Stochastic Approaches to Modeling Alternative Energy Systems

Elizabeth Stoltzfus

Department of Industrial Engineering and Operations Research, University of California, Berkeley, California 94720

gair@uclink.berkeley.edu

I. Introduction

There are many sources of uncertainty that discourage potential investors from financing alternative energy systems. Only a small proportion of our power consumption is generated by renewable means. According to the California Energy Commission, 12% of power in California in 2001 was provided by alternative energy systems ("Power Content Label"). These alternative energy systems include biomass / waste, geothermal, small hydroelectric, solar, and wind. Each of these sources individually contributes at most 5% of the overall power. Pacific Gas and Electric Company projects that it will continue to purchase power in the same proportions in 2003. Consequently, potential investors turn first to well established non-renewable energy systems. They tend to view renewable energy projects as riskier and gravitate towards the more established alternatives: coal, gas, or nuclear power. As a result, traditional fuel systems dominate not only current production, but current investment as well.

There are a number of aleatory factors to which alternative energy systems are particularly vulnerable due to the relative lack of data and experience. Though factors vary between fuel sources, there are common themes among them. The quantity and reliability are subject to fluctuations in the environment. The conditions and level of utilization are uncertain. For example, for a proposed geothermal system, the reservoir capacity, pressure, temperature, and salinity can not be precisely gauged in advance. The performance of relatively untried equipment can not be projected as reliably. There is more uncertainty for renewable energy sources because the technology tends to be newer and is produced on a much smaller scale, resulting in less accurate performance predictions. In contrast to traditional systems, alternative energy systems often have a backup traditional fuel system to be used as necessary. The utilization level of auxiliary power is uncertain. All of these factors are time varying.

One way to make the implementation of alternative energy systems more appealing is to minimize the risks. This can be done by creating accurate models of the uncertainty so that investors can make better informed decisions. Stochastic models provide an excellent tool to mitigate the risk inherent in the creation of a renewable energy system. Despite their applicability, stochastic models have not been extensively utilized.

This paper examines the work of some engineers who chose to pursue an interdisciplinary approach to problem solving. They led the way to applying stochastic models in the area of alternative energy. The first paper discussed illustrates the use of a Markov model as applicable to solar energy. The next article presents an approach to modeling uncertainty in a geothermal system. A final paper addressing the risk factors in wind power is considered. Each of the articles is summarized, reviewed, and possible extensions are presented.

II. Solar Energy

A. Summary

Dr. Gerard F. Lameiro and Dr. William S. Duff examined the use of solar energy for space heating. They sought to evaluate the long term system performance based upon stochastic models. Their approach was to develop a stochastic model of the solar energy space heating system, program the model, evaluate the accuracy of the model, and, if accurate, assess the program computational efficiency.

They designed their model for a designated type of space heating system. The system analysis was based upon the assumption that a water storage system was used. This allowed them to incorporate known collector efficiency curves.

The variable inputs to the model are the parameters of a specific system implementation. They include climate, insolation, hot water demand, and architectural data. The load on the system depends on the location of the system: colder climates produce larger demands for heat. In addition, the individual preference of the user can significantly influence the amount of heat desired. The site of the system also determines the range of solar radiation available to be absorbed. The insolation and weather are cyclic with a cycle time of one year. The architectural data, in contrast, is fixed and is comprised of the collector area and storage mass.

The model includes four stochastic variables: insolation, ambient temperature, storage temperature, and level of hot water demand. Each of these variables changes over time. The authors chose to use an hourly time grid based upon the data that was available. They created a transition probability matrix for each of the variables. This was done by taking the continuous time stochastic processes and binning them into discrete time intervals. They chose to use a discrete rather than continuous time Markov chain.

The model's governing equations describe the relationship between the stochastic variables and the other system variables such as collector area. The hot water load is defined as a function of the ambient and indoor temperatures. The four transition probability matrices are combined into one. Specific values of the four different stochastic variables make up each state. The performance values are functions of the state and other system parameters. These parameters are: useful energy collected per hour, auxiliary energy rate usage per hour, and load per hour. Cumulative totals can be calculated over a given time period.

