Micro-climate Solar Modeling Over Complex Terrain:

A Validation Study of ESRI Solar Analyst

By,

Nathaniel Vandal and William Hegman
May 2007

Abstract

Incoming solar radiation (insolation) is the single most important source of energy at the Earth's surface. Solar energy variation on the landscape is the primary determinant for the spatial distribution of almost all natural and biological processes, including human life. There is, therefore, tremendous importance in the ability to accurately model incoming radiation on the landscape. This paper analyzes the results of Environmental Systems Research Incorporated's (ESRI) *Solar Analyst* on complex topography that includes land forms and buildings. Little solar modeling has been conducted on this scale, and no validated results have been published to date. The analysis in this paper starts with the development of high a resolution complex DEM of the Middlebury College campus which has a topographical range of 285 feet and contains over 80 buildings. Radiation was calculated over the study area with atmospheric data available from the National Solar Radiation Database. Validation of the model was conducted using on-site solar pyronometers placed at strategic locations within the study area. In addition to the validation study, a sensitivity analysis of the user specified input parameters was conducted to help determine their effects on the accuracy of the model. Validation results revealed an average error of 11%. Of the potential sources for error identified in this study, snow events were determined to be the most significant with an average of 7% of the study days having snow events. These results are magnified by the short time interval of the study, highlighting the need for longer measurement durations in future studies. Results from the validation study helped to determine a hierarchy for model input parameter selections, which was as follows: hour interval, day interval, skysize, and calculation directions. The effects of zenith and azimuth divisions were not significant. Input parameter settings were chosen from the sensitivity analysis to achieve 1% convergence accuracy (day interval 1, hour interval 0.1, skysize 1000, calc. dirs. 48) and were used to create radiation maps for the entire study area. Radiation patterns in the area modeling results matched up with empirical interpretations from topography, with the exception of some abnormal ground shading on the south side of buildings, which requires additional validation before being determined erroneous. In conclusion, the results from this study highlight the need for enhanced DEM creation, and increased modeling of solar radiation on complex terrain.

Introduction

1.1 Importance of Solar Radiation

Solar radiation (insolation) is the single most important source of energy at the Earth's surface, providing 99.8 percent of incoming energy (Kumar, Skidmore and Knowles, 1997). Originating from nuclear fusion in the core of the sun, radiation waves travel over 150,000,000 km before reaching earth at the solar constant of 1,366 watts/m^2. About one fourth of this solar energy makes its way through the atmosphere and reaches the Earth's surface. Despite atmospheric reduction, solar energy remains "the primary driver for our planet's physical and biological processes" (Fu and Rich, 1999). While the energy at the outer atmosphere remains relatively constant, solar energy patterns on the landscape are incredibly varied. This variation plays a significant role in the establishment of the spatial distribution of many natural processes. Local gradients in insolation have direct and indirect affects on air and soil heating, energy and water balances, and primary production (Dubaya and Rich, 1995). These processes are essential for almost all life on earth and are almost entirely so for human life. Indeed, almost all of the energy that humans consume has either directly or indirectly come from the sun. Everything from photosynthetic creation of the food we eat to the fossil fuels we burn is a product of solar energy. Because of the importance of solar radiation, there is tremendous value in working towards a better understanding of insulation patterns on the landscape.

1.3 Design and Implementation of Validation Study

This paper discusses the results of a validation study of *Solar Analyst's* accuracy using a complex DEM that models natural terrain and buildings as an input for topography. This paper starts by describing the development of a high resolution DEM of the Middlebury College campus which, in many respects, is an exemplary study area for high resolution complex solar modeling. Physically, the topographic variation of the Middlebury Campus is well defined with a range of over 250 feet, including both natural and man-made terrain. In addition, the campus contains over 80 buildings, creating complex shading patterns. Monthly solar variation also changes rapidly due to Middlebury's northerly latitude, further testing the rigor of the model.

As a perfect site for validation of ESRI's *Solar Analyst*, high resolution DEMs of the Middlebury College campus were created of the study area. These terrain models were then used as the input to *Solar Analyst.* Insulation was calculated over the study region with model calibrations from local atmospheric data obtained from the National Solar Radiation Database. Convergence studies were also employed to help calibrate the model for use in area solar mapping. The results of the modeling at specific sites were then compared with measured results from Hobo pyranometers (Figure 2).

Methods

2.1 Modeling Ground Topography

Topographic information was collected from two separate sources. The first was a campus-wide survey with two-foot (0.6m) elevation contours in an AutoCAD file provided by Andropogon Associates. To bring the topography up to date in areas of recent change, a series of patches with updated contours helped to supplement the master survey. These two files were combined and vertices were converted to a point shapefile. Gridding was preformed using Dynamic Graphics earthVisions (EV). EV was chosen as a grid engine because it allowed for statistical analysis of the accuracy of the grid. Erroneous points were identified in preliminary gridding, and all points with 0.15m difference error were removed. This removed approximately 200 points in a dataset containing roughly 250,000 points. Final grid results were satisfactory with an average absolute z error to 0.002m and a maximum z error to 0.22m. Because EV produced the grid in a proprietary grid format, it was exported into a more universal ASCII x,y,z format. The *ASCII to Raster* command was then used in ArgcGIS for the creation of the final terrain raster.

2.2 Adding Buildings to Topography

Building data was acquired for the Middlebury College from a variety of sources. For a majority of buildings, data was collected for the height of the eves, on prominent buildings and ridgeline elevations were also collected. This data was compiled and stored in shapefile format.

The main difficulty with creating a terrain that contained buildings was modeling a vertical wall using a raster. Rather than trying to retroactively add the buildings as features to include in the gridding process, a method of simple vector conversion and raster overlay was the most functional. This also avoided potential problems that earth type gridding engines have with square, horizontal and vertical surfaces.

The first step in the overlay process was to convert the building footprints shapefile to a raster with 0.6m resolution that would match the raster DEM. The value for each of the buildings was assigned to the roof height or, in cases where roof data was not available, the height of the eves. In order to ensure that the extent of the building's raster overlapped the topography completely, two polygons were added just outside of the vertical and horizontal extent of the topography.

In order to effectively place the buildings onto the digital topography, some simple raster math had to be preformed. The first step was to 'reclass' the building's layer so that any values outside of the buildings contained a value of zero. This layer was reclassified to create a binary layer where the buildings were assigned a value of zero and all other cells were converted to a value one. Using the raster calculator, the binary layer was used as a multiplier for the topography. This effectively flattened the terrain, and built a zero-foot high foundation for the buildings to be built on. Once added, the result was a terrain that included flat-topped buildings for analysis.

While flat-topped buildings were satisfactory for modeling ground shadowing, they had no value as an inventorying tool for calculating rooftop insulation. Flat roofs tend to show equal insulation values across an entire roof, which are often lower than what might show up on a south facing pitched roof. In addition, flat roofs cannot model any shading on the north side of a pitched roof. After some basic test modeling, it was determined that east-west oriented buildings required additional ridgeline modeling.

A first attempt at modeling ridgelines proved to have some success on square buildings with one ridgeline, but little success on more complicated structures. Conversion of the building features was conducted using the Triangulated Irregular Network (TIN) toolset in 3D analyst. In order to constrict the TIN modeler to the specified structure, buildings were used as hard clip polygons, while the roofline polylines were entered as hard lines. Results were satisfactory for simple, rectangular buildings with one roofline, but unsatisfactory for complicated building geometry with more interesting rooflines. Improved results were achieved by splitting complicated buildings and rooflines into segmented shapes. While this worked well for moderately complicated buildings, it became very time consuming and difficult to do with large, highly complex buildings. (Figures 1a and 1b).

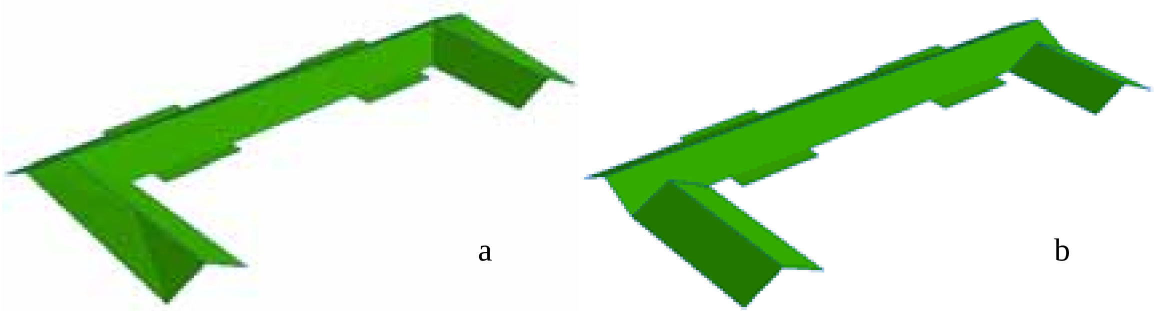


Figure 1a TIN network for Forrest Hall using one polygon and one ridgeline.
Figure 1b Triangulation using split polygons and rooflines.

