weibull shape and scale parameters α and σ.

we use the power curve of a common commercial wind turbine: the Vestas V90 turbine [Vestas, 2015]. The power curve can be seen in the middle panel of figure 3. At wind speeds up to 4 meters per second (m/s), no power is produced. Between speeds of 4 and 15 m/s, power output rapidly increases with respect to wind speed. The rated power output is reached at wind speeds of between 15 and 25 m/s, after which power is cut-out in order to protect the turbine.

Obtaining a prior predictive distribution of yearly power output is accomplished with the following algorithm:

1. Draw weibull parameters $\hat{\alpha}$ and $\hat{\sigma}$ from the appropriate gamma distributions.
2. Draw 1-year of (8640 hourly) wind speeds from a $weibull(\hat{\alpha}, \hat{\sigma})$ distribution.
3. Convert the hourly wind speed draws to power output via the power curve
4. Sum to get a yearly total
5. Repeat

The lower panel of figure 3 shows the distribution of yearly power outputs.

After having generated a prior predictive distribution of yearly power outputs. We define the following loss functions as in equation 6:

$$L_{invest} = I - (p_{kwh} - c_{kwh}) * kwh \tag{3}$$

$$L_{pass} = d((p_{kwh} - c_{kwh}) * kwh - I) \tag{4}$$

$$L_{wait-invest} = M + I - (p_{kwh} - c_{kwh}) * kwh \tag{5}$$

$$L_{wait-pass} = M + I - d(p_{kwh} - c_{kwh}) * kwh \tag{6}$$

These functions describe the losses associated with investing or passing in the first period as well as well as the decision to wait. I represents the lump-sum investment cost, p_{kwh} is the price per kilowatt-hour (kWh) for the electricity produced. c_{kwh} represents any running costs. d represents a discount rate on the opportunity cost of not investing - the intuition here is that the capital could be put in place in some other investment that gains a return. M represents the cost associated