The authors created three potential types of solar heat usage styles because of the lack of available data. Thus, the model input specifies which of the three Markov chains to utilize as well as defining the corresponding states. The storage temperature is determined by the architectural data and the collector efficiency curve. Since the collector efficiency curve is a function of insolation level, it inherits its aleatory nature. The ambient temperature and insolation transition probabilities were determined by analyzing hourly data.

The model is intended to be used to predict performance of potential solar heating systems. The information to be evaluated is derived from the limiting probabilities of the combined transition probability matrix. The Ergodicity Theorem, Theorem 4.1 (Ross 179), was used to determine the stationary probabilities for the combined Markov chain. The resulting proportion of time spent in each state was then utilized to determine the long-term system performance.

One of the main challenges in constructing the model was determining appropriate state definitions. The states were found empirically for the four stochastic variables. The author balanced the need to preserve the level of detail available in the data and maintain the manageability of the model. The possible number of states for the combined Markov chain was restricted by the limited computational resources available to them. Up to a point, more states would provide more accurate results. At the time the article was published, the resource limit was the overriding concern. As a result, the authors used only one ambient temperature state and two insolation states.

The accuracy of the model was assessed by comparing the results with other software packages available. There were thousands of experiments performed with STOLAR, the model described above, and two alternative software products, TRYNSYS and FCHART. All three software products achieved comparable results. The STOLAR model ran an order of magnitude faster than TRYNSYS. The FCHART software ran an order of magnitude faster than STOLAR.

The authors' contribution lay in applying stochastic models in the area of solar heating. They argued that this led to a simpler and approximately equally accurate model. They also claim to lead the way in terms of determining the best means of discretizing the continuous variables for their specific application.

B. Review

The paper lacks sufficient detail for an in depth examination of the techniques employed. This makes it hard to verify the technical soundness of the model. For example, it is unclear how one of the key variables, storage temperature, is determined. There is no indication that the details were intended to be published at a later time nor were they included as an appendix.

The authors' explanations of their design choices are incomplete. The assumptions and design options are clearly stated throughout. In some cases, limitations are cited and explained. Elsewhere, the design choices and assumptions are stated without justification or discussion of the implications. For example, the model is based upon the assumption that water storage is used in the system. The authors do not indicate how common this type of model is, why they chose it, or the implications of the decision. Additionally, they state that they have left out factors that influence the system, but provide no explanation for this choice or information regarding the ramifications.

The authors build upon results that are currently available. The experimental results are compared against those of other models on the market. However, there are no indications of how valid the alternative models are or how well respected they are in the industry. No mention is made of the other software being checked for accuracy. There is no comparison of the inputs for each of the models. A significant difference in the amount and type of input needed could prevent the model from being used in practice. Additionally, the amount of flexibility available in each of the models is not compared. One or both of the alternatives may be much more widely applicable.

They make comments throughout the paper that seem to indicate a lack of familiarity with stochastic processes. For instance, they write that it is worth noting that they did not use Monte Carlo methods to determine the transition matrices. When the Ergodicity Theorem is stated, they do not state that the limiting probabilities must sum to one.

The authors claim that the model allows some flexibility through the input of variables and parameters. However, it is unclear how much modification is necessary to vary the information as the designers may have hard coded some of the characteristics of the data. For example, the time interval of the data is not listed as a parameter. If the user decides to change the interval, it is not clear if the code would need to be modified.

The logic behind the experimentation as well as how it was performed is not clearly defined. They ran experiments for three cities, presumably for a "normal" system. It is stated that a poorly insulated version of STOLAR was run for one of the cities. It is not stated why this version was run, how it was different from the "normal" system, and if a similar version was run on the other two models for comparison.

C. Improvements

There are a number of possible improvements that could be made to the model to improve the accuracy and flexibility. The most basic is to reformulate the model without the computational limits that were imposed. This would allow the number of possible states to be increased to the point where the amount of detail captured is optimal.