In addition, there were errors in the TIN processor that occurred when building sections overlapped. In some cases the overlapping polygons were not treated as independent clip polygons, and the resulting TIN would represent a hybrid of the two polygons. The only solution was to split the buildings and run the TIN modeler individually for each section of building. Other modeling options were explored because segmenting the buildings and running multiple TINs was going to be extremely time consuming and did not guarantee results.

With no other way to generate a surface from the shapefiles in ArcGIS, the final method considered was to build models in a 3D environment, which could then be incorporated into the terrain. Sketchup by Google was chosen as the modeling program because of its compatibility with ArcGIS, as well as the ease of its use. As a preliminary test, the building footprint of Bicentennial Hall was exported from shapefile format into a Sketchup file. Using dimensions from the buildings plans, walls and rooftops were extruded to create an accurate 3-dimensional model.

Because the Sketchup plugin would not work with ArcGIS 9.2, the model was exported into multipatch shapefile format on an older ArcGIS 9.1 system. Once converted, the file could then be used in 9.2. While the multipatch format allowed for flawless display of the model in both ArcMap and ArcScene, limitations became evident very quickly. There were no tools within ArcGIS to generate a raster or TIN surface from multipatch files. While the feature heights and dimensions were embedded in the model, this information is hidden from the user and cannot be extracted. With no other surface modeling software available, it was determined that a combination of square extrusion and TIN creation would have to be used to model the buildings. As with previous TIN modeling, selected structures were split into simple polygons with independent rooflines. These were then processed by the TIN modeler and converted later using the *TIN to Raster* command.   Flat-topped or extruded buildings were converted to a raster environment using *Feature to Raster*. These files were combined and extruded onto the terrain using the same technique as described for flat-topped buildings.

2.3 Measuring Solar Radiation Using Silicon Pyranometers

Validation data was collected at three separate study sites using HOBO Silicon Pyranometer Smart Sensors. The HOBO Smart Sensors measure horizontal or flat plane radiation along a spectral range from 300 to 1100nm. The measurement range is from 0 to 1280 W/m^2, with an accuracy of +/- 10 W/m^2 or +/- 5%, which ever is greater in sunlight. The resolution of the sensor is 1.25 W/m^2 (HOBO). Measurements were recorded using HOBO Micro Station Data Loggers. Measurements were taken every three minutes and were averaged over the 30 minute logging interval before being recorded. This was done to reduce error on days with intermittent cloud cover.   Two separate data loggers were employed, each recording measurements from an attached pyranometer.

Measurement locations were chosen to have large variations in topographic shading and were spread across campus in an effort to measure the effects of topographic shading throughout the entire study area (Figure 2).

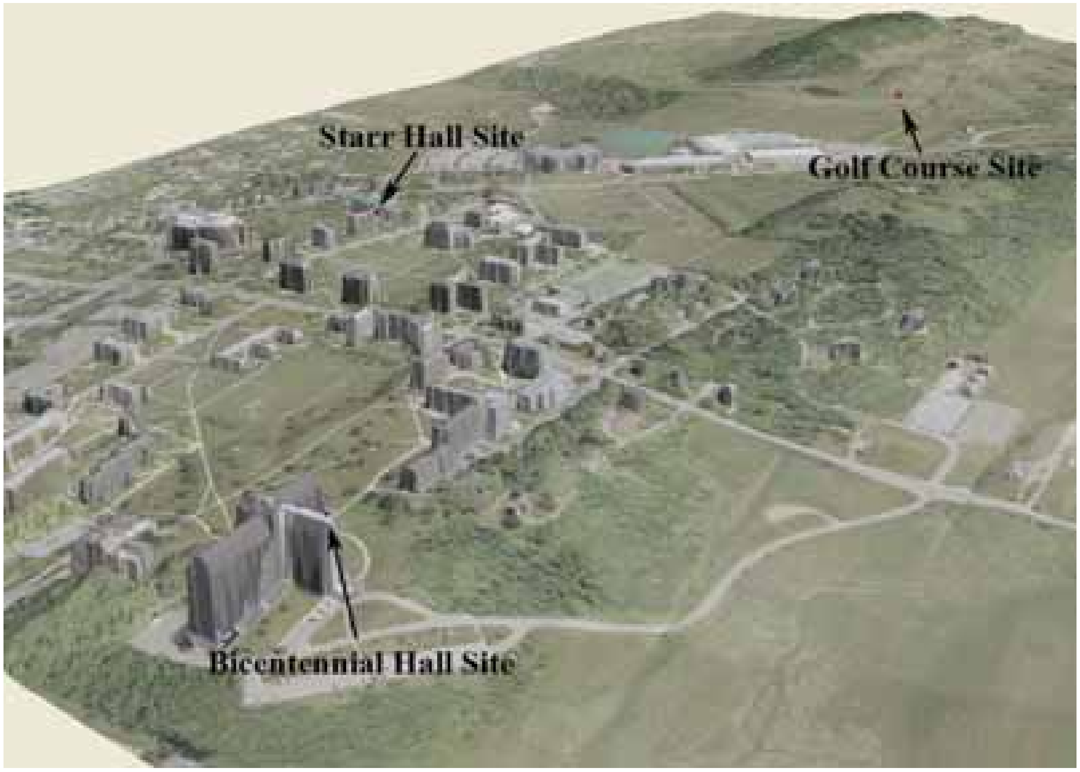


Figure 2. Location of the Bicentennial Hall, Starr Hall, and Golf Course Study sites within the study area.

In addition, sites were selected that reflect potential uses of the model. One of the pyranometers was deployed on the roof of Bicentennial Hall for the entire duration of the study to serve as a control point of measurement. Located at the highest point in the study area, the control pyranometer was chosen to minimize topographic shading. Because shading did not play a role in either the measured or modeled results, this helped to facilitate validation comparison and calibration of non-topographical inputs. Because of the site's high potential for solar energy, an additional consideration for the selection of this control location was to collect measured data for potential site development.

The second pyranometer/data logger pair was deployed for shorter time intervals at two other locations within the study area. These locations were also chosen to reflect some degree of utility. The golf course was chosen because it was a ground-based location with minimal shading effects from surrounding buildings. This helped to provide comparison results for a modeled environment with complex local topography, but with no buildings. As a ground-based measurement, the golf course site was also chosen to reflect the use of *Solar Analyst* as an agricultural tool. Starr Hall was chosen as the third study location in part because of its contrast with the golf course site. Starr Hall is located near the center of campus, providing a site with significant shading effects from the surrounding buildings, allowing for further contrast with the Bicentennial Hall site.

2.4 Modeling Point Based Radiation

The *Points Solar Radiation* tool in *Solar Analyst* was used to model global radiation for each of the validation sites. *Points Radiation* was chosen over *Area Solar Radiation* because it limited computation to individual points of interest, rather than the entire study area. *Points Radiation* also allowed for an offset elevation to be assigned to each point which was particularly helpful for modeling the golf course and Bicentennial Hall sites, where the pyranometers were mounted one meter above the terrain surface. The focused calculation of *Points Solar Radiation* drastically reduced computational time, which in turn, allowed for multiple iterations of the model at each study site. The pyranometer study sites were digitized in the model as point shapefiles which were used as the input for *Points Radiation*. Additionally, the study dates from each of the pyranometer datasets were used as the input time interval for direct comparison of the modeled and measured results.

The *Batch* command was used to systematically change each of the input parameters in an effort to determine the sensitivity of the user specified inputs. The parameters that were altered were sky resolution, day interval, hour interval, calculation directions, azimuth divisions, and zenith divisions. Other input variables were changed from the default setting, but these were held constant for each of the iterations within a batch. Because the data were to be compared with pyranometer data which collected radiation as a horizontal or flat plate sensor, the slope and aspect input were set to FLAT_SURFACE. Additionally, the STANDARD_OVERCAST_SKY model was chosen for the Diffuse Model type. All other input variables were left in the default configuration, with the exception of the Diffuse Proportion.

