

ESTIMATING THE UNCERTAINTY IN LONG-TERM PHOTOVOLTAIC YIELD PREDICTIONS

Didier Thevenard¹, Sophie Pelland²

¹ Numerical Logics Inc., Waterloo, Ontario, Canada

² CanmetENERGY, Natural Resources Canada, Varennes, Québec, Canada

Corresponding e-mail: spelland@nrca.gc.ca

Keywords: photovoltaic systems performance; simulation; uncertainty; PV yield prediction; statistical modeling.

ABSTRACT

The uncertainty in long-term photovoltaic (PV) system yield predictions was examined by statistical modeling of a hypothetical 10 MW AC photovoltaic system in Toronto, Canada. The goal of this research was to develop a methodology for estimating the uncertainty in PV yield predictions and to identify avenues for reducing this uncertainty. In this case study, uncertainties were estimated to be about 3.9% for year-to-year climate variability, 5% for long-term average horizontal insolation, 3% for estimation of radiation in the plane of the array, 3% for power rating of modules, 2% for losses due to dirt and soiling, 1.5% for losses due to snow and 5% for other sources of error. Uncertainties due to ageing and system availability were also considered. By performing statistical simulations with the Solar Advisor Model software, it was found that the combined uncertainty (standard deviation) is approximately 8.7% for the first year of operation, and 7.9% for the average yield over the PV system lifetime. While these numbers could vary significantly from one system to the next, the methodology developed is widely applicable. Moreover, a simpler methodology was also explored which should yield quick and fairly reliable estimates of uncertainty. Finally, avenues for reducing yield uncertainties were identified, including: increasing the reliability and resolution of solar radiation estimates, including measurements of irradiance in non-horizontal planes at high quality ground stations, reducing the uncertainty in module ratings and investigating losses that have not been well documented such as those due to dirt, soiling and snow.

1. INTRODUCTION

In 2006, the Province of Ontario (Canada) introduced subsidies for photovoltaics (PV) and other renewables via the Renewable Energy Standard Offer Program (RESOP). This was followed in 2009 by a Feed-In Tariff (FIT) program which led to a quick uptake both for residential and commercial systems and for Megawatt-scale PV parks ([1], [2]). As the number and scale of PV systems installed grows, so does the size of the associated investments, and the related opportunities and risks. It therefore becomes increasingly important for developers, investors and financing institutions to identify and minimize uncertainties that could affect the viability of their projects.

The uncertainty in PV system yield will be the focus of this article. Other uncertainties that can affect the cost and viability of a project include the eventuality of change in government policy, issues related to grid connection, uncertainties in capital and

construction costs, and the possibility of delays, to name a few. These uncertainties are discussed extensively in [3].

This article is structured as follows: Section 2 discusses the individual variables that contribute to PV yield uncertainty; Section 3 describes how the individual uncertainties from Section 2 were combined in a set of statistical simulations used to evaluate the global uncertainty in predicted yield for a typical PV park; Section 4 provides results and discussion, and Section 5 some concluding comments.

2. INDIVIDUAL VARIABLES CONTRIBUTING TO PV YIELD UNCERTAINTY

Modeling the uncertainty in long-term PV system yield requires information about the possible range or distribution of each of the variables involved in the model. These are discussed below, and form the basis of the statistical simulations in Section 3.

Note that in this study, only crystalline silicon PV modules were considered. The problems inherent to other technologies (e.g. amorphous silicon) were not covered, although the study could be extended to them. Note also that the term “uncertainty” in this paper will be used to designate root mean squared error (RMSE), unless otherwise specified.

2.1 Uncertainties in the solar resource

2.1.1. Long-term average insolation

There are a number of different sources of solar radiation data in Canada, including measured values, values estimated from terrestrial models, interpolated values, and values estimated from satellite-derived models (see [3] for a review). We will restrict our attention here to the Canadian Weather Energy and Engineering Data Sets (CWEEDS) produced by Environment Canada, and on typical meteorological year files derived from CWEEDS (these are known as Canadian Weather for Energy Calculations (CWEC) files).

The latest version of CWEEDS, produced in 2006, contains hourly data from the years 1953-2005 (at most) for 235 sites. The data include solar radiation, dry bulb temperature, wind speed, etc., and are freely available online at [4]. For about 40 sites, the radiation data is based on measurements from Environment Canada ground stations. For the remaining sites, radiation is generated synthetically using a model which relies on earth-sun geometry, cloud cover information (cloud amount and type by layer), and other parameters. The mean bias error (MBE) of the model is estimated at 5% or less and the RMSE is estimated at about 30% on an hourly basis, 16% on a daily basis, and 7% on a monthly basis [5]. As for the measured data, Myers et al. [6] evaluate the uncertainty of broadband irradiance measurements at about 3 to 5%¹.

¹ Myers et al. define uncertainty as the range associated with a 95% confidence interval.

2.1.2 Yearly and 20-year solar insolation variability

In an analysis of satellite-derived radiation data in the Mediterranean and Black Sea regions, Sürü et al. [7] have shown that yearly global horizontal insolation displayed low inter-annual variability, typically in the range 4 to 6%. This variability was found to be lowest in arid regions and highest (up to 10%) in coastal and mountainous regions. Similarly, standard deviations of yearly insolation for 13 Canadian sites based on CWEEDS data was found to range from 2.5% to 5.0% [3].

Solar radiation is also subject to decadal cycles and other long-term trends, however these seem very small compared to the other uncertainties discussed here. For instance, the standard deviation of CWEEDS solar radiation data averaged over 20 consecutive years for Toronto is only 0.44%; this uncertainty was therefore neglected in our analysis.

2.1.3 Uncertainties introduced by the use of transposition models

Transposition models calculate solar radiation in the plane of the PV array, given solar radiation on the horizontal surface. The hourly or sub-hourly models require a breakdown of solar radiation into diffuse and beam components, which is often done with a model as well.

A study by Gueymard [8] using one-minute measurements from Golden, CO, showed that when only global irradiance is known – i.e. the diffuse and beam components are modelled and the albedo is estimated – the MBE tends to be in the range of 0 to -6% for a South-facing surface with latitude tilt, and the RMSE in excess of 10%. Gueymard showed that there is in general no *best* model to use and that the results will vary according to the combination of models used, the characteristics of solar radiation at the site, ground albedo, etc., so it is difficult to eliminate or even reduce the error.

2.2 Uncertainties in PV system performance

2.2.1. Module rating

It is frequent for the measured power of a PV module to differ markedly from its nameplate power, with modules typically performing below rather than above their rated power at standard testing conditions (STC). For instance, module tests in the United States ([9], [10]) and Australia [11] revealed differences between measured and rated module STC power in the range of +4.9% to -19.7%, with average differences of about -3% to -5%.