One of the most significant limitations of the model is that the cyclic nature of the weather and insolation is not addressed. By using Markov chains, the next state of these stochastic variables is assumed to only depend on the current state and neglects the dependence upon the time of year and the current longer term trend, such as drought. The model could be extended to incorporate this functional dependency. The weather and insolation could have separate Markov chains for each season. For further accuracy, a separate chain for the seasonal weather could be made for each common long-term trend. These new chains would then be incorporated into the overall transition probability matrix. The temperature and insolation are also on a daily cycle. The probability of transitioning from the current state to the next depends upon the time of day. For instance, the probability that the temperature will increase is much greater in the morning than during the night. The model could be extended to incorporate this by creating separate Markov chains for the different parts of the day: night, day, period around dawn, and period around dusk.

The model could be formulated as a continuous time Markov chain instead of a discrete one. This would allow the time between states to be incorporated. This could improve the accuracy by taking into account how the system responds to different rates of change. For example, the change in the hot water storage between the same pair of ambient temperature states would be inversely proportional to the rate of change.

The fact that the model's results match its competitors' implies that it shares their limitations. The model could be more accurately judged by comparing it with actual performance data.

Some of the techniques used to model the data could have been improved. The method of guess and check was used to determine the best interval when discretizing the stochastic variables. The ideal bin size could be found by looking for the best trade-off between bias error and statistical error.

There are a variety of different types of solar heating systems that are not addressed by this model. The model could be extended to address all of the common types of solar heating systems. For instance, the type of collectors used for space heating are concentrating and flat-plate (Kraushaar and Ristinen 159). This would require increasing the input data to include the type of system that the user seeks to examine. For example, the model could be extended to incorporate pebble bed storage.

II. Geothermal Energy

A. Summary

Bernardo Fortunato and Giovanni Mummolo use methodologies from operations research to evaluate the economic viability and impact of different approaches to geothermal systems. The authors use stochastic analysis of performance parameters combined with success criterion to determine the probability of success. The inputs and outputs of the model are defined as probability distributions. The model is designed to evaluate the net present value (NPV) of implementing a geothermal system given the specific conditions.

Two types of input data are defined: specified and unspecified. Parameters that are more accurately known, such as the cost of equipment in the future, are considered fixed point values. The lesser known values, like future energy prices, are approximated with a probability distribution.

The same probability distribution is used to model all the stochastic variables. The authors call it a triangular distribution. The distribution is completely defined by three values. The peak of the triangle is placed at the expected value of variable. The end points of the distribution (the other two corners of the triangle) are placed at the maximum and minimum possible values.

The Monte Carlo technique is used to generate possible sets of the stochastic variables in the system. The specific values that make up each set are determined by generating random numbers then finding the corresponding value using the cumulative probability function. The model generates an “adequate” amount of sets of values.

The NPV is a key variable in this model. It is calculated as the cash flow scaled by the appropriate interest rate to put it in terms of current currency. The interest rate was modeled as a fixed value.

The probability distribution of the NPV is determined using the results of the Monte Carlo method. Monte Carlo simulations of the stochastic variables that determine the NPV are run. The resulting NPVs for each set of input variables are binned. The discrete points are then smoothed into a continuous probability density function. Since the simulation does not provide sufficient points near the ends of the distribution, the ends are extrapolated instead of being determined by binning. The expected value is assumed to be at the peak of the curve.

A specific example was used to demonstrate the model’s capabilities. The measures of success are determined using the NPV probability density function. The probability of exceeding the target value is found by dividing the area under the curve above the target by total area. Similarly, the probability of failure, the NPV falling below zero, is found by dividing the area under the curve below NPV equal to zero by the total area.

The authors then move on to discuss their use of decision analysis in the model. Multiple criteria are used to evaluate and rank potential implementations of a geothermal system. A methodology called Preference Ranking Organization Method for the Enrichment Evaluation (PROMETHEE) was selected for the model. It was chosen for its simplicity and limited amount of data required.

PROMETHEE is made up of three steps. The first step is to do a pairwise comparison for all possible sets of two alternatives. In the second step, a multicriteria preference index is created. It is calculated based upon the results of step one and user- defined weights that express the relative importance of the decision criterion. Finally, in the third step, the alternatives are ranked.