Diffuse proportion was calculated for each time interval from the National Solar Radiation Database (NSRDB) results for Burlington, VT. The NSRDB contains 30 years (1961-1990) of solar radiation data from 237 National Weather Service meteorological sites. NSRBD data was used because it is a readily available national data source which can be used to infer diffuse proportion at virtually any location in the United States. Monthly totals for the average daily total solar radiation for the global horizontal (Wh/m^2) were divided by the diffuse horizontal radiation (Wh/m^2) to find the diffuse proportion for each time interval. In the case that the monthly totals did not match up with the time interval, the monthly totals were averaged for all months that intersected the modeled time interval. Once calculated, the diffuse proportion remained the same for each Batch.

Because sky size was the only parameter which affected the creation of all three hemispherical maps, (i.e., viewshed, sunmap and skymap,) changes in the other 6 parameters were repeated for the complete range of sky sizes, 200 – 4000 at resolution intervals of 200 (i.e. 200, 400, 600, etc.) The day intervals used in the sensitivity analysis were 14, 10, 6, 2 and 1. During sensitivity modeling of other variables, the day interval was set to the highest resolution of 1 day. The hour intervals used were 0.8, 0.6, 0.4, 0.2, and 0.1. The hour interval was also set for its highest resolution, 0.1 hours, during sensitivity studies of the other parameters. The calculation directions modeled were 16, 32, 48, 64, and 80. Sixty-four calculation directions, twice what is "recommended for complex topography," was used as the standard for testing other variables (ESRI Solar Analyst Help). The zenith and azimuth divisions tested were 8, 16, and 32. Sixteen divisions, twice the default, were used as the standard for the number of zenith and azimuth divisions while changing other variables. For each sky resolution modeled, there were a total of twenty-two outputs, each in the form of a new shapefile (Figure 3). With twenty different sky sizes, there were 420 iterations of *Points Solar Radiation* for each study site, adding up to a total of 1260 total calculations for the study area.

Table 1. Example of batch calculations for skysize 200.

File Name	Sky Size	Day Interval	Hour Interval	Calculation Dirrections	Zenith Devisions	Aziumth Devisions
t1.0.shp	200	14	0.1	64	16	16
t1.1.shp	200	10	0.1	64	16	16
t1.2.shp	200	6	0.1	64	16	16
t1.3.shp	200	2	0.1	64	16	16
t1.4.shp	200	1	0.1	64	16	16
t1.5.shp	200	1	0.8	64	16	16
t1.6.shp	200	1	0.6	64	16	16
t1.7.shp	200	1	0.4	64	16	16
t1.8.shp	200	1	0.2	64	16	16
t1.9.shp	200	1	0.1	64	16	16
t2.0.shp	200	1	0.1	16	16	16
t2.1.shp	200	1	0.1	32	16	16
t2.2.shp	200	1	0.1	48	16	16
t2.3.shp	200	1	0.1	64	16	16
t2.4.shp	200	1	0.1	80	16	16
t2.5.shp	200	1	0.1	64	8	16
t2.6.shp	200	1	0.1	64	16	16
t2.7.shp	200	1	0.1	64	32	16
t2.8.shp	200	1	0.1	64	16	8
t2.9.shp	200	1	0.1	64	16	16
t3.0shp	200	1	0.1	64	16	32

Batch files were created in Microsoft Excel in tabular form, which were then pasted into the batch display window in ArcCatalog. This allowed for easy manipulation, and automated incrementations of each of the modeled parameters. In an attempt to conduct a sensitivity study of each of the individual variables, only one parameter was changed at a time, while the other parameters were left unchanged. The sensitivity studies served two purposes. The first is that they helped to verify that the results were as accurate as possible. By running the model using a variety of input parameters, the effect of each of the inputs could be measured, ensuring that the best combination of inputs was used for the final result. In addition, convergence thresholds were modeled and helped to indicate where increased accuracy no longer had a measurable effect on the results. By determining the lowest resolution that achieved convergence accuracy, modeling time could be reduced. This would be very helpful for determining the correct modeling thresholds for the more computationally intense *Area Solar Radiation,* as well as to give insight into the model accuracy needed for other users using similar high resolution DEMs.

Once the results from the *Batch* command had been compiled, a preliminary test for the influence that each of the input parameters was preformed. The range of each of the datasets were calculated as a percent of the highest resolution result. This was done for all twenty different sky resolutions for each input parameter. The percent range in no way determined the nature of the variables' effects on the graphical results, but it helped to serve as an indicator for horizontal linearity within each parameter. This proved to be a valuable indicator in the case of calculation directions, zenith angle and azimuth angle. For example, the range of zenith and azimuth divisions proved to be less than 0.3 percent at all of the modeled sky sizes at each of the different study sites. This proved that altering the zenith and azimuth divisions had a negligible effect on calculated radiation. Additionally, the number of calculation directions proved to be inconsequential for the Bicentennial Hall and Golf Course sites. These parameters would therefore not need any further statistical exploration.

The analysis of the remaining parameters was done graphically. Separate graphs were produced for each study site which represented the input variables' day interval and hour interval vs. the calculated radiation for each skysize. Based on the preliminary results, calculation directions were graphed only for the Starr Hall site. For each of the graphs, a +/-1% error bar was added to the highest resolution (4000).

The final comparison and validation of the modeled data was performed on the high resolution modeled data for each study site.   The input parameters used were as follows: Day Interval = 1, Hour Interval = 0.1, Sky Resolution = 4000, Calculation Directions = 64, Zenith Divisions = 16 and Azimuth Divisions =16. All other parameters were the same as previously described for each of the *Batch* calculations. This was done to ensure that the accuracy of the model was based on the best possible approximation of real data.

2.5 Modeling Area Solar Radiation

Area solar radiation maps were created for the Middlebury Campus study area using *Area Solar Radiation.* This was done in an effort to compare area solar results with the measured results from the pyranometers. The Starr Hall Study interval was not mapped, because the *Area Solar Radiation* maps created used slope input from the DEM as an input to the model. Unlike the Bicentennial Hall and Golf Course Sites, the pyranometer mounted on Starr Hall was not perpendicular to the DEM surface. As a result, the *Area Solar Radiation* values would not be comparable.

Input parameters for *Area Solar Radiation* were chosen based on a 1% convergence with the highest resolution modeled result. The settings used were Day Interval = 6, Hour Interval = 0.1, Sky Resolution = 1000, Calculation Directions = 64, Zenith Divisions = 8 and Azimuth Divisions = 8. With the exception of the aforementioned slope information being calculated from the DEM, all other input parameters were the same as in the *Points Solar Radiation* model.

Results

3.1 Comparison between Calculated and Measured Radiation

The results from the *Points Solar Radiation* comparison study showed an average difference error of 11% +/- 5% between measured and calculated radiation. Results are split into three distinct time intervals; the Bicentennial Hall (BH) time interval, the Golf Course (GC) time interval, and the Starr Hall (SH) time interval. The BH pyranometer data was also compared over GC and SH time intervals as a source for control (Table 2).

Table 2. Results from the *Points Solar Radiation* comparison study.

Study Site	Study Interval	Measured Radiation (Wh/m²) +/- 5%	Calculated Radiation (Wh/m²)	Percent Error +/- 5%
Bicentennial Hall	10/16/2006 - 3/20/2007	242,182	270,678	12%
Golf Course	10/17/2006 - 12/10/2006	35,671	41,051	15%
Bicentennial Hall	10/17/2006 - 12/10/2006	33,996	41,830	23%
Starr Hall	1/9/2007 - 3/20/2007	92,590	97,392	5%
Bicentennial Hall	1/9/2007 - 3/20/2007	175,784	179,261	2%

As demonstrated by Table 2, the lowest percent errors occurred over the SH interval from January 9, 2007 – March 3, 2007. Over this range, the error for both the SH and BH sites was less than or equal to 5%. Results from November 16, 2006 – March 20, 2007 at the BH site showed the next best results with an error range from 6% to 17%. The SH interval from November 17, 2006 – December 12th, had the worst percent error, with an average error for both sites of 19%. This is also the shortest study interval in the study. Interestingly, it is also the only time interval where the calculated results do not correspond with measured results, in that the measured radiation at the GC site was higher than the BH site, the opposite of what the model suggests.

3.2 Effects of Model Parameters on Calculated Results in Addition to Convergence Thresholds

The first step in assessing the effects of the model input parameters was to calculate the range of each of the individual parameters as a percent of the highest resolution result for all sky sizes. The percent range of zenith and azimuth divisions was less than 0.3% at all of the modeled sky sizes at each of the different study sites. Additionally, the percent range for the calculation directions parameter was less than 0.03% for the Bicentennial Hall and Golf Course sites. The relationships between other modeled parameters and model calculated radiation are demonstrated in the following graphs.

