

# ACCURACY OF LIGHT POLLUTION MAPS

## The Clear Sky Charts and The Dark Sky Atlas vs. Measurements

### SUMMARY

This reports describes a statistical evaluation of the accuracies of the two currently available forms of light pollution maps; the Clear Sky Charts (**CSC**, [cleardarksky.com/](http://cleardarksky.com/)), and the Dark Sky Maps (**DSA**, [www.taurus-tech.com/](http://www.taurus-tech.com/)). The CSC maps are based on a mix of satellite measurements and computer modeling. The DSA maps are based on the U.S. census and computer modeling. These maps are useful to amateur astronomers for finding observing sites with dark skies.

The light pollution levels reported on these maps have been compared to measurements made across the U.S. with the Sky Quality Meter (**SQM**, [www.unihedron.com/](http://www.unihedron.com/)). The results clearly demonstrate that the **Dark Sky Atlas** is twice as accurate (half the probable error) as the **Clear Sky Charts**.

### THE MEASUREMENTS

The CSC and DSA light pollution models were compared to 217 measurements. Most of the measurements used are listed in the publicly available SQM database web site, and others were made by this author and other contributors. The SQM data base is large, but most entries represent locations outside the U.S. Virtually the entire SQM U.S. database was used for these comparisons. The only exceptions were those measurements made under conditions that were inconsistent with the CSC and DSA model assumptions (see Measurement Issues below), and a few with obvious errors (wrong coordinates, LP values of 23, etc.).

Montana	Indiana	Pennsylvania
Idaho	Illinois	New Jersey
Utah	Tennessee	Delaware
Arizona	Texas	Maryland
California	Oklahoma	Virginia
	Colorado	West Virginia
	New Mexico	Georgia
		Florida

The 217 measurements were made in the 20 states shown in Table 1. Of these, 45 (21%) were made at altitudes ranging from 5,400 ft. to 10,500 ft., with an average of 7,500 ft.

### MEASUREMENT ISSUES

There is a natural tendency to consider any difference between the SQM measurements and the values predicted by the light pollution map as errors in the maps. The discussion below describes why this is not as straightforward as it seems, and lays out the procedures used to better qualify these differences.

A significant aspect of the comparison procedure is explained using the data shown in Table 2. This table presents some of the actual measurements used from the SQM data base. All 4 measurements were made in a 4 week period, in the same location, by the same person, using the same SQM instrument. The average and standard deviation of all 4 measurements is shown, as well as the differences of each from the average.

	Meas. (msa)	Meas-Ave
#1	21.24	0.24
#2	21.10	0.10
#3	20.70	-0.31
#4	20.98	-0.03
Average	21.01	
Std Dev	0.23	

All 4 measurements were made in a 4 week period, in the same location, by the same person, using the same SQM instrument

One might expect that any single location should have a single

value of sky brightness. Each of the light pollution maps only predicts a single value for this location (but not the same value). So the questions is: What caused the variations?

The fact is that the scatter in the measurements shown in Table 2 is common. There are 5 general categories of possible causes.

### Causes of Measurement Scatter

1. **The SQM meter is faulty** - This is not considered a significant effect. In those cases in which multiple SQMs were used, the results agreed closely.
2. **The meter is being used incorrectly** - This can happen and the results can be significant. There are two types of problems. One is technique, and the other is timing. The technique problem is that the SQM "sees" a wide light cone. At a dark site, if the user stands too close to a tree, building, or other tall structure, the meter will be shielded and the reading will be too high (indicating too dark). In a brighter sight, the SQM can pickup nearby light sources and the reading will be to low. Another problem is that the user may not measure in the zenith direction, but at some angle. The sky brightness at most locations is strongly non-uniform. The timing problem is when a measurement is made during astronomical twilight, or when the moon is present. Both models assume no twilight or lunar contributions.
3. **The sources of the light pollution (nearby towns, distant cities) are changing** - This actually does happen. On Sunday night, nearby cities are usually darker than on week nights. Even on week nights, the sky usually gets darker late, after businesses shutdown and people go to bed.
4. **The atmospheric conditions are changing** - This is the largest component of the variations. Light pollution is caused when light from a distant town or city travels up into the atmosphere and is then reflected back down to the observers location. Changes in the atmosphere can either reduce or enhance this transfer. Both the CSC and DSA models assume clear skies within at least 50 miles of the observing site, and average transparency.
5. **The background natural brightness is changing.** - One typical example is the presence of the Milky Way which can reportedly decrease the SQM reading by 0.15 msa. Aurora can have a similar effect.