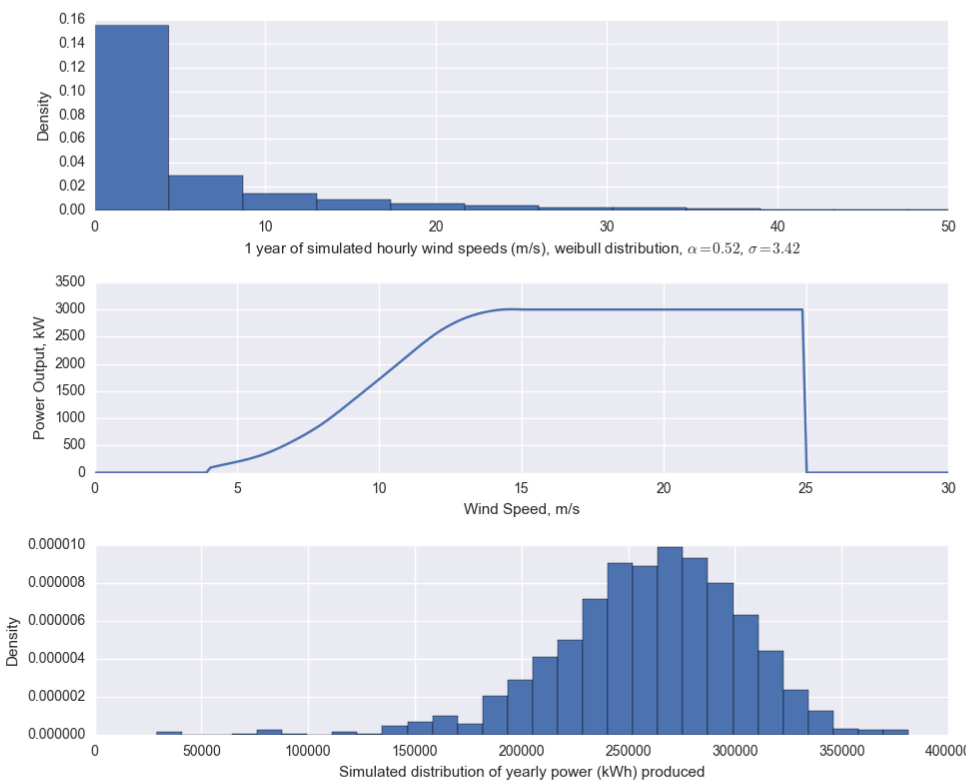

Figure 3: The top panel shows a histogram of 1-year of simulated wind speeds from a Weibull probability distribution. The middle panel shows the power curve of the V90 3 mW Vesta wind turbine. The bottom panel shows the prior predictive distribution of yearly power output from the wind turbine.

$E(loss_{invest})$	1394
$E(loss_{pass})$	-1254
$E(loss_{wait})$	-8129

Table 1: Expected losses from decision to invest, pass or wait at first stage

with waiting and gathering data. This could include the cost of installing and monitoring a wind-measuring station.

Given the loss functions for the different choices with parameters: $I = 80,000$, $c_{oper} = 0.005$, p_{kwh}, $d = .90$, $M = 0$ we then draw 500 sample years of hourly wind data. At the first stage the investor can either choose to invest, to pass, or to wait and gather more information. The distribution of the losses for each sampled year for these three choices are shown in figure 4.

With the chosen parameters there is a high degree of prior uncertainty about whether a project will be profitable or not. Thus there is a high degree of value in waiting. The strong assumption in calculating the expected loss of waiting is that the investor will know with certainty whether investing is profitable or not after having gathered more information. This is why the expected loss of waiting in this simulation is always less than zero. This assumption can easily be relaxed.

Taking the expected value from these samples, we get the values shown in table 1. Clearly, the value of waiting in this example are quite high. One way of interpreting the result would be to say that it would be worthwhile to wait and gather more information if the cost of gathering the information, M were less than 6875 (8129-1254) Otherwise, you would pass on making the investment.

Having decided to wait and gather more information, the next step is then to update the prior distribution with the likelihood obtained from the gathered data and obtain a posterior distribution with which to make a new investment decision.

The simple model can be expressed as in equation 9. The first term expresses the likelihood as a weibull function with shape parameter α and scale parameter σ. In turn these parameters are given informative prior distributions with means $\hat{\alpha}$ and $\hat{\sigma}$ that are estimated from the publicly available data on average wind speeds in the area, as detailed earlier.