3.2a Effects of Day Interval and Skysize on Calculated Solar Radiation

The effects of day interval and skysize on calculated radiation were very similar at all three study areas. In each of the study areas there is a general trend that increased day intervals corresponds to increase modeled radiation. Additionally, convergence of skysizes occurs at a lower threshold as day interval increases (Figures 3a, 3b and 3c). There is a general convergence (within the 1% error threshold) of skysize results above 1000 for all day intervals, at each of the three sites. In addition, for day intervals less than 6 days results are within the 1% threshold of the 1 day interval, highest resolution result. These results support the conclusions that there is minimal difference in modeled results at day intervals less than 6 once the sky size exceeds 1000.

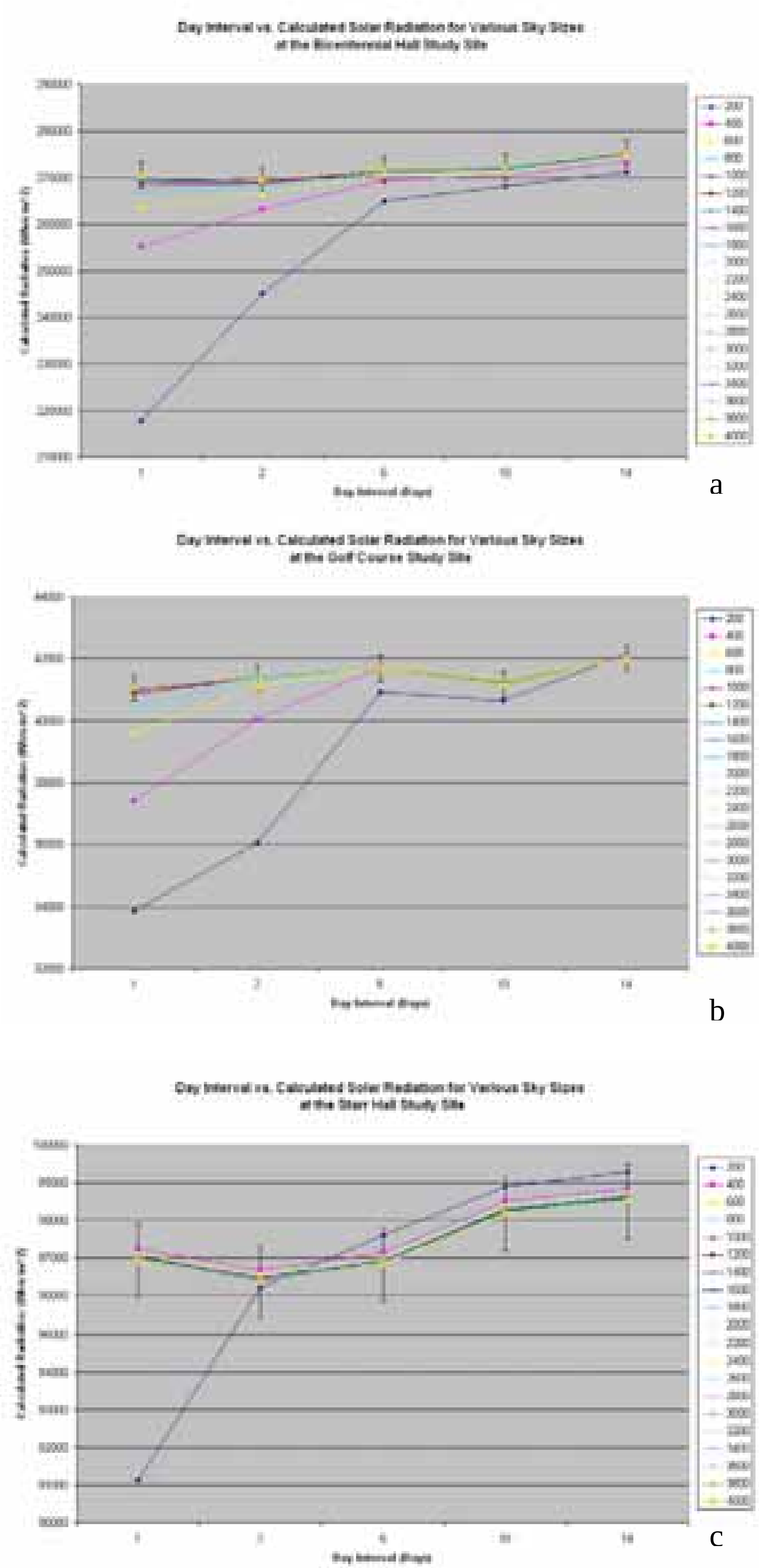
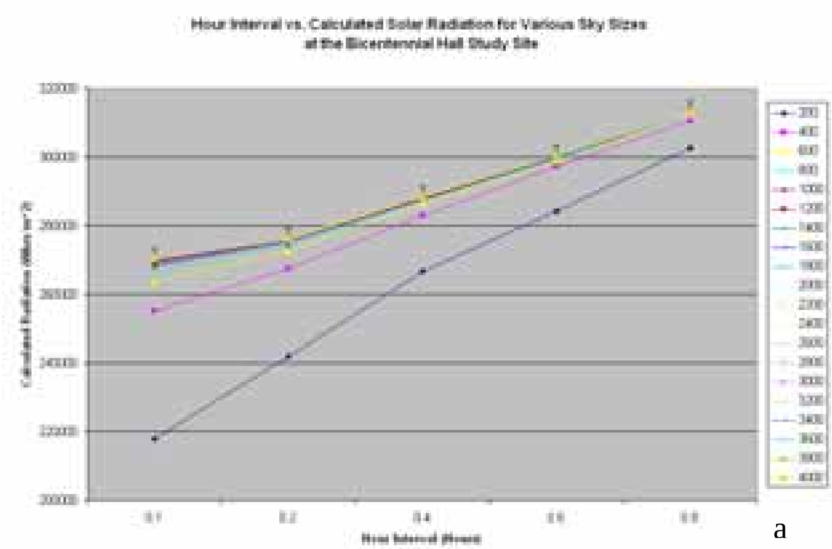


Figure 3. Graphs of the effects of Day Interval and Skysize on Calculated Radiation at the a) BH, b) GC and c) SH Study Sites. 1% error bars are provided.

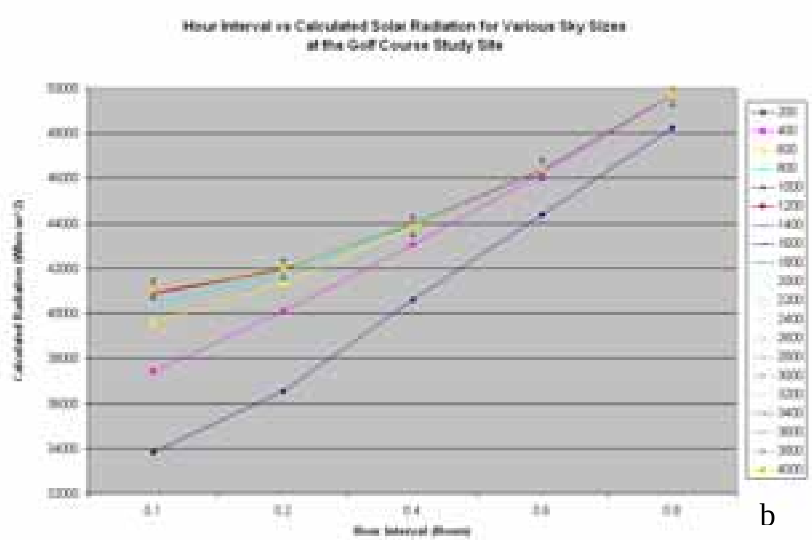
What is interesting is that the SH site shows skysize convergences at 400, where results for each day interval are within the 1% error threshold. What is also interesting is that for the skysize 200 results for day intervals greater than 6, calculated radiation is greater than all other sky sizes. This difference is most likely attributed to the greater topographical shading at the SH site. At low model resolutions, viewsheds and sunmaps are generalized and the effects from model shading are decreased. Overall, the results from all three locations suggest that at larger sky sizes, the effects of the day interval are less pronounced on calculated data. In order to achieve an average difference error close to 1% at all three sites, a minimal sky resolution of 1000 would be required with a maximum day interval of 6.

