MASTER

SOLAR INDEX PREDICTION METHODOLOGY PRELIMINARY ANALYSIS

by

Loren J. Lantz

SOLCOST Service Center
Solar Environmental Engineering Company, Inc. 950 2528
2524 E. Vine Drive
Fort Collins, Colorado 80524

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Currently the Solar Index is generated utilizing established weather wire data which is reduced on an hourly basis to estimates of ambient temperature and total horizontal radiation by computer models which have been thoroughly presented and described in a previously presented paper entitled "Solar Index Generation and Delivery". At sundown each day the average ambient temperature for the day is estimated and the total horizontal radiation is summed for the daylight hours and converted into percent of possible sunshine. These two estimates as well as yesterday's percent of sunshine are transformed into the Solar Index for the particular day for a particular city. It is desirable to accomplish this task in mid-afternoon for each location and then to disseminate the predicted daily Index so that it can be included in the early evening news broadcasts. Currently the software which sums the total daily radiation on a horizontal from the hourly values has a capability of filling missing data. This is accomplished by utilizing a statistical approach which assumes that the weather does not change drastically from hour to hour. That is, if it is cloudy at 1:00, then it is likely to be cloudy in a similar manner at 2:00. This procedure for filling radiation data has worked well, the ambient temperature on the other hand is handled in a much more continuous manner. The methodology used for calculating the average daily ambient temperature is to take the average

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

of the minimum and the maximum observed values throughout the day. Even if weather data sets of 20 observed values throughout the day are available, similar results are achieved when averaging the 20 values versus taking the average of the maximum and minimum. However, for certain locations there are systematic weather patterns which occur at sundown. For example: late afternoon cloudiness. The filling technique described could not account for this particular type of systematic behavior for each location. Consequently, it is desirable to develop a methodology which will minimize the mean error and the variance of the individual errors over time.

Proposed Methodology

By utilizing the existing software, estimates are available for expected average ambient temperature for the day and percent sunshine and Solar Index. Then at the end of the day, after sundown in the evening, updated values of these parameters and final results become available. Following is a description of several methodologies which could be utilized to minimize the error. Data is being collected at the current time for various locations throughout the country so that these particular errors can be defined and then adjustments can take place. The next phase will be to define an error minimization methodology that will indeed produce the best results, taking into account the cost of implementation. If one particular methodology produces slightly better results, but costs ten times as much to implement, it is not likely that it will be implemented in the end result.

In an effort to be more specific, we have predictions for each day. They can be expressed respectively in vector form as percent sun, average ambient temperature and the Solar Index;

$$[\%S_p, \bar{T}_p, SI_p]$$

and at the end of the day, the resultant values are obtained and illustrated in vector form as well as the predictions

$$[\%S, \bar{T}, SI]$$

This can be extended in a vector sense to represent a daily series such as

$$[\%S_p, \bar{T}_p, SI_p]$$

and

$$[\%S, \bar{T}, SI]$$

for example:

$$\%S = \{\%S_1, \%S_2, \%S_2, \dots, \%S_k\}$$

and etc., for the other vectors.

Now the Solar Index prediction on an average one day has an unknown prediction error (ξ) associated with it.

$$SI_k = SI_{p,k} + \xi$$

This error can possibly be biased and its statistical properties may be non-stationary (vary with time).

The desired properties of the prediction scheme are to force the mean error or expected value to zero; i.e.,

$$E[\xi] = 0$$

and to minimize the variance of the expected error.

$$\min E[\xi - E[\xi]]^2$$

Forcing the mean error to zero is a simple process and can be accomplished by a weighted feedback loop. The feedback function can be allowed to change with time if the process is nonstationary.

If the process is stationary, the estimate of the mean error $\hat{E}[\xi]$ can be calculated recursively.

$$\hat{E}[\xi]_k = \frac{k-1}{k} \hat{E}[\xi]_{k-1} + \frac{\hat{\xi}_k}{k}$$

where

$$\hat{\xi}_k = SI_k - SI_{p,k}$$

If the process is nonstationary, a moving average technique could be employed, such as

$$\hat{E}[\xi]_k = \frac{1}{n} \sum_{i=k-n}^k \xi_i$$

which could also be accomplished in a recursive manner. Another avenue which will result in a lessening of the error variance as compared to the above method is to calculate, by least squares, the expected error as a function of the inputs when calculating the predicted Solar Index.

$$\hat{E}[\xi]_k = \underline{M} \underline{x}_k$$

where $\underline{M}$ is a vector of coefficients determined by least squares and $\underline{x}$ is a

vector of functions of the inputs. One possibility for $\underline{x}$ could be

$$\underline{x}_k = \{\bar{T}_{p,k}, \bar{T}_{p,k}^2, \%S_{p,k}, SI_{p,k}, SI_{p,k}^2, \dots\}$$

As above, it will have to be determined whether or not this process is stationary or not and to compensate accordingly.

One additional procedure to minimize the variance of the error would be to implement a Kalman filter, which is an optimized procedure for accomplishing this task. The filter, however, may offer little benefit as compared to the least squares procedure described above. The Kalman filter is also expensive to implement and can be very tricky (easily becomes numerically unstable).

Implementation

After it is determined which methodology will produce the most economical and desirable results, as presented in the methodology descriptions above, it will come time to implement this particular methodology. This will require additional software which will be executed in mid-afternoon, which will automatically create the filtered Solar Index prediction which will be delivered to the news media through existing services, and software which will keep track of existing errors and parameter updates for the individual sites of interest. Enough data will be available by the middle of May to begin the determination phase as to which methodology will produce the best results. That determination and further necessary definition of software will take approximately 30 days to accomplish and then the predicted methodology will be implemented and be running smoothly sometime during June. Results and recommendations will be presented in the monthly progress reports per contract requirements.

Schedule
SOLAR INDEX PREDICTION

	April	May	June
1. Gather Data	● —		
2. Determine Methodology		● — ●	
3. Implement Methodology		● — ●	
4. Start Delivery of Predicted Index			Δ