The lack of strict requirements in current PV module standards may partly explain this situation: The current (2nd) edition of the IEC 61215 standard – which specifies the requirements for the design qualification and type approval of terrestrial crystalline PV modules – lacks a pass/fail criterion that would ensure that the modules meet the nameplate rating given by the manufacturer. As long as the degradation of the measured maximum output power does not exceed a prescribed limit after each test (usually 5%), and the modules meets a number of additional criteria described in the standard, the module design is deemed to have passed the qualification tests. This loophole will be fixed in the next edition of IEC 61215 under development by requiring that modules meet

a minimum rating value specified by the manufacturer. Meanwhile, the UL 1703 PV module safety standard, and the equivalent ULc 1703 document recognized in Canada, require that, under standard test conditions (STC), the power output of the modules tested be at least 90% of their nameplate rating. This criterion has however been criticized as being too permissive.

Note also that there is currently a shift among some PV module manufacturers to move towards positive tolerances for the power rating of their products, such as [0%, +5%] instead of $\pm 3\%$.

2.2.2 Degradation of PV modules

Degradation of PV modules is actually a combination of two phenomena:

1. An initial, very rapid decrease in efficiency within the first few days of exposure.
2. A long-term, gradual decrease in efficiency over the years.

Initial degradation

In a field test of 66 crystalline silicon modules, Dunlop et al. [12] observed initial degradation of $2.6\% \pm 1.3\%$. Since this degradation occurs only after the first few days of module exposure to natural light, it is not accounted for by indoor tests of module STC power and it is currently up to manufacturers whether or not to include it in their module tolerance ratings.

Long-term degradation or ageing

In a field test of 204 crystalline silicon PV modules, Skoczek et al. [13] found that 70% of modules had an annual maximum power degradation rate lower than 0.75%, with only 35 out of 204 modules tested losing more than 20% of their rated power after 25 years. Ageing was fairly consistent among modules of the same type (i.e. manufacturer / construction). Based on this study and others [3] it seems reasonable to assume that module degradation will typically be in the range of 0.3 to 0.8% per year.

2.2.3 Availability

Annual availability was around 95% for 116 systems in Germany, Italy and Switzerland [14]. For recent installations, over 90% of systems had availability greater than 90% and about 55% had availability greater than 99% [14]. For a 3.51 MWp plant, Moore and Post [15] report overall system effective availability of more than 99.4% over a five year period.

2.2.4 Presence of snow

Snow was reported as a factor for high losses in winter in an analysis of 500 systems in Japan [16], but these still accounted for less than 2.2 % overall on a yearly basis. Another Japanese study by Sugiura et al. [17] looking at 25 sites estimated annual losses due to snow at less than 0.7% in northern Honshu (near latitude 44°N) but as high as 3.5% in Hokkaido (a few degrees further North). Meanwhile, a study of a 1 MWp system in

Munich, Germany estimated annual losses due to snow at 0.3-2.7 % [18]. Another study of two systems in Austria found losses of 1% for one system and 3.3% for another, at 1550 m elevation [19].

There are numerous parameters at play: type of snow (light, heavy), age of snow, radiation, temperature, tilt angle of the modules, type of mounting system, distance of modules to the ground. The complex interactions of these phenomena make it difficult to generalize the findings of the above studies, and render a more precise calculation of snow effects difficult.

2.2.5 Dirt and soiling

Like the effects of snow, those of dirt and soiling depend on the site and climate, and are difficult to model or extrapolate from case studies. As described by the IEA PVPS Task 7 [20], module soiling can result from various mechanisms: pollution, the accumulation of dust or pollen, bird droppings or the growth of lichen (particularly at the lower edge of framed modules). They note that bird droppings represent a serious problem as, contrary to dust, they tend not to be washed away by rainfall. They mention that the average impact on yearly energy production for the systems in Germany's "1000-roofs-Programme" was typically less than 2%, but reached up to 2 to 6% for modules with a tilt of 30° or less and high horizontal frame or cover profiles. In the worst case a string delivered 18% more power after cleaning.

By studying PV systems in the South-Western USA, Kimber et al. [21] showed that soiling involves a complex interaction of dust and rain. There is a marked decrease in module efficiency during the dry season. Similarly, systems located in deserts are very susceptible to dust build-up, with performance loss ranging from 0.1 to 0.3% per day. Meanwhile, systems receiving frequent rainfall (at least once a month) do not experience significant losses. The average annual loss according to the model developed by these authors ranges from 1.5 to 6.2%. Kimber [22] confirmed that figure in a controlled experiment of three systems in Los Angeles, CA (one washed, the other two not) with annual losses of 3.5% and 5.1%.

Hammond et al. [23] note that the effect of soiling increases with the angle of incidence, from 2.3% under normal incidence to 7.7% when the angle of incidence reaches 56°.

2.2.6 Shading

Some row-to-row shading cannot be avoided in multiple row systems. This shading can be minimized by leaving more space between rows, and using a lower PV array tilt. The effect of shading is best estimated for the location under consideration with the help of a simulation program. Very little or no uncertainty is attached to this as the occurrence of shading can be precisely modelled, so one can design systems to be shade-free most of the time. Nevertheless, array performance under partial shading conditions is still difficult to evaluate.

2.2.7 Other PV system losses

Other losses worth mentioning include thermal losses, spectral losses, losses due to reduced performance at low irradiance, losses due to reflection on the front surface of the modules, wiring and mismatch losses and losses due to DC to AC conversion by the inverter. On a yearly basis, thermal losses are around 2 to 10%, losses due to reflection are around 0 to 5% and spectral effects are usually in the range of +1 to -3% [24]. These losses are either small and/or can be well modelled by simulation programs (ex: thermal, angle of incidence and inverter losses), so that their individual contribution to the uncertainty in annual yield estimates is modest. The overall uncertainty in these “other losses” is estimated to be of the order of 3 to 5%.

2.2.8 Post-inverter losses

There are two main types of loss that can arise between the inverter and the connection point to the public grid (PCC) where electricity generation is metered: wiring losses and transformer losses. Wiring losses are regulated in Canada by Rule 8-102 of the CE Code [25], which mandates a maximum of 3% voltage drop at nominal power for the branch circuit and a maximum of 5% total voltage drop between the PCC and the point of utilization (inverter).

The main variables affecting voltage drop in a given system are the selection of the nominal resistance of the wires and temperature fluctuations. Knowing the wire size and length, detailed simulations can account for the dependence of wiring losses on the inverter output power. However, accurately accounting for conductors’ temperature variation is definitely more demanding and is beyond the capabilities of PV simulation tools; this introduces uncertainty into wiring loss estimates. In what could be considered a worst case scenario of -20 to 90°C temperature span, assuming a 5% maximum voltage drop design, losses would vary between ~2 and 3.5%. On the other hand with a more realistic 1% maximum voltage drop design, losses would only be in the range of ~0.5 to 0.75%.