3.2b Effects of Hour Interval and Skysize on Calculated Solar Radiation

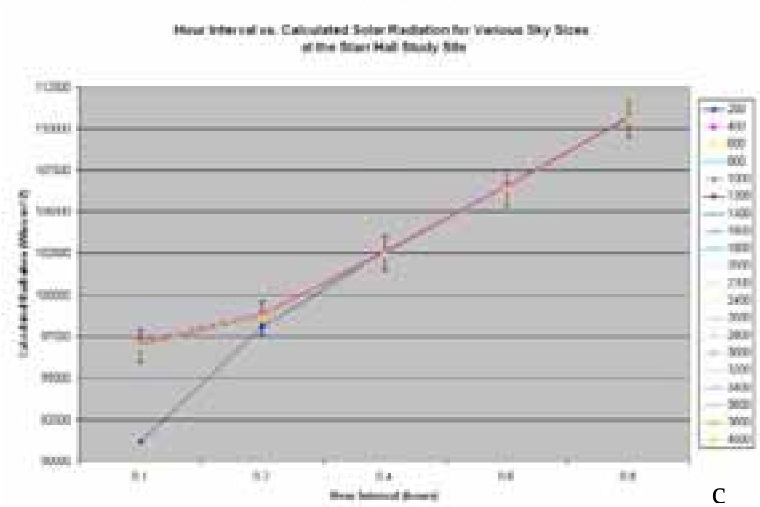
The effects of hour interval and skysize on calculated radiation were also very similar at all three study areas. As with day interval, decreased resolution, (larger hour intervals,) increased the modeled result. In the case of hour interval, the effects of this relationship are pronounced at all sky resolutions, suggesting that the choice of the hour interval is always pertinent (Figures 4a, 4b and 4c). Additionally, there is a convergence of results for all sky resolutions within the 1% error threshold at sky resolutions greater than 1000.



a



b



c

Figure 4. Graphs of the effects of Hour Interval and Skysize on Calculated Radiation at the a) BH, b) GC and c) SH Study Sites. 1% error bars are provided.

At all three sites, no other hour intervals are within the 1% error threshold of the highest resolution (0.1 hour) result. This suggests that in order to achieve 99% accuracy, the choice of hour interval should always be confined to 0.1 hours. At each of the three sites, however the results from hour interval 0.2 are very close to the 1% error threshold. Overall, the results from all three study areas suggest that in order to achieve results with an average difference error less than 1% at all three sites, a minimum sky resolution of 1000 would be required with a maximum hour interval of 0.1.

3.2c Effects of Calculation Directions and Skysize on Calculated Solar Radiation

The effects of calculation directions and skysize on calculated radiation were only pronounced at the SH site. Preliminary analysis (described in Methods) revealed that that for the BH and GC sites, all results were within 0.03 percent of the highest resolution result. At the SH site, however, the choice of calculation directions played a significant role in the accuracy of the model (Figure 5).

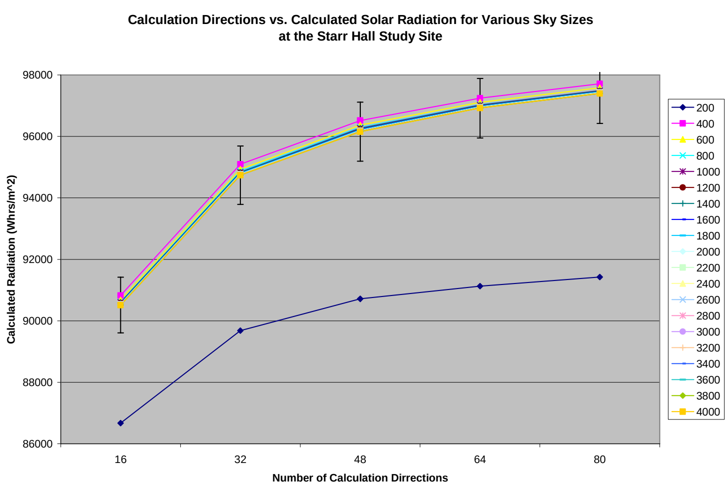


Figure 5. Graph of the effects of Calculation Directions and Skysize on calculated radiation at the SH site. 1% error bars are provided.

The graphical results from Figure 5 show a non-linear trend between increased calculation directions and change in modeled radiation. At all sky resolutions, there are significant changes in modeled resolution from 16 to 48 calculation directions. The changes above 48 calculation directions become much less dramatic, with results from 48 and 64 directions, very close to and within the 1% error threshold of 80 calculation directions. Additionally, there is a tight convergence of results at sky sizes above 400, that all fall well within 1% of the highest resolution, 4000 skysize.

It conclusion, the number of calculation directions played a significant role in the modeled results at the Starr Hall Site. In order to achieve a 1% difference error at the Starr Hall Site, a minimum sky resolution of 400 is required with a minimum number of 64 calculation directions.

3.3 Results from Area Solar Analyst

The radiation maps created for both the BH and the GC study interval showed results that were consistent with empirical interpretations from terrain. In both maps, as expected, radiation values on Bicentennial hall were generally high, especially on the south sloped portions of the roofs. Additional shading and highlighting occurred in topographically anticipated patterns throughout the study area. The radiation maps were also draped over the raster DEM in ArcScene. This helped to further illustrate the effects of topography on modeled radiation are more visible (Figure 6).

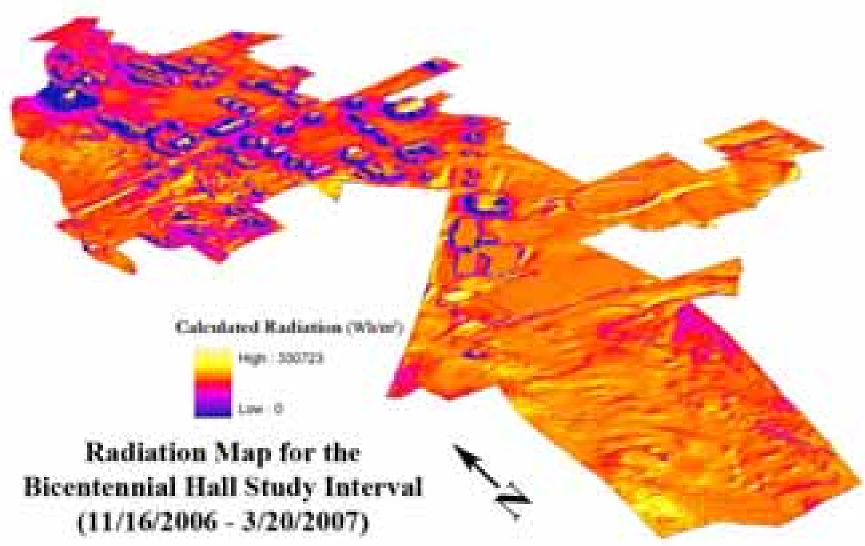


Figure 6. Radiation map for the BH study interval (11/16/2006 – 3/20/2007) draped over the DEM surface. Note the topographic shading on the north side of buildings, and the highlights on south facing surfaces.

One area where topographic shading results appear to be wrong is at the base of the south side of buildings. These effects are very pronounced in Figure 6 where many of the entire south walls of building appear shaded. A closer look in both 3-dimensional and planer views reveals that this shading is mostly limited to ground surfaces (Figure 7a and 7b). These errors are more pronounced in Figure 6 because the display resolution is twice as large as the cell size. This creates generalizing of pixels along the border of the building and the ground surface.

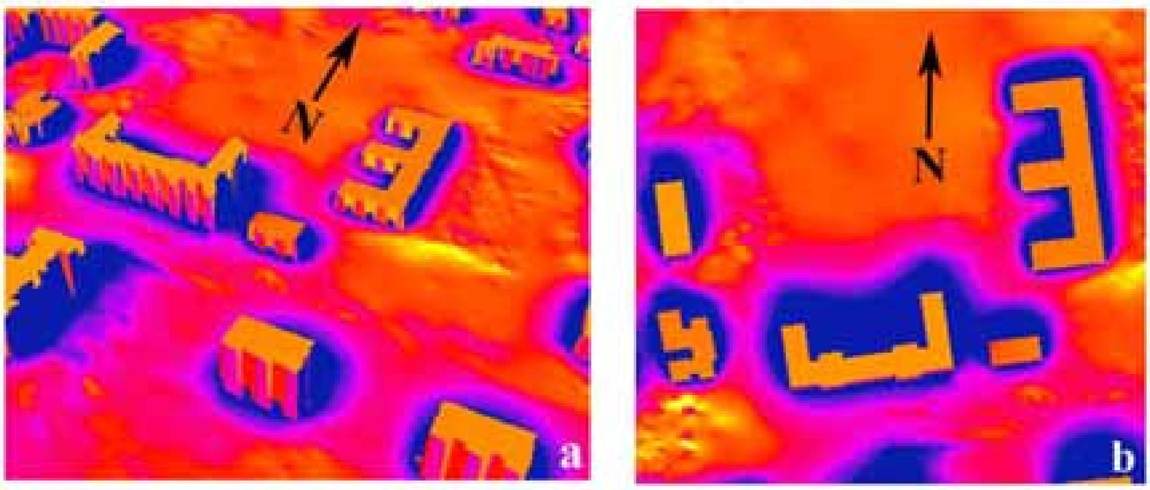


Figure 7a. 3-dimensional view of calculated radiation on buildings in central campus. Figure 7b. Planar view of calculated radiation at the same site. Note the dark shading in both images on the ground surfaces and the slight shading on the south facing walls. shows building south faces. (Color definitions are the same as Figure 6).