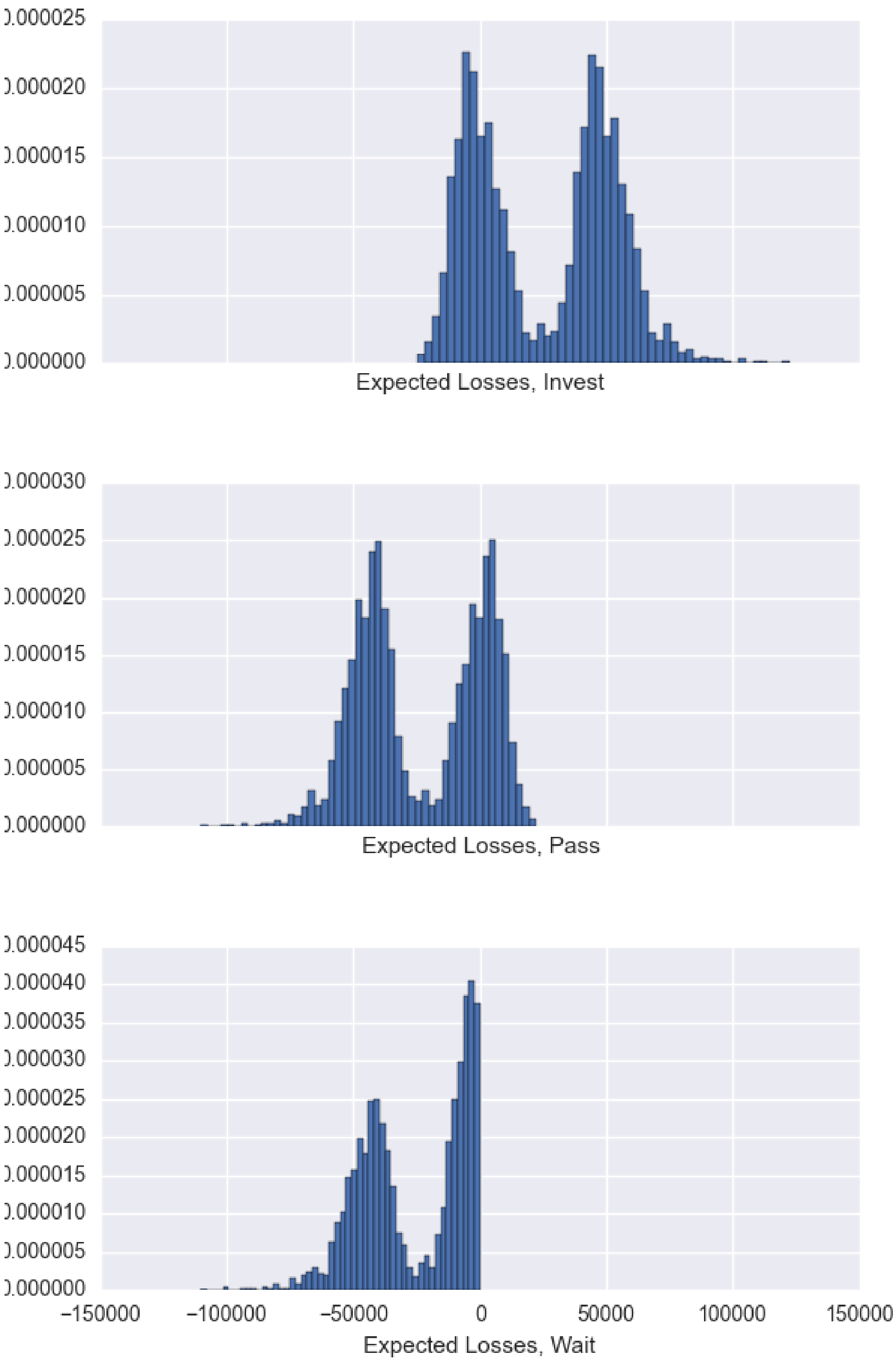

Figure 4: Distribution of losses for the decisions to invest, pass or wait.

$$p(wind - data|\theta) \sim weibull(\alpha, \sigma) \tag{7}$$

$$\alpha \sim gamma(\hat{\lambda_\alpha}, \hat{\zeta_\alpha}) \tag{8}$$

$$\sigma \sim gamma(\hat{\lambda_\sigma}, \hat{\zeta_\sigma}) \tag{9}$$

To sample from the posterior distribution $p(alpha, sigma|wind - data)$ we use the bayesian programming language and Markov Chain Monte Carlo (MCMC) sampler Stan via the python interface pyStan [Stan Development Team, 2014a]. Stan utilizes Hamiltonian Monte Carlo to efficiently simulate sampling from the posterior.

Other popular simulation tools are widely available like BUGS and JAGS, but we choose to use Stan because of the flexibility of its modeling language and the efficiency of its sampling routine. While the case outlined hear is quite simple, the model can be easily extended with Stan to include other sources of uncertainty relevant to the realistic investment decision.

The wind data is taken from actual wind measurement data at the appropriate height at the site of the proposed Andmyran wind farm. We use 1 year of hourly measurements - or 8780 data points to update the posterior distribution.

Figure 5 shows results from the MCMC sampling in the form of a histogram of draws from the posterior distributions of the shape parameter α in the top panel and the scale paramter σ in the bottom parameter.

To simulate wind speeds from the posterior distribution, we first draw values of α_{draw} and σ_{draw} from their posterior distribution. We then draw 8760 simulated wind speed observations from a $weibull(\alpha_{draw}, \sigma_{draw})$ distribution, representing a year of hourly observations. These values are converted to power output with the power curve function detailed earlier. The power output is then summed and returned. We then repeat this to generate a measure of uncertainty over the yearly wind conditions.