Transformer losses are divided into two parts: no-load (core) losses and ohmic (winding) losses. No-load losses are constant, and depend primarily on the magnetic material and geometry of the core. Meanwhile, ohmic losses are proportional to the loading of the transformer and depend on the material and size of the wires used in the transformer windings. Nominal losses (no-load and ohmic) can be estimated based on manufacturer specifications. However, these are subject to some uncertainty due to variability between transformers of the same type [26]. As a rough estimate, for a 1 MVA unit and production typical of a 1MW PV system in Ontario, an amorphous metal based transformer will generally have total losses of around 0.85-1%, a good silicon steel based transformer around 1.5-2% and a standard transformer around 2-3%.

To summarize, for given wiring and transformer specifications and design, uncertainties in post-inverter losses are relatively modest (of the order of 1% or less).

2.3 Modeling uncertainty

In addition to the uncertainties mentioned in sections 2.1 and 2.2, modeling of the system itself with simulation tools is a source of uncertainty. The accuracy of a given PV system model is difficult to estimate, since it is difficult to distinguish between modeling errors and errors in the model inputs. Estimates of this uncertainty therefore vary widely, but typically range from 3 to 5% if the modules are well characterized. For example, validations of the PVSYST model against data from seven well-characterized systems yielded mean bias errors with magnitudes ranging from 0.7% to 12.8% [27].

3. STATISTICAL SIMULATIONS OF LONG-TERM PV YIELD

Statistical simulations were used to combine the uncertainties described in the previous sections into a global uncertainty in photovoltaic yield predictions. This was implemented by using the parametric and statistical analysis options of the Solar Advisor Model (SAM) software [28]. SAM's statistical modeling uses a variant of the Monte Carlo method called the Latin Hypercube Sampling or LHS method. The method was illustrated for a solar farm located in Ontario, Canada. Starting from a base case, we studied first the uncertainties associated with the first year of operation and then the uncertainty associated with the performance of the system over a 20 year lifetime, with ageing and system availability taken into account.

3.1 Base case

The base case is a 10 MW AC solar farm located near Toronto, ON. The system comprises 20 SMA SC500U inverters rated at 506 kW_{ac}, and 60,160 BP Solar SX 3200N modules rated at 200 W_p; each inverter is connected to 188 strings of modules in parallel, each string having 16 modules in series. For the purpose of simulation only 1/20th of the system (i.e. one inverter) was considered. As SAM does not allow detailed modelling of post-inverter components (transformers, wiring, etc.), these were excluded from the current analysis, and the post-inverter derating set to one. Simulation parameters for this subsystem are summarized in Table 1. The pre-inverter derate factor in Table 1 was chosen in part to yield a realistic performance ratio for a large-scale PV park of the order of 0.8, as reported for instance by Moore and Post [15] for a 3.51 MW_p system in Arizona (see also section 3.2.6 on how the various deratings add up to the value listed in Table 1).

Table 1 - Default system parameters.

<i>Parameter</i>	<i>Value</i>
Inverter	SMA SC 500U
Inverter size (kWac)	506.7
PV Modules	BP Solar SX 3200N
Modules per string	16
Strings in parallel	188
Array peak power (kWp)	601.4
Tilt	35°
Azimuth	Due South
Pre-inverter derate factor	0.85
Post-inverter derate factor	1
Ground reflectance	0.1 (no snow) 0.15 (with snow)
Climate file	Toronto Airport (WBAN 04714) CWEC file

3.2 Modeling uncertainties with the Solar Advisor Model

3.2.1 Uncertainty in solar radiation

While an hourly simulation was performed to estimate annual PV system yield as accurately as possible, annual yield is to a good approximation primarily influenced by the yearly total insolation in the plane of the array, and relatively insensitive to the underlying hourly distribution. Thus, solar radiation uncertainty was represented by the uncertainty in the yearly total insolation. Although mostly measured data was used in this analysis, an uncertainty of 5% was modelled as representative of a sub-optimal case where data must be either interpolated from other stations or derived from satellite observations. The uncertainty was assumed to follow a normal distribution.

SAM does not allow the user to stochastically modify solar radiation read from the climate files. The uncertainty associated with solar radiation therefore had to be introduced by varying the input files manually. The procedure starts by dividing the standard normal cumulative distribution function (mean of zero, standard deviation of 1) into five regions of equal probability, and taking the central point of each region as the representative sample for which simulations were run (this is somewhat akin to stratified sampling; see [29]). This is illustrated in Figure 1. Then, the climate file used in the simulation was modified to generate five files corresponding to these five points on the cumulative distribution function (CDF), by multiplying *all* irradiance values in the climate file by a constant factor according to the point considered on the CDF. For example, the fourth point on the curve corresponds to a CDF of 0.7 and is equal to 0.5244; given the 5% uncertainty considered, the multiplicative factor in this case was $1 + 0.5244 \times 0.05 = 1.0262$.

This method of accounting for uncertainty in solar radiation is somewhat simplistic, in particular because it can potentially lead to unrealistically high values of solar irradiance. However it can receive a simple physical interpretation. For example, the multiplicative

factor applied to all radiation values can be viewed as an improper calibration of the pyranometers measuring solar radiation. It also has the merit of being easy to combine with straightforward Monte-Carlo simulations to study uncertainties associated with other parameters.

