Propagating Uncertainty in Solar Panel Performance for Life Cycle Modeling in Early Stage Design

Tomonori Honda¹, Heidi Q. Chen¹, Kennis Y. Chan², and Maria C. Yang³

¹Department of Mechanical Engineering, Massachusetts Institute of Technology
Cambridge, MA 02139
tomonori@mit.edu; heidiqc@mit.edu
²ATAC Corporation
Sunnyvale, CA 94085
kennis.sf@gmail.com
³Department of Mechanical Engineering and Engineering System Division, Massachusetts Institute of Technology
Cambridge, MA 02139
mcyang@mit.edu

Abstract
One of the challenges in accurately applying metrics for life cycle assessment lies in accounting for both irreducible and inherent uncertainties in how a design will perform under real world conditions. This paper presents a preliminary study that compares two strategies, one simulation-based and one set-based, for propagating uncertainty in a system. These strategies for uncertainty propagation are then aggregated. This work is conducted in the context of an amorphous photovoltaic (PV) panel, using data gathered from the National Solar Radiation Database, as well as realistic data collected from an experimental hardware setup specifically for this study. Results show that the influence of various sources of uncertainty can vary widely, and in particular that solar radiation intensity is a more significant source of uncertainty than the efficiency of a PV panel. This work also shows both set-based and simulation-based approaches have limitations and must be applied thoughtfully to prevent unrealistic results. Finally, it was found that aggregation of the two uncertainty propagation methods provided faster results than either method alone.

Introduction
A key aim of sustainable design for both the engineering design and AI communities is to develop design methods and tools that can aid cradle-to-cradle design, thereby minimizing environmental impact throughout the entire product life cycle. To this end, a number of research efforts have been made to quantify the environmental impact of the product over its life cycle, from design to retirement. Some of these metrics include Life Cycle Assessment (LCA) (Curran 1993; Pennington et al. 2004; Rebitzer et al. 2004; White and Shapiro 1993), Life Cycle Sustainability Assessment (LCSA) (Heijungs et al. 2010) and Life Cycle Commonality Metric (LCCM) (Wang and Tseng 2009).

Among these metrics, LCA has become a standard, but there continues to be research on developing it further (Cooper and Fava 2006; Pennington et al. 2004; Rebitzer et al. 2004). In particular, Pennington, et al (2004) considers the role of uncertainty in a life cycle model. Typically, the actual usage and disposal/recycle of a product cannot be predicted by a product’s designers, which means that there can be a considerable degree of uncertainty associated with any sustainability metric. Furthermore, creating a traditional LCA model is a data intensive process which requires considerable additional effort for new products, though there have been attempts to develop learning surrogate models to reduce building time (Eisenhard et al. 2000). Ideally, the fidelity (accuracy) and building time for such a learning surrogate LCA model should be balanced with irreducible uncertainties associated with sustainability metrics. An example of an irreducible uncertainty might be using a product under unanticipated weather conditions. In comparison, a better understanding of the inherited unavoidable uncertainties in a model will shorten the time to build it as there is limited benefit to improving the accuracy of a model if its irreducible uncertainty is very high.

This paper is a preliminary study of ways to propagate uncertainty into overall system performance to support sustainable design, with a focus on high uncertainty products. It compares set-based and simulation-based approaches, and aims to help design teams to evaluate the
benefits and limitations of sustainable design applications in AI. This work is conducted on a case study of an amorphous photovoltaic (PV) solar panel cell.

Background

There are many different uncertainties associated with engineering design. One type of irreducible uncertainty is uncontrollable variation that occurs during the manufacturing and usage stage. Designers cannot control or remove this type of uncertainty, and so must focus on managing it. The management of this type of variation is crucial for the development of sustainable design. Another type of uncertainty associated with engineering design is modeling uncertainty. This is directly related to fidelity of the model. Furthermore, the engineer can sometimes reduce this type of uncertainty. Work in understanding uncertainty has focused on classifying (Klir and Folger 1988; Thunnissen 2003), quantifying (Capaldi et al. 2010; Giunta et al. 2004; Russi 2010; Wojtkiewicz et al. 2001), propagating uncertainty into system performance (Feeley 2008; Frenklach et al. 2002; Phillips 2003; Thunnissen 2005), and optimizing design under these uncertainties (Allaire and Willcox 2010; Du and Chen 2001; Enevoldsen 1994; Lee et al. 2002; Liang et al. 2008; Rajnarayan et al. 2008; Tu et al. 1999).