The authors discuss the role of social-economic and technical characteristics in the decision process. These characteristics influence the options considered, the selection criteria, and, ultimately, the decision itself. Each of the exploitation schemes embody a number of potential uses such as greenhouse heating. Once the alternatives are determined, PROMETHEE is applied to perform the evaluation.

The criteria were decided based upon the conflicting interests of the anticipated stakeholders. The first criterion is the energy generated by the system. The magnitude of this energy is measured in terms of tons of oil equivalent in order to determine the impact of switching to a renewable source. The second criterion is the return on investment (ROI). The ROI is defined as the annual profits divided by the investment capital. The last criterion is the number of jobs created by the potential system. The authors state that they initially started with additional criteria such as environmental impact. Since the rankings for these additional criteria were approximately equal for each of the alternatives, they were not included in the final set of criteria.

The preference data is scaled in order to compare dissimilar criteria. This is done by defining two parameters: a lower and upper threshold. A higher preference index indicates a stronger preference. The lower thresholds for the indices are defined as the value below which the options are considered equally appealing. The upper thresholds define the cutoff for a strict preference.

The challenge of reaching consensus with multiple decision makers is addressed. Ideally, a compromise can be reached resolving any differences of opinion. When that approach is no longer an option, the groups' preferences are grouped into a joint set of preferences based upon the will of the majority.

The authors' contribution is to combine NPV estimation with decision analysis in the area of geothermal energy systems.

B. Review

The authors provide helpful background for understanding the motivation and context for the research. The picture becomes much less coherent once the methodology is discussed. There are a number of details of the model formulation that are unclear. The authors built their model upon previous results. They carefully cite all the previous research used. However, it is not always clear which results from the references were used.

The decision criteria for choosing between specified and unspecified data, as described above, are unclear. The authors state that both involve uncertainty, but the unspecified data is less accurately known. No explicit rule for allocating the data types to each category is provided.

The derivation of the probability distribution of the NPV could be more accurate. The function used to find the NPV is not provided in the paper. In addition, the accuracy of the probability distribution function for the NPV is not clearly defined. There is no mention in the paper of determining the error inherent in the probability density function generated. The paper states that an adequate number of points are generated to determine the probability distribution function. The two ends of the distribution are extrapolated. However, neither an explanation as to how the adequacy was established nor the method of extrapolation are stated. It is also unclear where, along the range of values, the extrapolation begins. The authors indicate that they calculated the cumulative probabilities by taking the area under the curve up to the desired point and diving through by the total area. Since the total area under the probability density curve should sum to one, it is unclear whether the authors forgot this fact or the area under their curve doesn't sum to one.

The paper states that a triangular distribution is used for all stochastic variables. The vague description of the distribution provided does not match the usual definition (Feller 502). The distribution used in the model is not symmetric about zero.

The authors' description of the selection criteria for alternative geothermal fields includes much greater detail than those of the rest of the document. The authors state that assumptions are reasonable without any justification. The naming convention used for some of the performance measures are misleading. For example, the amount of energy produced is referred to as "energy savings".

The two central concepts that form the basis of the model are the net present value (NPV) and the decision analysis. The authors cite other papers for each of these concepts.

C. Improvements

A large improvement in the model could be achieved by choosing more accurate distributions than the triangular distribution to represent the stochastic variables. This could be done by analyzing the data for each of the variables to determine the best distribution and the parameters for that distribution. The best fit is found by minimizing the error in the distribution. These distributions would then be used to generate a more accurate NPV. Because the NPV is utilized throughout the model, the improved accuracy would affect the entire model.

NPV accuracy could be further improved by performing more Monte Carlo simulations. The increased number of points generated could eliminate the need to extrapolate the ends of the distribution curve. Throughout the range of the NPV, the curve will be more accurate due to the increased overall number of points that are binned as well as the improved distributions for the variables upon which NPV is dependent.

The utilization of the NPV introduces even more error into the calculations. The mean should be used for the expected value instead of the usual mode. The cumulative probabilities were calculated using the area under the curve. Since this area is based upon smoothing, a more accurate approach would be to sum the contents of the bins that lie within the designated NPV range.