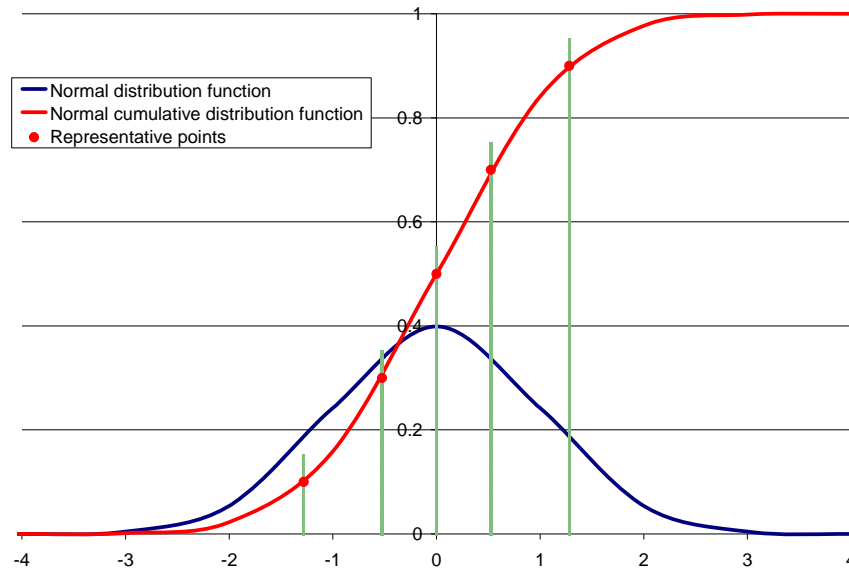


Figure 1 - Points chosen on normal CDF

3.2.2 Transposition model and dirt

As in the case of global horizontal radiation (3.2.1), the transposition model contributes uncertainty to the simulation primarily through its long-term bias. Following the discussion in Section 2.1.3, the uncertainty associated with the transposition model was therefore simply represented by a normal distribution with a mean of -2% and a standard deviation of 3% (-2%, 3%). As for dirt, it is believed that rain will wash the arrays often enough that the derating should be lower than in California; however some sites may be more susceptible to dirt because of the proximity to busy roads, agricultural operations or heavy industries. This was reflected by using a fairly significant standard deviation (2%) on top of an average loss of 3%. The uncertainty associated with dirt was therefore represented by a normal distribution (-3%, 2%).

These two uncertainties affect the amount of solar radiation available in the plane of the array, before it reaches the solar cells. SAM does not allow the user to explicitly apply stochastic variations to this quantity. However, the program can be tricked into doing so by modifying two of the parameters of the PV module performance model, namely the reference irradiance and the NOCT irradiance (by modifying the latter, one simulates the effect of a lower or higher reference irradiance on cell temperature). The two uncertainties for transposition and dirt were combined into a single normal distribution; the average of the distribution is the sum of the averages of each individual distribution (i.e. -2%-3%=-5%), whereas its standard deviation is the square root of the sum of the squares of the standard deviations for each distribution [30], (i.e. $[3\%^2+2\%^2]^{1/2} = 3.61\%$).

3.2.3 Power rating of modules

Uncertainty in the power rating of the modules was modeled as a normal distribution (-3%, 3%). The 3% standard deviation used here is typical of the tolerance usually given by module manufacturers. It assumes that the modules will indeed perform according to their rating, which is the case for most modules from reputable manufacturers (large PV developers may also be able to enforce this to some extent). The -3% bias is typical of light-induced degradation of PV modules in the first few days of operation.

3.2.4 Albedo

The albedo was represented by a uniform distribution between 0.1 and 0.15. This range is simply an educated guess; it accounts for the fact that rows beyond the second one receive little reflected solar radiation, so albedo should be lower than the often-used value of 0.2. In SAM, uncertainty can be applied directly to the albedo. Note that albedo in the presence of snow was varied in the same proportion; a correlation between the two uncertainties was set up with SAM.

3.2.5 Snow, modeling errors and other

As indicated earlier, the presence of snow seems to have a limited impact on energy production. The resulting uncertainty was modeled as a normal distribution with a standard deviation of 1.5% around an average of -2%, which seems representative of the values found in the literature. Modeling errors and the remaining PV losses (including post-inverter losses and losses discussed in 2.2.7) were represented by a normal distribution (-5%, 5%). The -5% bias represents modeling errors such as mismatch losses and wiring losses [31], while the 5% standard deviation is typical of the uncertainty associated with modeling, as discussed earlier.

These two errors were combined and were applied as a post-inverter derating. This is probably the most appropriate way to take them into account, since they represent global errors that cannot be attached to any particular sub-model in the program (snow is very seasonal, and is therefore not treated in the same way as dirt, although the mechanisms are somewhat similar). The resulting error is a normal distribution; the mean and standard deviation were combined as indicated earlier (-2%-5% = -7%, $[1.5\%^2+5\%^2]^{1/2} = 5.22\%$).

Finally, the uncertainties due to ageing and failures/outages were treated separately, as discussed in Section 3.4.

3.2.6 Number of realizations required

SAM was run in statistical mode with the various distributions described above. It should be noted that the means of the distributions were ignored in SAM, i.e. all distributions were centered around zero. The reason is that the base case already includes a pre-inverter derating of 15%, which is equal to the sum of the biases from the various distributions. *Not* centering the distributions around zero would be equivalent to counting the biases *twice*.

The number of realizations necessary to represent properly the uncertainty of the yield is not easy to determine. The Appendix shows that the *average* yield determined from statistical simulations is estimated with a confidence inversely proportional to the square root of the number of realizations, and that the *standard deviation* of the yield seems to follow the same pattern. The Appendix indicates that using 100 realizations provides a reasonable confidence in the results while keeping the run-time short (less than 10 minutes). These 100 realizations have to be run for each climate file considered.

3.3 Modeling uncertainty in year one yield

The first set of statistical simulations aimed at modeling the uncertainty in the first year of operation. For this, one has to combine the uncertainties due to system parameters and modeling (modeled with SAM as described above and summarized in Table 2) with that due to yearly climate variability.

When only long-term system performance is of interest, it is customary to run simulations using a so-called 'typical meteorological year', which is created from a long-term climate record (30 years in this case) by concatenating months selected from individual years. The months are selected based on statistical criteria, to be representative of long-term conditions in terms of solar radiation and average temperature.

On the contrary, when the year-to-year variability of solar radiation is of interest, it is best to run simulations for individual years from the 30-year time span. This however leads to a very large number of simulations. To speed things up, a CDF of yearly yields for the 30-year period was calculated (see Figure 2), and 10 years were chosen using a stratified sampling procedure to statistically represent the whole range of yearly solar conditions and yields expected at the site. The years selected by this procedure were 1972, 1970, 1984, 1982, 1985, 1979, 1975, 1974, 1960 and 1978. The procedure to account for uncertainty in solar radiation, explained previously, was applied to each individual year. The total number of realizations was therefore 5,000 (10 individual years \times 5 modifications to account for uncertainty in solar radiation \times 100 realizations).

Table 2 - Types of uncertainties modeled.