In the early stages of design, subsystem and system models can vary widely in their levels of fidelity. Two contrasting perspectives on managing the fidelity of subsystem models (Klatt and Marquardt 2009) include a) simply creating the highest fidelity model possible (Kahrs and Marquardt 2008; Mogk et al. 2002; Tan and Li 2002; Tulleken 1993) and b) building approximate models to estimate the output of the high fidelity model by balancing computational cost with fidelity (DeLaurentis and Mavris 2000; Sasena et al. 2002; Wang 2003). These subsystem models may be approximated to ensure they match the fidelity of the rest of the system. However, such approximation must be balanced against the potential loss of accuracy of having a system level model with the highest fidelity for its subsystems (Prusha 2005).

This paper takes the view that creating a model that considers accuracy of system performance will reduce cost and effort. Therefore, it presents a method for creating models, that takes into account system-level fidelity. This method quantifies the impact of fidelity on system performance by estimating overall uncertainty, and in future work will also consider the role of subsystems in overall system performance. The goal of this work is to provide a method to aid design teams in allocating time and effort in improving critical subsystem models.

Methods

The steps for this experiment include propagating uncertainty in estimating solar radiation intensity using two different methods, comparing the PV power output (system performance), and combining the two methods to create a new approach.

Step 1 – Quantify uncertainty in estimating solar radiation intensity. An uncertainty distribution for solar radiation intensity was determined for each hour and month for the Dane County Airport in Wisconsin, USA using data from the National Solar Radiation Database (NSRDB) (National Renewable Energy Laboratory). This data source was chosen because it is a high uncertainty region with unpredictable weather in winter. This weather uncertainty was estimated using an empirical cumulative density function (CDF) and probability density function (PDF).

Step 2 – Create an initial PV efficiency model. Physical experiments was performed on an amorphous silicon photovoltaic cell to capture its efficiency as a function of solar radiation intensity.

Step 3 - Propagate uncertainties from the solar radiation intensity and experimental PV models. To determine the required fidelity of the PV model, uncertainties from the solar radiation intensity model and the current PV model were propagated using both a set-based approach (Agarwal et al. 2004; Saleghaftari and Rais-Rohani 2010; Thunnissen 2005; Ward et al. 1994; Ward et al. 1990) and a simulation-based approach using Monte Carlo Simulation.

Step 4 – Assess the benefits and limitations of the two methods. The aim is to determine a good method for obtaining the uncertainty bound.

Models

This study examines ways of decreasing uncertainty in the two key models: the solar radiation intensity model and the photovoltaic model.

Solar Radiation Intensity Model. Because weather for one year may vary drastically from one year to the next, even for same location, the uncertainties in solar radiation intensity data can be considered irreducible. More importantly, the inherent inaccuracy of weather prediction models may dictate the overall accuracy of the entire solar energy system. If the error of the weather and solar radiation intensity model is too large, minimizing errors for a photovoltaic model may be cost-ineffective and unnecessary.

The measured solar radiation intensity data for this study comes from Madison, WI via the NSRDB. Historical data has shown that for any given hour of any given day, there is more than a 25% chance that it is too dark to produce...
any useful power output. This is a good location to test how various types of uncertainties propagate through system performance, such as power output from a PV system.

The main challenge associated with fitting a probability distribution over solar radiation intensity data is the fact that there is finite probability that the intensity will be exactly zero. This is problematic for fitting the distribution because the probability density function must either contain the Dirac Delta function for continuous distributions, or it must mix discrete and continuous distributions. Furthermore, for many hours, the shape of the distribution is bimodal and asymmetric. Thus, because we have over 3000 samples for each hour, the empirical cumulative distribution was adequate to represent the uncertainty rather than fitting the distribution.

Photovoltaic Model. For this study, an experimental PV model was used to collect realistic data. A lux probe, lux meter, and solar panel were set up on a flat roof (Figure 1). The amorphous silicon solar panel used is the Sunforce 12V Battery Trickle Charger, with an area of 0.09 m². It lies horizontally on the roof, away from obstacles that might cast shadows. Output from the PV panel is sent to a resistor via a smart meter which converts the data into a digital reading of voltage, current, resistance, and power. Different resistances were used in the range of 1 to 950 ohms, and changed every morning so as to obtain a spectrum of data points. The experiment was conducted in October 2010 in Cambridge, Massachusetts, and again in December 2010 – January 2011 in Goleta, California. Only data after noon was used for the Cambridge location because shadowing from a flagpole caused the morning data to suffer from systematic error. The Goleta location experienced no such issue, and the full day of data was collected.