Of the above potential causes of variation in the SQM measurements, only items 1 and 2 are measurement **errors**. Items 3, 4, and 5 represent actual variations in sky brightness. For simplicity below, I will refer to the variations in measurements due to items 3, 4, and 5 as "**measurement scatter**".

### THE ANALYSIS PROCEDURE

Ideally, if one compares a *perfectly accurate* model to a large number of *error-free* measurements, then the differences (model-measurement) would only be due to the measurement scatter. This means that the differences should have two characteristics. First, the average difference should be zero (because the scatter is defined relative to the average). Second, the standard deviation of the differences should be the same as that of the measurement scatter.

The point of the above statement is that one must know something about the measurement scatter in order to evaluate the accuracy of the CSC and DSA models. This requires two sets of comparisons. First, the measured values are compared to model predictions. These differences represent the *combined effects* of model accuracy *and* the measurement scatter. The second set of comparisons involves only the measurements. If there are subgroups of measurements made over time, at the same location, by the same observer, with the same SQM instrument, then these are compared to their own average value to characterize the measurement scatter.

The second set, the measurement-only scatter, is then subtracted statistically from the first set, the combined effects of model accuracy and the measurement scatter, to obtain the model-only accuracy component.

## THE RESULTS

Two sets of analyses were performed. A composite analysis compared both DSA and CSC maps to all 217 measurements. A mountain analysis compared them to only those measurements made above 5,000 ft.

The results are shown in two formats. Table 3 and 4 summarize a semi-rigorous statistical analysis, and Figures 1 and 2 show the same results in a more intuitive graphical form.

Of the 217 measurements, 132 were in groups of multiple measurements made in 11 different locations across the U.S.. These were used to estimate the standard deviation of the measurement scatter value of 0.129 mpsas. In Tables 3 and 4, this scatter value is statistically subtracted from the "Model - Meas." standard deviation to obtain the "Model Only" standard deviation for both the DSA and CSC models, using the standard *square root, differences of the squares technique*.

The final answers are the numbers in the in the green and rows of Tables 3 and 4.. The green row numbers are what we can see. This is the range of differences between model and measurement. These differences are due to the combined effects of errors in the model prediction plus all of the measurement scatter components (items 1 - 5 above). The numbers in the red rows are the implied model-only errors, after the measurement-only scatter has been subtracted. In other words, if you observe differences between the DSA, or CSC, and measurements that are within the green range, the the actual model accuracy is in the red range.

The conclusion from the results in Tables 3, and 4, is that the **Dark Sky Atlas** is twice as accurate (half the probable error) as the **Clear Sky Charts**.