<i>Variable</i>	<i>Uncertainty distribution</i>
Solar radiation, long-term average	Normal (0%, 5%)
Solar radiation, yearly variability	Normal (0%, 3.9%)
Transposition model	Normal (-2%, 3%)
Power rating of modules	Normal (-3%, 3%)
Dirt and soiling	Normal (-3%, 2%)
Snow	Normal (-2%, 1.5%)
Albedo	Uniform (0.1, 0.15)
Other (modeling errors, spectral effects, etc.)	Normal (-5%, 5%)

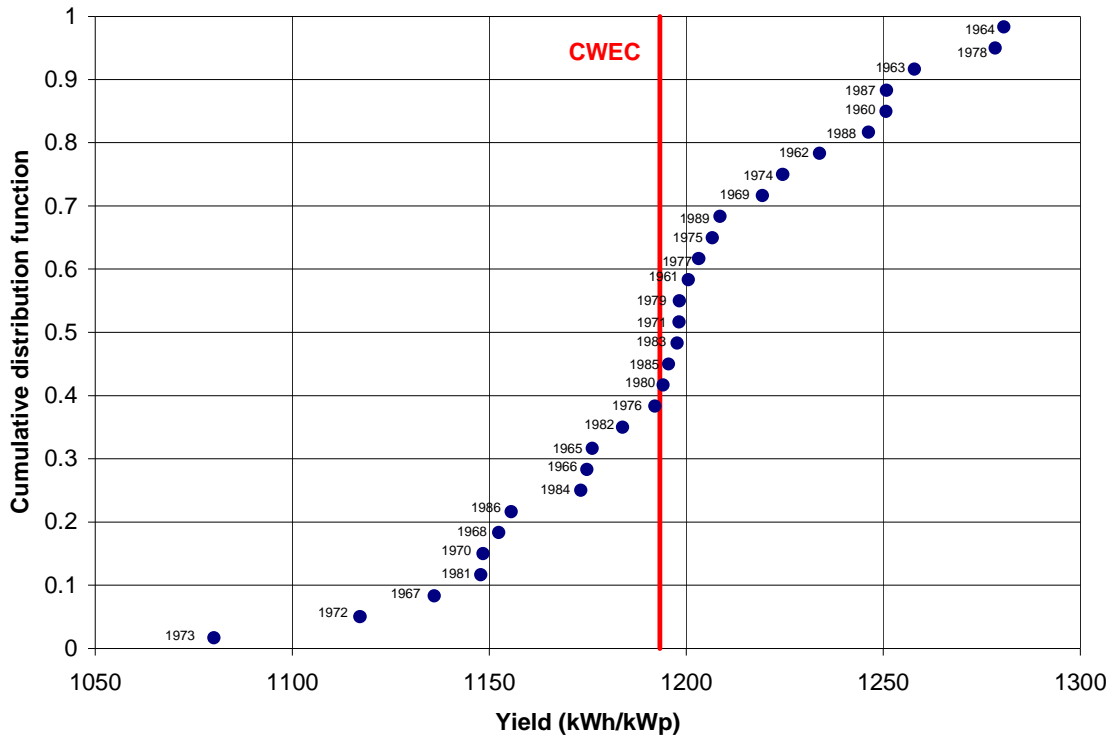


Figure 2 - CDF of yield depending on year used for simulation, for the base case.

3.4 Modeling uncertainty in lifetime average yield

The second set of statistical simulations aimed at determining the uncertainty attached to the 20-year, or lifetime average yield. For the purpose of illustration, ageing was modeled as a uniform distribution $[-0.3, -0.8]$ %/year, and availability was modeled as a uniform distribution $[0.985, 0.995]$.

Other uncertainties were modeled statistically as before, however fluctuations of the solar resource from one 20-year period to the next could be ignored as discussed in Section 2.1.2. Therefore, the statistical simulations were run only for the 'typical year', and the uncertainties related to ageing and availability were modeled by superimposing the distributions for ageing and availability to the results of the statistical simulations.

4. RESULTS AND DISCUSSION

4.1 Base case

For a typical year, with nominal parameters as per Table 1, SAM predicts an annual energy output of **717.6 MWh**, i.e. a yield of **1,193 kWh/kWp** and a system performance ratio of **0.80**.

4.2 Uncertainty in year one yield

When yearly climate variability and all uncertainties are taken into account, annual yield assumes a statistical distribution represented by the CDF shown in Figure 3. The mean yield over all realizations is **1,200 kWh/kWp**, very close to the yield predicted for the base case; the standard deviation is **105 kWh/kWp** or **8.7%** of the mean.

This figure can also be used to determine confidence levels or energy exceedance probabilities. For example, there is roughly an 80% chance that the yield of the system, for the first year, will be between 1,068 and 1,337 kWh/kWp. Similarly, the P90 level is 1,068 kWh/kWp, that is, there is only a 10% chance that the first-year yield would fall under that threshold.

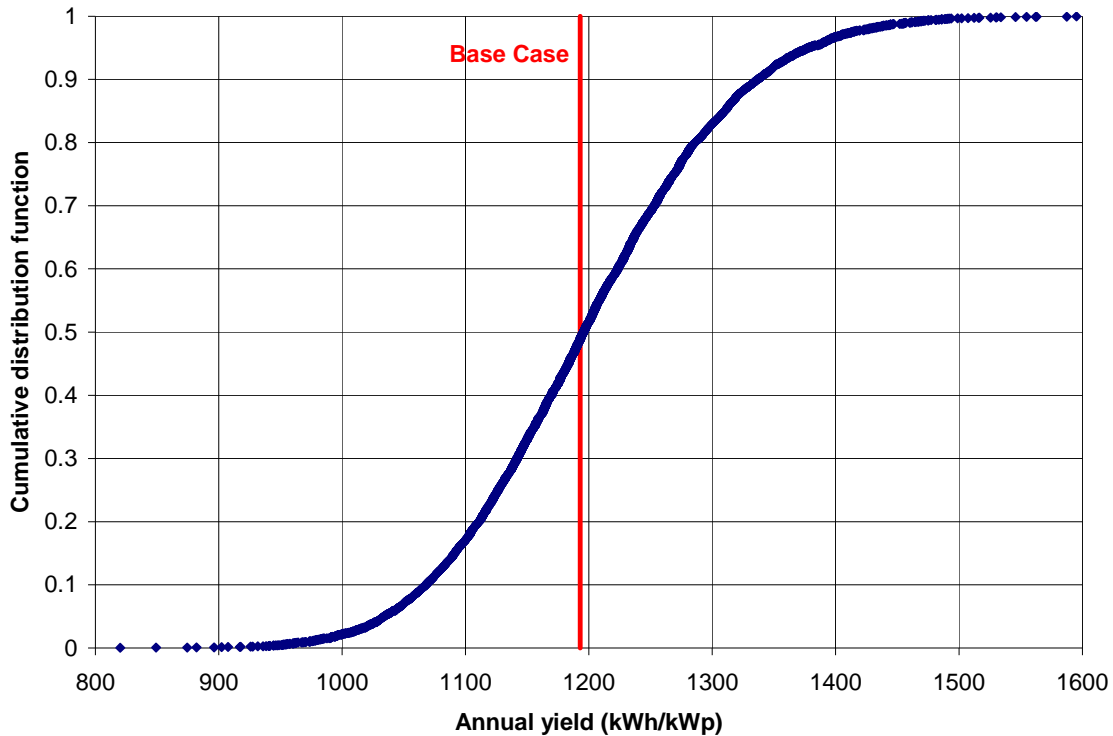


Figure 3 - CDF for year one yield considering all uncertainties and climate variability.