Sample results are shown in Figure 2A and 2B. Note that there is uncertainty associated with the experimental results due to naturally occurring factors like the angle of the sun, clouds and dust or pollen gathering on the surface of the solar panel. These uncertainties are taken into account in the analysis, and reflected in the model.

The data from both Cambridge and Goleta was combined. The fluctuating weather conditions at each location resulted in similar levels of uncertainty, and the data could be combined without loss of fidelity. Illuminance was converted to irradiance at a conversion rate of 93 lux to 1 W/m² (Zenith Solar).

For a range of irradiance intensities from 200 to 700W/m², the data points at that radiation intensity (±1W/m²) were isolated, and plotted on a voltage-current graph (Fig 2B). For each resistance, the power from the 5th to 95th percentile at intervals of 5 was obtained. The set of power values for the range of resistances at each percentile was then fit using the least squares method to Eqn 1 in order to get a curve that represents the voltage-current (v-i) relationship (Fig 2C) (Martil and Gonzalez 1992).

\[ i = a_1 e^{a_2 v} + a_3 \]  
*Eqn 1*

Where \( a_1, a_2, \) and \( a_3 \) are constants.

A least squares fit based on vertical offsets was used to fit the v-i curve instead of a fit based on perpendicular offsets in order to simplify the problem. As the number of noisy data points was reasonably large, the difference between vertical and perpendicular fits was small. This was checked by swapping the voltage and current for the independent and dependent variables and fitting the curve via the least squares method again. The two curves were plotted on top of each other and found to be similar in the maximum power range.

From this v-i fit, the maximum experimental power at that radiation intensity and percentile was obtained. Combining the power values from the range of intensities and percentiles, each percentile was fit with a linear polynomial least squares fit. Having derived the trend of power with radiation intensity, the power for an extended range of intensities from zero to 950W/m² in intervals of 25 W/m² was then calculated.
Figure 2: Photovoltaic model method: deriving the characteristic mean and standard deviation from the cumulative distribution function of power output.
For each interval in the extended range of radiation intensity, power values from the percentile curves were extracted and fit into a normal cumulative distribution function (Fig 2E). The characteristic values of mean and relative standard deviation were found for each curve. These were then plotted against radiation intensity (Fig 2F and 2G). The mean values were fit using linear regression model, and the relative standard deviation to an exponential curve, shown in Eqns 2 and 3.

\[ \mu = b_1 I + b_2 \]  
- Eqn 2

Where \( b_1 \) and \( b_2 \) are constants

\[ \frac{\sigma}{\mu} = c_1 e^{-c_2 \mu} + c_3 \]  
- Eqn 3

Where \( c_1, c_2, \) and \( c_3 \) are constants

Using measured solar radiation intensity data of Wisconsin in the winter months from December to February from the NSRDB, along with the mean and standard deviation fitted curves, a comparison of simulation-based and set-based approaches to uncertainty propagation was made.

**Uncertainty Propagation**

The goal is to determine how uncertainty propagates over the whole system. In this case, uncertainty is propagated from the solar radiation intensity data into total power output over Madison WI for one winter.

**Monte Carlo Simulation.** Uncertainty was propagated by sampling from the distribution obtained for both the solar radiation intensity model and the PV model. Using a sample size of 10,000, random numbers were generated from a uniform distribution and mapped to solar radiation intensity and power output using the mean and relative standard deviation obtained from the cumulative distribution functions created in the previous step. One sample included power outputs from all the hours of daylight for all the days of winter. These were then summed to obtain 10,000 samples of total output power over the whole winter. From this, the probability density function of output power was plotted.

**Set Based Approach Using Percentiles.** Using the cumulative distribution functions of solar radiation intensity, different percentile bounds of radiation intensity were obtained, such as [5 percentile, 95 percentile], and [25 percentile, 75 percentile]. Five values of radiation intensity at the 5th, 25th, 50th, 75th and 95th percentile bounds were calculated for each hour of daylight. The power output uncertainty was propagated for each upper and lower bound with the same percentile bounds, but applied to the cumulative distribution functions of power output given radiation intensity. The total power over each day was summed, and multiplied by the duration of winter (89 days).