One possible explanation of the shading results in Figure 7 is modeling error within *Solar Analyst*. The lesser radiation on building walls could be explained by the effects of incidence angle on the near vertical building walls in the campus DEM. Because these walls are close to 90 degrees, incidence angles of radiation are much lower during the middle of the day when the sun is highest in the sky and solar intensity is largest. While this explains the results for building walls, it does not explain the dark ground shadowing. Without any explanation for these errors, the only possible reason is model error. Without pyranometer validation for these areas however, this finding is not conclusive.

In addition to the 3-dimensional modeling of the radiation maps, a classification of the results was preformed to help demonstrate similarities within the radiation maps. Radiation values for each cell were grouped by their topographic definition, along with manual matching from the histogram (Figure 8).

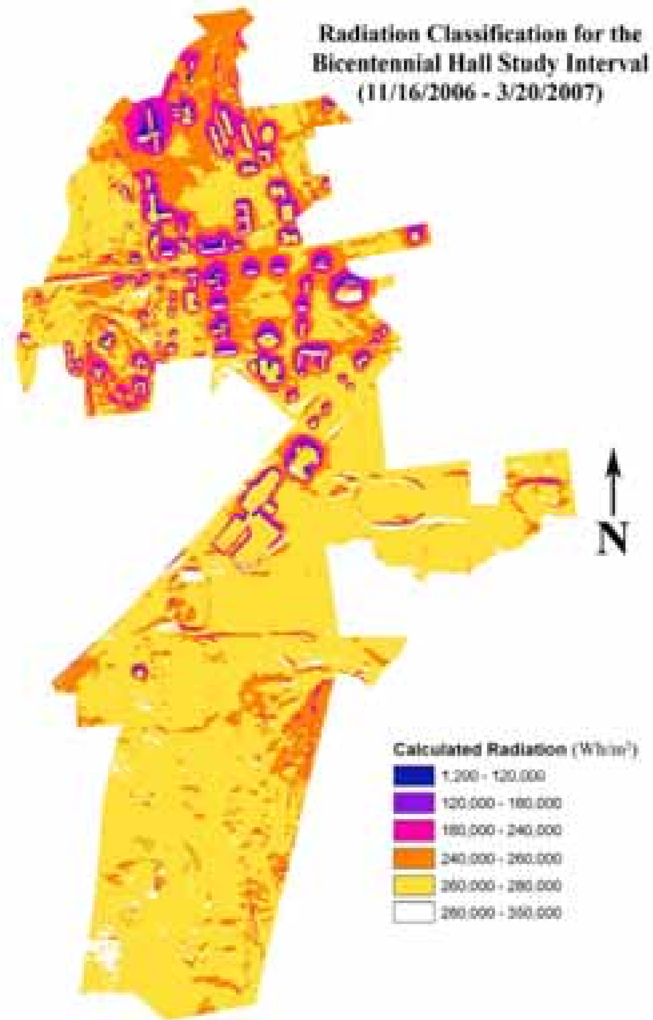


Figure 8. Pixel radiation value classifications of the BH Study Interval (10/16/2006 – 3/20/2007).

As demonstrated in Figure 8, rooftop radiation values that ranged from 280,000 to 350,000 Wh/m^2. These values were also ~15% higher than the measured radiation at the BH site of 241,182 Wh/m^2. Ground topography located away from buildings showed higher radiation values ranging from 260,000 to 280,000 Wh/m^2. Values for ground topography located between buildings ranged from 240,000 to 260,000 Wh/m^2. In addition, areas of topographic similarity are interspersed with small pockets of radiation that deviated from the average because of their slope and aspect. This is especially visible near the GC site in the southern portion of the study area. In addition, complex topographical areas near buildings show expected radiation patterns with high radiation values on the south sides of buildings and lower values to the north where the effects of shading were more pronounced.

Discussion

4.1 Validation Results from the Comparison Study

On average, the results from the point radiation validation study proved that *Solar Analyst* is an effective tool for modeling incoming solar radiation. With an average percent error at all three sites of 6% to 17%, the modeled results are very close to the standard of 5% error. While the average error is a telling indicator of the overall accuracy of *Solar Analyst,* the results from each of the individual sites is perhaps more revealing. The topographic differences between the sites, in addition to the different time intervals at each of the sites, were an excellent test of *Solar Analyst* over a wide variety of conditions.

A brief comparison of all results shows that the GC time interval produced the largest difference errors. From November 17th, 2006 through December 10th, 2006 the averaged modeled radiation was for the GC and BH sites were 19% higher than measured. What is also interesting is that the GC pyranometer measured over 4% more radiation than the BH pyranometer. This result is puzzling because the BH Site is the highest elevation in the study area and should have the highest measured radiation. A simple explanation of these abnormal results is random error. With the logging interval set at 30 minutes and data polling at 3 minute intervals, it is entirely possible that by random chance, the BH pryanometer happened to average measurements that had significantly more cloud cover.

A second potential explanation of this error is that of the pyranometer itself. The difference range of 4% falls within the +/-5% error brackets of each pyranometer, indicating potential calibration differences between the two pyranometers. To test whether or not the pyranometers were calibrated to the same measurements, each was placed at the BH Site for two days. A comparison of the two results proved that the pyranometers were within 0.6 percent of one another. The only other possibility is that there are measured differences at other temperatures. A longer comparison period would be required to test this hypothesis.

Another source of radiation variance is reflected radiation. Since the GC pyranometer was mounted only 2m above the ground, there is more chance that reflected radiation, particularly from high albedo snow cover that would affect the reading. This is the opposite of the BH pyranometer which was mounted over 27m above the ground surface, with few nearby surfaces that held snow for a prolonged period of time. Without measurements of reflected radiation, it is impossible to validate this theory.

Despite radiation variances between pyranometer data it is still important to note that the GC time produced the largest difference errors. Because both sites are chosen to have almost no solar shading from nearby buildings, and minimal topographic influence, the error in the pyranometer readings is most not likely due to topographical errors. As with all of the sites, there are certain features which were not incorporated into the terrain. The most obvious of these features are trees. Both sites, however, were located in open areas without close proximity to trees. Therefore, any influence from trees should be minimal, though any quantifiable influence would support the study results which showed that measured radiation was lower than modeled radiation.

A more practical source for topographic error is that campus DEM does not include the large topographical influences located outside of the scope of campus; namely the Green Mountains to the east and the Adirondacks to the west. These features were left out of the DEM because of file size constraints, and as a result, were not included in the horizon angle calculations used to create the individual viewsheds. Consequently, there is potential that some of the viewsheds in the model might show sky as visible, when in reality, it is obscured by an obstacle outside of the study area. These changes, however, only affect the model at the margin of the viewshed when the sun is at a very low incidence angle. Because of the low angle of incidence, the effects of extra radiation would be small.

In an effort to quantify the effect of the missing mountains on the horizon, an output duration feature was calculated using *Points Solar Radiation* for the first day of the study time period. The output duration value represents the modeled direct radiation on any particular site. The output duration for the GC was 8.93 hours. This fit well within the range of 8.5 – 9 hours of direct radiation that can be inferred from the pyranometer data. It is important to note that the data loggers only logged radiation averages every 30 minutes. In the future, it would be more useful to use shorter logging intervals so that modeled and measured durations could be compared with better accuracy. Further quantification of the effects was determined by analyzing the radiation during first and last hour in comparison with daily total radiation. Results at the GC for November 17th revealed that low incidence angle radiation accounted for only 3% of the total measured radiation, exemplifying the lesser effect not including horizontal features outside the study area.