4.3 Uncertainty in lifetime average yield

When the long-term performance of the system is considered, including ageing and unavailability, annual yield assumes a statistical distribution represented by the CDF shown in Figure 4. The quantity plotted is the yield averaged over the expected lifetime of the system (20 years). Lifetime average annual yield is **1,119 kWh/kWp**, with a standard deviation of **89 kWh/kWp** or **7.9%** of the mean. It is worth noting that because of ageing, the average yield is significantly lower than the year 1 yield. On the other hand, the standard deviation is smaller than that of year one despite the added uncertainty associated with ageing, because of the reduced uncertainty in the solar resource when comparing a 20-year period to a single year.

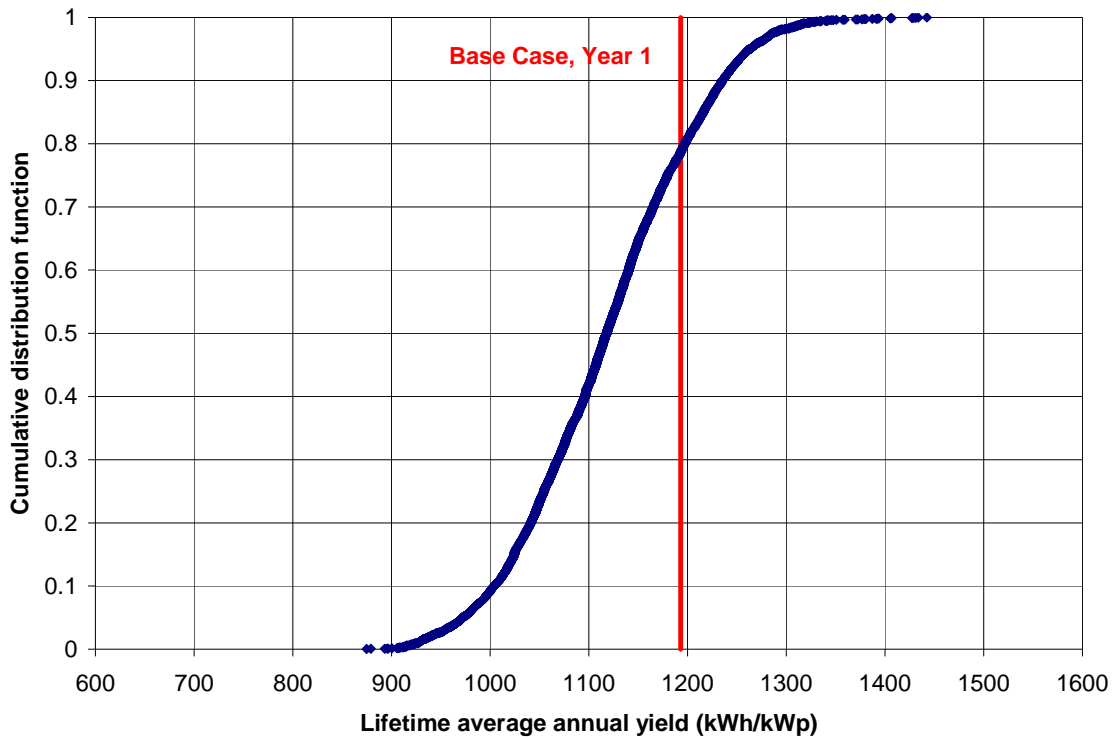


Figure 4 - CDF for lifetime average annual yield.

4.4 Discussion: Comparison with a simpler method

Since the statistical simulation approach used in this study is fairly involved, it is worth investigating whether a simpler approach to combining uncertainties could be used to yield a quick estimate of the overall uncertainty.

With the exception of the transposition model, PV output can be represented to a good approximation as the product of linear factors (output = input \times proportionality factor - offset, with the offset relatively small compared to the phenomenon itself). It is well known [30] that if a quantity X is the product of N independent variables X_1, X_2, \dots ,

X_n , as in $X = \alpha \cdot X_1 \cdot X_2 \cdot \dots \cdot X_N$ with α a constant and with uncertainties $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$, then the combined uncertainty ε_t in X can be calculated according to the 'rule of squares':

$$\frac{\varepsilon_t}{X} = \sqrt{\frac{\varepsilon_1^2}{X_1^2} + \frac{\varepsilon_2^2}{X_2^2} + \dots + \frac{\varepsilon_N^2}{X_N^2}}$$

The main uncertainties associated with year one uncertainty, for example, are:

- 3.9% for climate variability
- 5% for solar radiation
- 3% for transposition model
- 3% for power rating of modules
- 2% for dirt and soiling
- 1.5% for snow
- 5% for other errors

Using the rule of squares the combined uncertainty is:

$$\sqrt{3.9^2 + 5^2 + 3^2 + 3^2 + 2^2 + 1.5^2 + 5^2} = 9.45\%$$

which is close to the estimate (8.7%) obtained with statistical simulations, notwithstanding the fact that not all the uncertainties were normally distributed. This indicates that, should different numbers be used for the various uncertainties mentioned in this report, the 'rule of squares' can be used in lieu of full statistical simulations to obtain a quick estimate of the global uncertainty attached to the yield of the system.

The 'rule of squares' is also useful for understanding the relative importance of various uncertainties in the global uncertainty. For example, climate variability and uncertainties in solar radiation combined account for almost half the term under the square root. On the other hand, the uncertainty introduced by snow and dirt accounts for less than 7% of the same term.

5. CONCLUSION

Statistical simulations were used to estimate the uncertainty of the annual yield of a typical large PV park in Ontario. It was found that the combined uncertainty (in terms of standard deviation) is of the order of 8.7% for individual years, and of 7.9% for the average yield over a 20 year system lifetime. The numbers above are thought to be representative of good-quality installations in Ontario, and could vary significantly from location to location or from system to system. However, the methodology that was used should be widely applicable once the values of the relevant uncertainty distributions have been assessed for a given case. Moreover, as highlighted in Section 4.4, a less sophisticated approach based on simply combining uncertainties seems to yield a good estimate of overall uncertainty and could be used to obtain quick estimates.