From this sample of of yearly wind conditions, we calculate a sample of expected losses for the second stage decisions of whether to invest or pass. The results are shown in the form of histograms in figure 6. The updated posterior beliefs indicates that the investment will be profitable with high certainty.

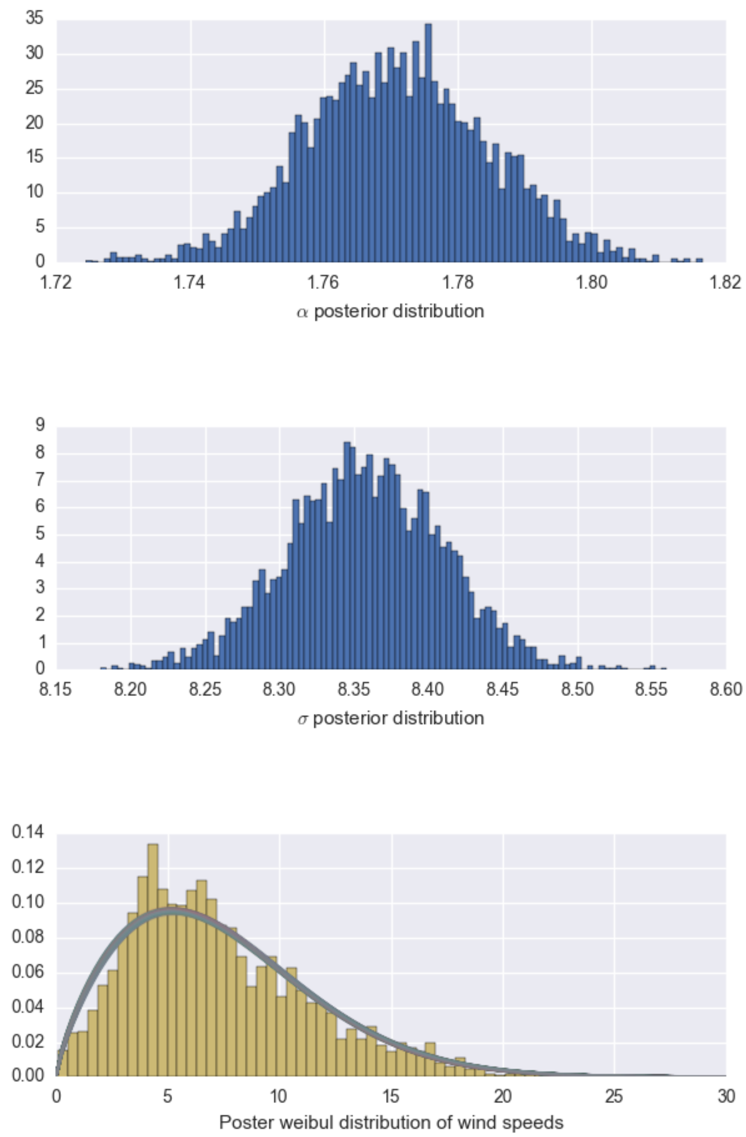

Figure 5: Posterior distributions of the α (top panel) and σ (middle panel) parameter of the weibull distribution. The bottom panel shows the weibul distribution with draws from the α and σ posterior distributions. The histogram of the actual data is overlaid.