Other explanations of the abnormally large errors associated with GC study interval are related to the duration of the measurement period. The measurements were made over the shortest interval in the study, a period of 24 days. One possible source for measured errors at both sites that is magnified by the short time interval is the diffuse proportion calculated from the NSRDB. Since the NSRDB's data is available in monthly intervals, the diffuse proportion cannot be calculated for the exact time interval of each study. In addition, the diffuse proportion calculated from the NSRDB represents a much broader average of diffuse proportion from thirty years of record, and does not relate to this year's conditions. If the weather of November and December were generally cloudier than the average, the calculated diffuse proportion from the NSRDB would result in a higher than measure global radiation. Additionally, as with the snow events, the shorter time interval is potentially problematic. Two or three very cloudy days in a row could have dramatic effects on the data, which might average out if the study interval were lengthened. This theory is supported by the results from the BH and SH Sites, where the longer interval study durations showed an overall reduced percent error. Interestingly, however, both Starr and BHs showed modeled values that were greater than measured values. Without measured values for the proportion of diffuse radiation, there is no way to determine whether or not the Fall of 2006 and Winter of 2007 had greater than average cloud cover as the results suggest.

While the longer time intervals of the SH and BH Sites help to increase the potential that the diffuse proportion will be closer to the seasonal averages, the positive difference error from all three sites is perhaps a stronger indicator that snow cover is the primary source for error in all of the measured datasets. The effect of snow events are also magnified over the shorter time interval from 11/17/2006 – 12/10/2006. Over the course of the study there were a few days with significant snow events which may have covered the pyranometer, effectively blocking incoming radiation. The snow error is magnified over the shorter study interval, because there are less total days of observation included in the total. In the GC measurements, there are two abnormally low radiation days on November 28th and December 1st (Figure 9). These two days (and subsequent snow events) corresponded with measured precipitation data from local weather stations. These two days account for 8.33% of the total duration of the study interval. Assuming that the low readings are indicators of snow cover, there is a high potential that actual radiation values were larger.

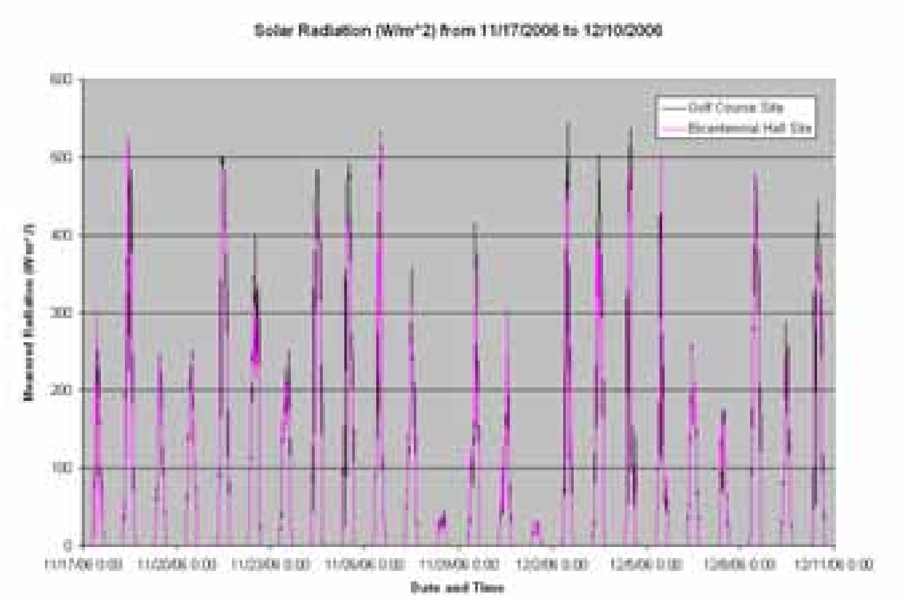


Figure 9. Measured solar radiation over the GC Study interval (11/17/2006 – 12/10/2006) for the BH and GC Sites. Notice the abnormally low radiation values on 11/28 and 12/2 which correspond with precipitation events.

In a closer analysis of the pyranometer data collected at SH, there are four days that show extraordinarily low radiation. These four days comprise 5.6% of the total study duration, suggesting a small, but measurable, source of error. Additionally, a closer look at the BH data reveals that there were ten days over the longer study interval that showed abnormally low solar gain. These days comprised a total of 8.0% of the total duration of the study. When comparing these results to the percent error of the modeled calculation, there is a strong correlation between the percentage of days with potential snow events, and the degree of error. BH and the GC show higher percentages of possible snow days, indicating that the modeled results should be higher, which is indeed the case. The results from SH show a lesser percentage of potential snow error, which is mirrored by increased modeled results by about the same percent (Table 3).

Table 3. Snow Events and Validation Error

Study Site	**Number Snow Events (Days)**	**Duration of Study Interval (Days)**	**Percentage of Days With Precipitation Events**	**Validation Error**
Golf Course	2	24	8.33%	15%
Bicentennial Hall	10	125	8.00%	12%
Starr Hall	4	71	5.63%	5%

Assuming that the percentage of days with snow events is a measure of their induced error on measured radiation, the average validation error at all three sites would drop from 11 percent to 4 percent.

4.1 Validation Results from the Comparison Study

The results from the *Area Solar Radiation* maps showed empirically that the modeled radiation accuracy from the study sites can be extrapolated to the entire study region. Solar shading patterns showed excellent highlighting and shading trends that were consistent with observations of the terrain. Results on and around buildings showed that that the model was able to reproduce the effects from the vertically extruded buildings at a raster resolution of 0.6m.

While the *Area Solar Radiation* maps were consistent with empirical observations from the terrain, it was very difficult to link the model directly to the measured radiation levels. This was mostly due to the area solar model's incorporation of slope into the model calculations. Because all of the pyranometers were mounted horizontally, there was no way of determining how well *Solar Analyst* was able to model sloped aspects. It was also very difficult to compare the area model to the measured results because of the lack of additional equipment. With only two pyranometers, it was impossible to make more than a few comparisons within the study area. In a more ideal setting, many more pyranometers would have been deployed, oriented at many different slopes and aspects, and all over the same time interval. This would facilitate much better categorization and comparison of the modeled results.

4.2 Interpretation of the Effects of Model Parameters on Calculated Radiation

The results from the batch iterations of *Points Radiation* reveal perhaps more about *Solar Analyst* than the validation study. Because the validation study used extremely high-resolution model parameters, it was assumed that these results were the most accurate that the model could produce. While this strategy worked well for *Points Solar Radiation* because the calculation times are relatively small, the use of the same parameters in *Area Solar Radiation* would take hundreds of hours to run for the Middlebury College Study area. The demand for accurate solar radiation maps drives the need for a better understanding of the results from the batch iterations. The conclusions that can be drawn from the batch results are extremely valuable for helping to pick model parameters that balance computational time with model accuracy. This is extremely important for high resolution DEMs, where the step from 1.5% model convergence to 1% model convergence, an increase of only 0.5 %, can mean a doubling of calculation time.

Of the six input parameters tested, skysize is the most important variable to balance with computational times. As described in the ESRI *Solar Analyst* Help, increasing the resolution of the skymaps and sunmaps increases the resolution of the model (ESRI Solar Analyst Help.) As shown in the results, however, there is a point of convergence where increasing resolution does nothing more than to further subdivide pixels. Choosing a sky resolution that is too high will not decrease the accuracy of the model, though it will have tremendous decrease on the modeling time. On the other hand, choosing a sky resolution that is too small will run extremely quickly, though it can have very detrimental effects on modeled accuracy. This is because skysize will effect the calculation of sunmap intervals (personal communication with Ryan DeBruyan.) There is a point at which skysize is too coarse to be able to represent all unique values for every time interval. This results in a grid which cannot represent all sunmap intervals and results in a sunmap that has overlapping days. Because sunmap intervals are so dependent on skysize for their accuracy, skysize should be chosen after all other parameters to ensure that the sky resolution is high enough to accurately model each sunmap interval.

In a more practical analysis, the results from this study suggest that a skysize of 1000 can accurately portray sunmap intervals (at least under 1% error,) at the highest resolution day and hour intervals. Interestingly, the percent error for each skysize at each of the sites is the same for each input parameter. In other words, the percent error associated with a jump in skysize is the same for each variable, whether the tested parameter is day interval, hour interval, etc. This further supports the suggestion that the primary effect of skysize on the model is the accuracy of the sunmap intervals. As the resolution of the day and hour intervals becomes more and more coarse, the skysize diverges. For example, at the SH Study Site, 1% convergence is visible at a skysize of 200 for day interval 14, and hour interval 0.8. These findings support the conclusion that skysize should be chosen after the initial choice of day interval and hour interval.