Uncertainties could be further reduced by exploring the following avenues:

- Solar resource: current estimates from various sources vary sometimes widely. Effects of micro-climates are not well known. Reliable data sources are often interpolated over large distances, or supplemented by satellite-derived data which still suffer from serious shortcomings [3]. More work is needed to increase the reliability and spatial coverage of solar radiation estimates.
- Transposition models: As pointed out by Gueymard [8], it is difficult to know which combination of models and parameter estimates will yield the best value for irradiance in the plane of the array, since this depends both on the available data (global horizontal irradiance only, or diffuse and beam as well) and on the location. In order to support the development and testing of such models for Canada, it would be useful if a few high quality ground stations in the country could include measurements in a non-horizontal plane as part of their dataset.
- Module rating: there is still some confusion as to the tolerance on the initial power of modules. More clarity is needed to put all manufacturers on a level playing field and to reduce the uncertainty faced by developers. This clarity could be provided for instance by tightened requirements in PV module standards, such as the planned improvement to the IEC 61215 standard discussed in Section 2.2.1. In particular, these standards should deal explicitly with the issue of initial degradation of crystalline PV modules following the first few days of exposure.
- Some of the losses experienced by systems are poorly understood and difficult to estimate. This concerns in particular losses due to dirt and soiling, and losses due to snow. Current numbers are extrapolated from other jurisdictions, but their effects need to be actually measured across the variety of climates experienced in Canada.

ACKNOWLEDGEMENTS

Financial support for the report project was provided by Natural Resources Canada through the ecoENERGY Technology Initiative, which is a component of ecoACTION, the Canadian government's actions towards clean air and greenhouse gas emission reductions. We also wish to thank Anton Driesse, Yves Poissant, Lisa Dignard-Bailey and Dave Turcotte for their contributions and comments.

REFERENCES

[1] Ontario Power Authority Feed-In Tariff website, <http://fit.powerauthority.on.ca/> (August 13, 2010).

[2] Ayoub J and Dignard Bailey L. Photovoltaic Technology Status and Prospects - Canadian Annual Report 2009, CETC Number 2010-023 / 2010-04-28. Available online at http://canmetenergy-canmetenergie.nrcan-rncan.gc.ca/eng/buildings_communities/buildings/pv_buildings/publications.html?2010-023 (August 13, 2010).

- [3] Thevenard D, Driesse A, Pelland S, Turcotte D, Poissant Y. Uncertainty in long-term photovoltaic yield predictions, report # 2010-122 (RP-TEC), CanmetENERGY, Varennes Research Center, Natural Resources Canada, March 31 2010, 52 pp. Available online at: http://canmetenergie-canmetenergie.nrcan-nrcan.gc.ca/eng/renewables/standalone_pv/publications.html?2010-122
- [4] Environment Canada, National Climate Data and Information Archive, CWEEDS and CWEC datasets, http://climat.meteo.gc.ca/prods_servs/index_e.html, (July 20, 2010).
- [5] Morris RJ and Skinner WR. Requirements for a solar energy resource atlas for Canada. *Proc. 16th annual conference of the Solar Energy Society of Canada, Inc.* 1990; Halifax, Canada: 196-200.
- [6] Myers DR, Reda I, Wilcox S and Andreas A. Optical radiation measurements for photovoltaic applications: instrumentation, uncertainty and performance. *Proc SPIE's 49th annual meeting* 2004. Denver, U.S.A. Available online at <http://www.nrel.gov/docs/gen/fy04/36321.pdf> (July 20, 2010).
- [7] Sürri M, Huld T, Dunlop ED, Albuissou M, Lefèvre M, Wald L. Uncertainties in photovoltaic electricity yield prediction from fluctuation of solar radiation. *Proc. 22nd European PVSEC* 2007. Available online at http://re.jrc.ec.europa.eu/pvgis/doc/paper/2007-Milano_6DV.4.44_uncertainty.pdf (July 20, 2010).
- [8] Gueymard C. Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. *Solar Energy* 2009; **83**: 432-444.
- [9] Atmaram GH, TamizhMani G, Ventre GG. Need for uniform photovoltaic module performance testing and ratings. *Proceedings of the 33rd IEEE Photovoltaic Specialists Conference* 2008; San Diego, U.S.A.: 1-6.
- [10] Detrick A, Kimber A and Mitchell L. Performance Evaluation Standards for Photovoltaic Modules and Systems. *Conference Record of the Thirty-first IEEE Photovoltaic Specialists Conference* 2005; Coronado Springs Resort, U.S.A.: 1581-1586.
- [11] Carr AJ and Pryor TL. A comparison of the performance of different PV module types in temperate climates. *Solar Energy* 2004; **76**: 285-294.

[12] Dunlop ED. Lifetime performance of crystalline silicon PV modules. *Proceedings of the 3rd World Conference on Photovoltaic Energy Conversion 2003*; Osaka, Japan: 2927-2930.

[13] Skoczek A, Sample T and Dunlop E. The results of performance measurements of field-aged crystalline silicon photovoltaic modules. *Prog Photovolt Res Appl* 2004; **17**: 227-240.

[14] Jahn U and Nasse W. Operational performance of grid-connected PV systems on buildings in Germany. *Prog Photovolt Res Appl* 2004; **12** : 441-448.

[15] Moore LM and Post HN. Five years of operating experience at a large, utility-scale photovoltaic generating plant. *Prog Photovolt Res Appl* 2008; **16**: 249-259.

[16] Ueda Y, Kurokawa K, Kitamura K, Yokota M, Akanuma K, Sugihara H. Performance analysis of various system configurations on grid-connected residential PV systems. *Sol Energ Mat Sol C* 2009; **93**: 945-949.

[17] Sugiura T, Yamada T, Nakamura H, Umeya M, Sakuta K, Kurokawa K. Measurements, analyses and evaluation of residential PV systems by Japanese monitoring program. *Sol Energ Mat Sol C* 2003; **75**: 767-779.

[18] Becker G, Schiebelsberger B, Weber W. An approach to the impact of snow on the yield of grid-connected PV systems. *Proc. European PVSEC 2006*; Dresden, Germany. Preprint available from <http://www.sev-bayern.de/content/snow.pdf> (July 20, 2010).

[19] Wilk H, Szeless A, Beck A, Meier H, Heikkilä M, Nyman C. Field Testing and Optimization of Photovoltaic Solar Power Plant Equipment, Progress Report 1994. *12th European Photovoltaic Solar Energy Conference 1994*; Amsterdam, Netherlands.

[20] IEA PVPS Task 7. Reliability Study of Grid Connected PV Systems - Field Experience and Recommended Design Practice. Report IEA-PVPS T7-08: 2002. Available online at: http://www.iea-pvps.org/products/download/rep7_08.pdf (July 20, 2010).

[21] Kimber A, Mitchell L, Nogradi S and Wenger H. The effect of soiling on large grid-connected photovoltaic systems in California and the Southwest region of the United States. *Proc. 4th IEEE World Conference on Photovoltaic Energy Conversion 2006*; Waikoloa, Hawaii, USA.

[22] Kimber A. The effect of soiling on photovoltaic systems located in arid climates. *Proceedings 22nd European Photovoltaic Solar Energy Conference 2007*; Milan, Italy.