**Results**

By looking at how different lines from the percentile result cluster together (Fig 3), we can determine the significance of each uncertainty. Because the percentiles for PV uncertainty are clustered near each other compared to the percentiles from solar radiation intensity uncertainty, this implies that uncertainty in solar radiation intensity is much more significant than the uncertainty in the efficiency of the PV panel. This makes sense because the impact of sunny vs. cloudy can make drastic differences in the power output. This result shows that by propagating percentiles, we can determine dominating uncertainties. This also shows that the fidelity of the PV model is accurate enough for computing the daily power output.
Figure 3 shows that blindly applying the Monte Carlo approach can give misleading information. It is shown that the probability for zero output from PV system within a day is zero. This result implies that Madison, WI will not have day in the winter season during which it snows or rains throughout the day. In real life, this is unrealistic, but this highly unlikely result is caused by standard assumptions for independence of the events. In real life, the probability for rain or snow in the next hour is highly influenced by current weather conditions. Thus, there is a high degree of coupling between the hourly solar intensities that this Monte Carlo model is missing. Furthermore, the actual transition probability between hourly solar radiation intensity is impractical to obtain.

On the other hand, when we look at seasonal total uncertainty (Fig 4), we see another side of the story. We can observe that the lower bound created by the 5th and 25th percentiles leads to zero total solar radiation intensity. As Madison, WI is not located at a very high latitude, this is an unrealistic result. The cause of this phenomenon is that fact that a set based approach assumes perfect correlations between worst case scenarios. In other words, this means that if it snows on the first day of winter (no sunlight), then there will continue to be no sunlight for the rest of winter. However, the assumption for independence is more realistic when we are considering a longer duration like a whole winter. Thus, this shows that assumptions are critical when we propagate the uncertainties.

Another interesting question is whether we can combine these approaches together coherently. It can be shown that the set-based approach using percentiles and a simulation-based model are not necessary mutually exclusive, and that they can work together effectively. In this example, because the uncertainty for radiation is more significant than the uncertainty of the PV model, we can propagate the cumulative distribution of solar radiation intensity into power output using a percentile based approach while ignoring the uncertainty in PV efficiency (see Fig 5 and 6). After obtaining the hourly cumulative distribution, we can propagate this information using Monte Carlo Simulation to get the power output of PV system performance. The result converges to a distribution similar to a pure Monte Carlo Simulation. This is useful as a percentile based approach can be more efficient even for just a portion of uncertainty propagation. Furthermore, we can remove the Monte Carlo Simulation to predict the total power distribution from the hourly cumulative distribution of power. By utilizing the Central Limit Theorem, we can deduce that the distribution will approximate a Gaussian distribution. Furthermore, the mean and variance for independent and identically distributed events will grow proportionally to the number of samples. This result implies that rather than blindly applying any one technique for propagating uncertainty, the designer should understand the nature of uncertainty in order to choose the most appropriate techniques. Finally, considering just the median or mean value, will be provide the designer with misleading information, leading to poor product development choices.

**Conclusions**

The results from the study show that designers of high uncertainty products and systems should aim to better understand the nature of the uncertainty before applying set-based or simulation techniques. Simply considering percentile, mean or median values will lead to misleading results. Dominating uncertainties can be determined by first propagating them, and scrutinizing the underlying assumptions of the methods. Set-based and simulation methods may be effectively combined in order to reduce...
time and cost of propagating uncertainties so as to improve critical subsystem models. In the case of sustainable design of PV cells, a better knowledge of uncertainty propagation will allow designers to better approximate the power output, especially for regions with fluctuating weather conditions where the current PV use is low.

Future work includes adding components such as the inverter subsystem model to the PV cell subsystem model, then considering the entire fidelity of the system as a whole. The impact of subsystem fidelity on system performance may then be quantified by estimating overall uncertainty in the system.

Acknowledgments

This work was supported in part by the Center for Scalable and Integrated Nanomanufacturing (SINAM) at UC Berkeley which is an NSF Nanoscale Science and Engineering Center. The opinions, findings, conclusion and recommendations expressed are those of the authors and do not necessarily reflect the views of the sponsors.

References