Of the two parameters that affect the choice of skysize (day and hour interval), hour interval has the most pronounced effects on model accuracy. Of all six tested parameters, hour interval is the only input parameter showing a near linear relationship with resolution and modeled accuracy. While it is true that an hour interval of 0.1 is always the most accurate, the results from the study suggest that the slope of the line is gentle enough that an hour interval of 0.2 might also be an appropriate choice for achieving satisfactory modeled results, with calculations at hour interval 0.2, which is close to within the 1% error range of hour interval 0.1. This is important because moving from hour interval 0.1 to hour interval 0.2 allows a decrease in skysize from 1000 pixels to 800 pixels, further reducing calculation time. Beyond allowing for a lower resolution skysize, the effects of reducing the model accuracy by moving from a 0.1 to a 0.2 hour interval doesn't have significant effect on calculation time. This is because only one sunmap is created for overlay, and the initial calculation time of that sunmap is negligible compared to the calculation time required for all of the individual viewsheds (Personal communication with Ryan DeBruyn).

For this same reason, the effect of day interval on calculation time is also very small. Since hour interval has more pronounced effects on modeled accuracy, the conclusion can be drawn that the choice of hour interval dictates, to some degree, the choice of day interval. For example, if an hour interval of 0.2 is chosen, then this dictates a skysize of 800. The choice of the day interval should then be the highest resolution interval that achieves modeled convergence for a skysize of 800. To continue with the SH example from before, a day interval of 2 would achieve satisfactory results for a skysize of 800. While there is minimal difference between the results of a 2 day interval and a 10 day interval for at the SH Site, the minimal increase in model calculation time due to the creation of only one sunmap per area is not worth the sacrifice in modeled accuracy.

The results from the iterations of calculation directions are the primary indicator for the model's performance in complex terrain. It is not coincidental that the SH Site was the only site where calculation directions played a role in the accuracy of the modeled results. Out of the three study sites, the SH Site has the most complex topography. Increasing the calculation directions, therefore, was essential for the creation of an accurate viewshed. This ensured that every feature was mapped, reducing the potential errors involved with excluding features because they were skipped over with too few calculation directions. What is interesting is that skysizes larger than 400 did not have any additional benefit to model accuracy for all calculation directions. This says two things about the model. The first is that the topographical influence can be accurately portrayed at relatively low sky resolutions. The second conclusion is that the number of calculation directions does not rely on any other input parameters.

Unlike the choice of day and hour interval, however, the choice of the number of calculation directions has a significant effect on calculation time. This is because viewsheds are created for every cell. A doubling of the calculation directions would double the number of calculations for every cell. This in turn will be multiplied by all cells in the DEM and, depending on the size of the DEM, could significantly increase the model time. The results from this study show that 48 calculation directions is satisfactory for complex terrain, though more testing is required for extremely complex environments, such as cities. In addition, the interpolation process that fills in a viewshed between calculation directions attempts to create a smooth surface which mimics natural topography, further requiring greater calculation densities when modeling man-made structures.

In conclusion, the results from this particular study revealed a hierarchy for choosing model parameters. The most important model parameter is hour interval. Once hour interval is determined the next step is to choose an appropriate day interval from convergence studies. These resolutions then help to determine the choice of skysize. Calculation directions should be chosen last, and should be chosen based on convergence studies from topographically complex locations within the study area. While the sensitivity analysis results from this study have many implications for the usage of solar analyst on terrain similar to the Middlebury College Campus, it is still very important to perform a sensitivity analysis when modeling radiation on new terrain.

Conclusion

Overall, the results from this study show that there is a tremendous potential for using *Solar Analyst* on complex terrain. Low sources of topographic error in the study help to suggest that *Solar Analyst* can effectively model complex topography. Additionally, the results from both the modeled and measured radiation suggest that error might be lessened if the study intervals were increased. This reveals the weakest aspect of this study, which was that the time intervals of all three sites were relatively short. This type of study would benefit greatly from data collected over an entire year. This would help to clear up issues such as snow cover, and would be very useful for testing the model in a variety of seasonal solar positions.

While this study is limited to a small study area, over a short time interval, these results, in addition to the validation done by Fu and Rich, (1999) suggest that there is potential for use of *Solar Analyst* in a wide spectrum of applications. With *Solar Analyst* now available as a standard tool in the Spatial Analyst Extension, the availability of the model has increased dramatically. As the use of *Solar Analyst* continues to grow, it is important that users have a solid understanding of the mechanics of the model, in order to better choose an appropriate configuration of input parameters to match their needs. Perhaps the most important result from this study is that errors exceding 20% can be attributed to poor choices of input parameters (e.g. skysize 200, day interval 14, hour interval 1, and calculation directions 8). If other users of *Solar Analyst* are to use this study to help guide their own modeling projects, it is extremely important that they use a sensitivity study similar to the one in this study to help guide their modeling. Additionally, the results from sensitivity studies conducted in other modeling environments could lead to increased information and improved default configurations.

While the results from Middlebury College help to draw important conclusions about the potential accuracy and uses of *Solar Analyst*, it is also very important to note the limitations of this study. The primary weakness of this study is that the measured radiation does not span an entire year. By comparing radiation for only the winter months, this study assesses only one phase of solar positioning during the year. There are many differences between summer and winter solar energy that have significant potential to affect the model. During the summer months, radiation values are much larger because of higher angles of solar incidence, decreased cloudiness and longer days. Both solar position and diffuse proportion are extremely important for *Solar Analysts* calculations of radiation and could lead to other interesting conclusions about the model.

Another major drawback of this study is the lack of overlapping measurements. Having only two pyranometers significantly reduced the ability to compare results between multiple sites. It would be highly beneficial to have more than two pyranometers that could be mounted for an entire year at multiple locations within the study area. Additionally, it would be beneficial to collect data over a range of inclination angles perpendicular to the surface, in order to facilitate comparison of results that included slope and aspect in the calculation. Another advantage to more equipment would be the ability to measure diffuse proportion. This would eliminate the potential differences between actual diffuse proportion and the 30 year summaries from the NSRDB.

In conclusion, this study brings to light some important information about *Solar Analyst's* potential uses. At the same time, it also reveals the need for more testing, and increased use of the model. In addition, there is a lot of potential for increased ability to model buildings more accurately within a DEM. This will be particularly important as more users start to apply *Solar Analyst* to even more complex terrain. With a growing library of 3-dimensional SketchUp and Google Earth models of city buildings, these types of raster based models could be significantly improved if software were developed to rasterize SketchUp models.

As the ability to model increasingly more complex landscapes continues to develop it is important that these surfaces be incorporated into models such as *Solar Analyst.* While more validation and testing of *Solar Analyst* needs to be preformed, the results from this initial study help to highlight the tremendous potential and validity of this toolset.

References

Solar Radiation (Spatial Analyst). ArcGIS 9.2 Desktop Help. Environmental Systems Research Incorporated. 2006

DeBruyn, Ryan
April 2006. Personal Communication.

Dubayah, R. and P.M. Rich. 1995. Topographic solar radiation models for GIS. *International Journal of Geographic Information Systems* 9:405-413

Flint, A.L. and S.W. Childs. 1987. Calculation of solar radiation in mountainous terrain. *Agriculture and Forest Meteorology* 40:233-249

Fu, Pinde
Fall 2006. Personal Communication.

Fu, P. and P.M. Rich. 1999. Design and implementation of the Solar Analyst: an ArcView extension for modeling solar radiation at landscape scales. *Proceedings of the 19th Annual ESRI User Conference*, San Diego, USA

Hetrick, W. A. , P. M. Rich, and S. B. Weiss. 1993b. Modeling insolation on complex surfaces. *Thirteen Annual ESRI User Conference*, Volume 2, pp. 447-458

Kumar, L., A.K. Skidmore and E. Knowles. 1997. Modeling topographic variation in solar radiation in a GIS environment. *International Journal of Geographic Information Science*, 11:475-497

Swift, L.W. 1976. Algorithm for solar radiation on mountain slopes. *Water Resources Research* 12:108-112

Rich, P.M. 1989. *A manual for analysis of hemispherical canopy photography*. Los Alamos National Laboratory Report, LA-11733-M

Author Information:

Nathaniel Vandal
Undergraduate Student
Independent Scholar
Middlebury College
Middlebury, VT 05753
USA
nathaniel.vandal@gmail.com

William Hegman
GIS Specialist
Department of Geography
Middlebury College
Middlebury, VT 05753
USA
whegman@middlebury.edu