[23] Hammond R, Srinivasan D, Harris A, Whitfield K and Wholgemuth J. Effects of soiling on PV module and radiometer performance. *Proc. 26th IEEE PVSC 1997*; Anaheim, CA, U.S.A.

[24] King DL, Boysen WE and Kratochvil JA. Analysis of Factors Influencing the Annual Energy Production of Photovoltaic Systems. *Proc. 29th IEEE PVSC 2002*; New Orleans, U.S.A.

[25] *CE Code. C22.1-09 Canadian Electrical Code, Part 1, 21st Edition*. Canadian Standards Association : Toronto, Canada, 2009.

[26] teNyenhuis EG and Girgis RS. Measured Variability of Performance Parameters of Power & Distribution Transformers. *2005/2006 IEEE PES Transmission and Distribution Conference and Exhibition 2006*; 523 – 528.

[27] PVSYST V4.2. Study of Photovoltaic Systems. Available online at: <http://www.pvsyst.com/5.2/index.php> (July 20, 2010).

[28] National Renewable Energy Laboratory (NREL), Solar Advisor Model version 2010.4.12. Available online at <https://www.nrel.gov/analysis/sam/> (July 20, 2010).

[29] Gentle JE. *Random number generation and Monte Carlo methods, 2nd edition*. Springer-Verlag: New York, U.S.A, 2003.

[30] Crandall KC and Seabloom RW. *Engineering fundamentals in measurements, probability, statistics, and dimensions*. McGraw-Hill: New York, U.S.A., 1970.

[31] California Energy Commission. A guide to photovoltaic system design and installation. 2001. Available online at http://www.energy.ca.gov/reports/2001-09-04_500-01-020.pdf (July 20, 2010).

APPENDIX

The shape of the CDF curves generated in a statistical simulation, and the resulting average and standard deviation, depend somewhat on the number of realizations used. Intuitively one can understand that if too few realizations are used, then they will not statistically cover the whole range of possibilities for the various distributions considered. On the other hand the use of an extremely large number of realizations will give a better representation of the statistics of the problem, but at the expense of run time.

The variation of statistics as a function of the number of realizations was studied in the case of the statistical SAM runs used in this report (however to save time the uncertainty attached to solar radiation was not considered). The study is based on the use of *batches*, as explained in [25] (p. 237).

When N realizations are used, a CDF of the yield can be plotted and the statistics of the yield distribution can be calculated. Let us call \bar{Y}_N and $\sigma_{Y,N}$ the mean and standard deviation of the yield thus calculated with N realizations. If a different set of N realizations is run, the calculation of \bar{Y}_N and $\sigma_{Y,N}$ will lead to slightly different results, because the parameters used for the runs will be different. If one runs enough sets of N realizations, one obtains statistical distributions of \bar{Y}_N and $\sigma_{Y,N}$.

The distributions themselves depend on the number N of realizations (intuitively, if N is larger the distributions will be narrower). To characterize these distributions, a simulation with 6,000 realizations was run. Then the results were divided into batches of $N = 10, 20, 50, 100,$ and 200 realizations; for example there are 600 batches of 10 realizations, 300 batches of 20 realizations, etc., and 30 batches of 200 realizations. For each set of batches corresponding to a value of N , the standard deviation of \bar{Y}_N and of $\sigma_{Y,N}$ can be calculated, and the results plotted against N in a log-log graph as shown in Figure 5 **Error! Reference source not found.**

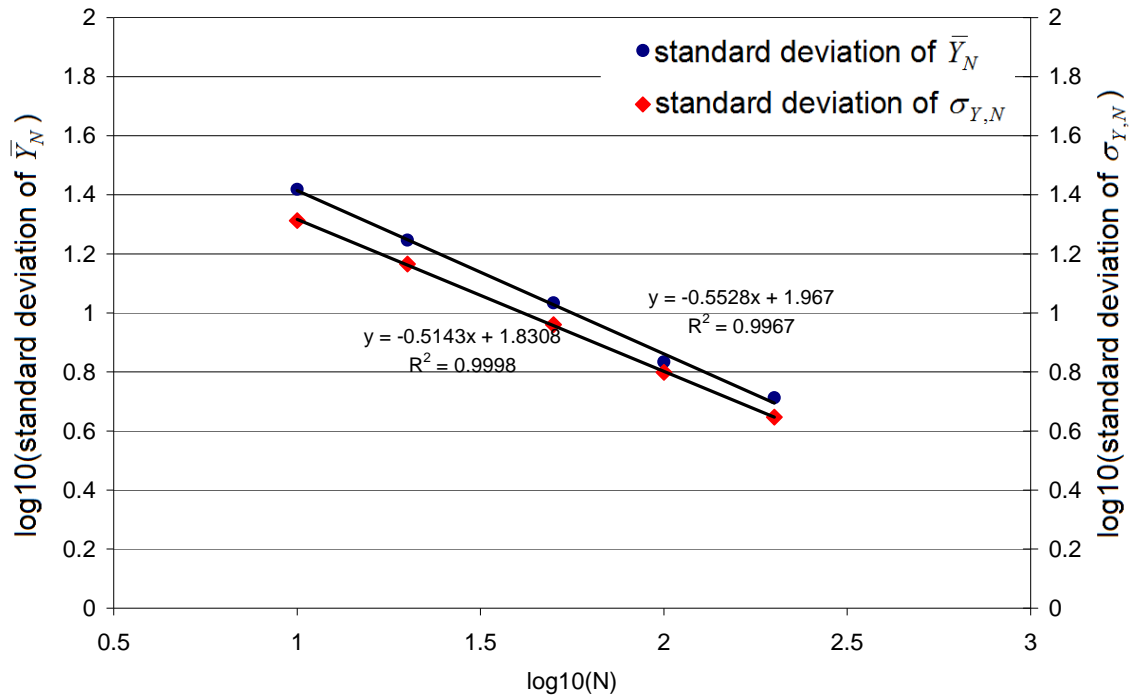


Figure 5 - Standard deviation of \bar{Y}_N and $\sigma_{Y,N}$ as a function of the number of realizations N .

The figure indicates that the standard deviation of \bar{Y}_N is inversely proportional to $N^{1/2}$. This does not come as a great surprise as the uncertainty on the mean of a sample of p values varies as $1/\sqrt{p}$. As for the standard deviation of $\sigma_{Y,N}$, its behaviour is slightly less obvious but it seems to be also inversely proportional to $N^{1/2}$.

With $N = 100$ as chosen for the study, the estimated standard deviation of \bar{Y}_N is 6.8 kWh/kWp; and that of $\sigma_{Y,N}$ is 6.3 kWh/kWp. Those numbers are small compared to the average value of the yield (~1200 kWh/kWp) and to its typical standard deviation (~93 kWh/kWp), so the number of realizations seems adequate.