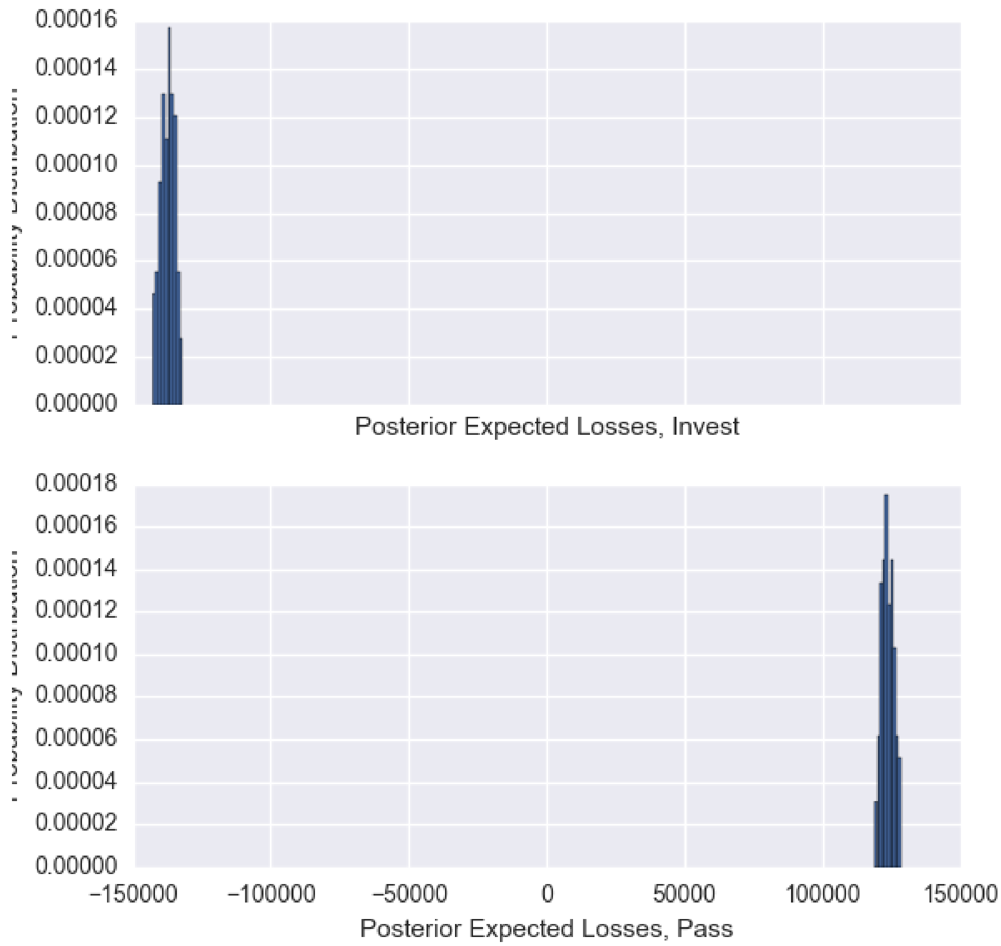

Figure 6: Posterior distribution of expected losses. Investing is now expected to be profitable with high certainty.

5 Conclusion

This article has three main contributions. First, it recognizes that wind power investments decisions, and potentially other renewable energy decisions, are subject to a large degree of spatial uncertainty. Even with the case study of a proposed site in a region which has some of the best average wind resources in the world, a high degree of uncertainty exists about the viability of any given project. While this point is more or less common sense in the industry, it has not been widely acknowledged in the literature, and it is not, to our knowledge, included in explicit academic models of wind power investment.

The second contribution is to model the sequential nature of a wind power investment by way of a Bayesian decision model where the spatial uncertainty can be modeled explicitly and where the option of waiting and gathering more data in order to update prior beliefs can be modeled naturally.

Finally, our implementation of the model using the open and flexible Bayesian modeling language and MCMC sampler Stan allows for straight-forward extensions that allows for me realism in the model. We hope and expect that the model could be easily extended by both other academics and investors looking for a way to formalize their decision making processes. We have posted all our code at `jmaurit.github.io#wind_invest_model`.

References

Stan Development Team. *PyStan: the Python interface to Stan, Version 2.5.0.* 2014a. URL `http://mc-stan.org/pystan.html`.

Stan Development Team. *Stan: A C++ Library for Probability and Sampling, Version 2.5.0.* 2014b. URL `http://mc-stan.org/`.

Vestas. Vesta V90 3.0 MW Brochure, 2015.

Waloddi Weibull. A Statistical Distribution Function of Wide Applicability. *Journal of applied mechanics*, 1951.