An improvement that could be made to the decision analysis is to incorporate some sensitivity analysis. This would allow the user to determine how much small changes input values could affect the decision recommended by the model.

The model is very general. The main concepts utilized are the NPV and decision analysis. A natural extension is to include the capability to analyze other types of energy systems, such as wind energy. This could be a helpful tool for a potential investor or community trying to decide which type of energy system is best suited to their needs. For such initial decisions, the model would likely prove useful despite its highly approximate nature.

III. Wind Energy

A. Summary

Dr. Goumas, Dr. Lygerou, and Dr. Papayannakis use stochastic analysis to evaluate the performance of an isolated wind power plant. The model is then applied to identify the most economically sound plant configuration.

The great variability in the energy production cost for wind power systems has prevented their widespread implementation. This variability is due to the wind potential, load required, and system configuration.

The paper addresses a specific implementation of a vertical axis wind turbine that is connected, by means of a gearbox, to a self-excited induction generator operating at constant frequency and voltage. The load is applied directly to the system. The surplus energy that exceeds load requirements is stored in batteries. When the battery charge level falls below a given value, a backup generator provides the energy to the load and the batteries until the maximum level is reached. There is a direct relationship between the storage system size and the reliability and operation continuity.

The technical and economic performance of a wind power system depends upon wind potential, load, and system configuration. Two of the key stochastic variables are the wind speed and the power required. Each is defined by a probability density function. A Weibull distribution is found to be a good match for the wind speed distribution, especially when the minimum value and mean are known. The electric power is defined as a function of the system parameters and the wind speed.

Two different designs are considered for the gearbox. One is to hold the transmission ratio between the turbine and the electric generator constant and the other is to use a variable transmission rate. A distinct distribution for the power difference is calculated for each option. Both are determined to be Rayleigh distributions.

The wind generator can supply power to the load if the wind speed is higher than the minimum value necessary for the turbine's operation and below the turbine's maximum speed. Once the maximum storage capacity is reached, the surplus wind power is dissipated unless a variable transmission ratio is used.

The amount of power produced by the generator is defined as the difference between the power produced and the power required. The distribution for the power difference is found using a transformation of variables.

The probability density function of the average daily power is found by randomly generating and analyzing 24-hour load power requirement curves. The average daily power is found to fit a Beta distribution.

The return is a function of the total annual cost and the energy provided. The total annual cost is comprised of fixed costs of purchasing the equipment and operating and maintenance expenditures. The elements of the system are considered to have the same fixed life except for the storage system and diesel generator. Their lives are expected to be shorter and their life expectancies are random variables that depend upon their frequency of use. These calculations incorporate fixed values for the equipment and discount rate. The amount of wind power generated and the fuel consumption by the diesel generator are stochastic variables. The fuel consumption depends upon the storage system capacity, wind speed, and transmission ratio. The difference in stored energy level over time is assumed equivalent to the energy produced during that time. In other words, the storage system is not dissipating energy. Fuel consumption is computed as the average value between the two limits corresponding to the two fixed and variable transmission ratios.

An example of the model inputs and outputs is provided. Various average and minimum wind speeds were evaluated. The results demonstrate how the model can be used and provide some sense of the system performance. The outcome shows how strongly turbine size influences fuel consumption. The unit cost of the energy is highly dependent on the turbine size and average wind speed.

Turbine size should be chosen to optimize the trade-off between equipment and operating costs. The turbine size also depends upon the average and minimum wind speed and the transmission ratio. For a fixed ratio, a larger turbine is helpful to offset low minimum wind speed.

The authors' contribution is the extension of the basic electric power performance measure to a more accurate model of system performance. Additionally, the model that they provide gives insight into how various environmental factors and component choices affect the performance of the system. Using the model, it is possible to evaluate the relationship of the wind speed variability and the optimal size of the wind system. The relative economic benefit of a single or multiple gearbox is also available for consideration.

B. Review

This article is much more technically sound than the previous two. While the first two articles reached conclusions about their models, this article reaches conclusions about the design of the system under consideration. There is much more detail and clarity provided in the text. Many more tools for stochastic analysis are utilized. The model is both more complex and more fully explained. The function definitions for the variables and parameters are clearly stated. The authors clearly state the data input into the model to achieve the stated results. They compared their results with actual data.