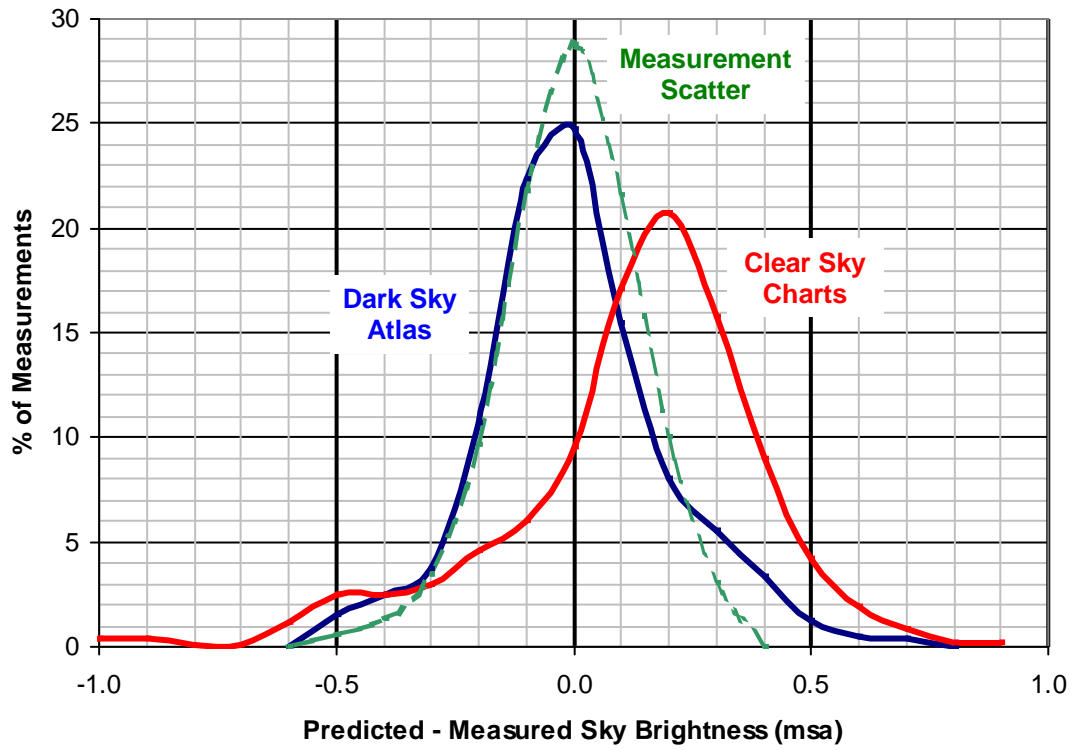
<b>Table 3. Composite Analysis</b> (n=217)			
	Meas. Scatter (mpsas)	DSA Model (mpsas)	CSC Model (mpsas)
		<b>Model - Meas.</b>	
Average Difference	0.00	0.000	0.119
Std. Deviation	0.129	0.188	0.274
<b>Max Observed (95%)</b>		<b>±0.37</b>	<b>±0.57</b>
		<b>Model Errors</b>	
Average		0.000	0.119
Std. Deviation		0.137	0.242
<b>Max Implied (95%)</b>		<b>±0.27</b>	<b>±0.52</b>

<b>Table 4. Mountain Analysis</b> (n=45)			
	Meas. Scatter (mpsas)	DSA Model (mpsas)	CSC Model (mpsas)
		<b>Model - Meas.</b>	
Average Difference	0.00	0.029	0.254
Std. Deviation	0.129	0.173	0.175
<b>Max Observed (95%)</b>		<b>±0.37</b>	<b>±0.60</b>
		<b>Model Errors</b>	
Average		0.029	0.254
Std. Deviation		0.115	0.119
<b>Max Implied (95%)</b>		<b>±0.25</b>	<b>±0.45</b>

The same conclusions can be better visualized from Figures 1 and 2. These are difference-histograms or frequency curves. They describe the % of differences (model - measurement) that had a given magnitude. The green curve with the dashed line represents the measurement scatter only. These are differences of individual measurements from the average measurement of a group made at the same location. The blue and red curves represent the differences between the models and the measurements.

As was discussed above, the histogram for a perfectly accurate model would be identical to the green curve.. In other words, the observed differences would be due only to the measurement scatter. So the green curve provides a frame of reference from which we can judge model accuracy. Model errors are indicated by the difference between the green curve and either the blue (DSA) curve, or the red (CSC) curve. In both Figures 1 and 2, both the blue (DSA) and green curves are centered horizontally on the zero line, with horizontal differences of about 0.1 mpsas. In contrast, the red (CSC) curve is significantly offset from the zero-line, with horizontal differences of about 0.25 mpsas, which is about twice that of the DSA model, in agreement with the results in Tables 1 and 2.

**Figure 1** LP Maps vs. Measurements - All Data



**Figure 2** LP Maps vs. Measurements - Above 5000 ft.