The authors provided thorough background for understanding the problem. However, some of the bold statements contained therein do not have corresponding references to support the information. For example, the authors claim, without stating the source of the information, that wind energy costs are similar to those of traditional energy sources in the United States.

The determination of the average daily power distribution could have been more clearly defined. The authors do not state how they decided that a Beta distribution with the given parameters was the best choice. Also, they do not mention examining the error inherent in the limited data from which it was derived.

The authors clearly delineate their assumptions and the model's limitations. For example, they share their reservations about the accuracy of the fuel consumption variable. They discuss the implications and how they addressed the issue.

The authors start with the simple and common definition for the power generated. They build their model upon it by adding refinements. Finally, they compare their results with the generally accepted formulation to measure the level improvement obtained.

C. Improvements

The model is very specific to the application for which it was intended. The wind energy system is evaluated for a small load in an isolated area. The model could be extended to be more generally applicable. The model could address less isolated areas by allowing the auxiliary energy to be provided by the power grid. This would mean having an option for whether or not to include an auxiliary energy source in the system. The model could also be extended to address much larger plants. In addition to the requisite modifications for scaling, the model would need to provide a corresponding option to have the production occur independently of the load. In other words, the load would not be directly connected to the generation system.

The model is also specific to the type of equipment available in Italy, where the research was performed. A set of various types of equipment available worldwide could be built into an expanded model.

Further research could be performed in the area of fuel consumption. A more accurate model of the fuel consumption would significantly improve the model since its uncertainty was one of the key model limitations.

The model allows the system performance to be examined as a function of the gearbox and the size of the generator. The current model only permits two different options for the gearbox, variable or fixed. More options could be added to the model of the gearbox. Additionally, the model could be extended to incorporate the analysis of the performance as a function of other system components.

IV. Conclusion

Stochastic techniques can play an important role in the realm of alternative energy by more accurately modeling the inherent risks. Historically, there have been few applications of stochastic processes to the area of alternative energy. This paper examined one of the applications to the areas of solar, geothermal, and wind power.

The first two articles provided a stochastic approach to problems that have already been solved. In contrast, the third article achieves new insights into the system through the use of stochastic methods.

The analysis of the wind power system began with the standard estimate for performance. The authors added many extensions to the model through the use of stochastic techniques. In addition to increasing the accuracy, they were able to examine the system performance as a function of various system parameters. The decision maker can use the model to gain an intuition for how the properties of the environment and different model component selections affect the system performance. For example, the authors derive a graph of energy cost versus turbine frontal area. The resulting clearly-demonstrated relationships allow the users to make much more informed design decisions.

The third article provides an example of how stochastic techniques can be used to improve our understanding and ability to accurately model alternative energy systems. Hopefully, this article will lead the way to other similar research in wind as well as other forms of alternative energy, ultimately leading to more extensive use of renewable energy.

References

1. Pacific Gas and Electric. Power Content Label (2003).

2. Lameiro, Gerard F., and William S. Duff. "A Markov Model of Solar Energy Systems." Sun: Mankind's Future Source of Energy: Proceedings of the International Solar Energy Society Congress: New Delhi, India: January 1978 3 (1978): 1514-1522.

3. Ross, Sheldon M. Introduction to Probability Models. 7th ed. Burlington: Harcourt Academic Press, 2000.

4. Kraushaar, Jack J., and Ristinen, Robert A. Energy and Problems of a Technical Society. 2nd ed. New York: 1993.

5. Fortunato, Bernardo, and Giovanni Mummolo. "Technical-Economic Optimization of a Wind Power Plant by Means of a Stochastic Analytical Model." Energy Conversion and Management 38 (1995): 813-827.

6. Feller, William. An Introduction to Probability Theory and Its Applications. Vol. 2. New York: John Wiley & Sons, 1971.

7. Goumas, M. G., V. A. Lygerou, and L. E. Papayannakis. "Computational Methods for Planning and Evaluating Geothermal Energy Projects." Energy Policy 27 (1999): 147-154.