Calibration Procedures for Large Aperture Scintillometers

John Chilleri¹, Brian Borchers¹, Sebastian Hendrickx-Rodriguez², Jan Hendrickx³

¹New Mexico Tech Department of Mathematics, ²New Mexico Tech Department of Physics, ³New Mexico Tech Department of Earth and Environmental Sciences

Abstract

Evapotranspiration is a large component of New Mexico’s water balance, but is quite difficult to measure in the field. Through the use of large aperture scintillometers, one can reliably measure the daily sensible heat flux, which in turn can be used to derive evapotranspiration. Currently, New Mexico Tech has nine first generation Kipp & Zonen scintillometers, as well as one second generation Kipp & Zonen scintillometer. The second generation scintillometer is more advanced and accurate in its measurements. In order to ensure high quality measurements from the first generation scintillometers, a procedure has been developed to calibrate the first generation scintillometers against the second generation scintillometer, so that the ten scintillometers can be optimally deployed for water resources research in New Mexico.
Contents

1 Introduction 3
  1.1 Scintillometers ............................................ 3
    1.1.1 Scintillometer Transmitter ............................ 4
    1.1.2 Scintillometer Receiver ............................... 5

2 Methods 6

3 Results 9

4 Discussion 9

5 Conclusion 10

6 Bibliography 11
1 Introduction

Evapotranspiration is difficult to measure in the field. In a 2015 Student Research Award Project Report, Reid Brown presented evidence that current Large Aperture Scintillometers (LAS) have improved tremendously and can be used for reliable daily sensible heat flux measurements. Subtracting the daily sensible heat flux from the daily net radiation, measured on the ground or determined through satellite remote sensing, will yield the daily actual evapotranspiration.

Calibration is required between the older, first generation, and current, second generation scintillometer, to assure high quality measurements of the sensible heat flux.

1.1 Scintillometers

A scintillometer is a scientific instrument used to measure the fluctuations of the refractive index of air caused by variations in temperature, humidity, and pressure. A scintillometer consists of a laser transmitter and a receiver at opposite ends of a clear horizontal or inclined path ranging from 500m – 5000m. The magnitude of the fluctuations of the refractive index is usually referred to as $C_n^2$, known as the structure parameter of air that is – among other things – responsible for the twinkling of stars and for the mirage effect on a hot day. The structure parameter can be used to quantify astronomical seeing conditions at the Magdalena Ridge Observatory or to calculate – together with meteorological measurements – the sensible heat flux at the land surface. The latter is ultimately used in an energy balance for the estimation of actual evapotranspiration in the Middle Rio Grande Valley.
1.1.1 Scintillometer Transmitter

Figure 1: Second Generation Kipp & Zonen Scintillometer - Transmitter.

The transmitter sends a signal to the receiver, measuring numerous quantities, which are used to calculate the daily sensible heat flux and estimate evapotranspiration.
1.1.2 Scintillometer Receiver

Figure 2: First Generation Kipp & Zonen Scintillometer - Receiver.

The receiver and transmitter are ideally, perfectly aligned. In order to achieve good alignment, the signal strength (analog on first generation) must be maximized, which can be done with slight adjustments to the direction of the receiver, which will alter the signal strength ever so slightly. The signal strength is also worsened by the air, so it will realistically never be 100%.
2 Methods

Initially, multiple data sets between a first and second generation scintillometer were acquired. Before a calibration procedure can be developed, the data must first be considered for bias and accuracy. Accordingly, the data was filtered for unjustifiable outliers and unrealistic conditions, and a fit was attempted. Unfortunately, the fits were highly inadequate when completed for polynomials of higher degrees (linear - quintic, with and without intercepts). It was decided to stop at quintic as the $R^2_{adj}$ began to worsen, which is generally an indication that the polynomial is of high enough degree. Regardless, all of the fits produced undesirable $R^2$ values. Clearly, further considerations were necessary before fitting.

After consultation with experts in the field, further theoretical limitations for reasonable data were determined and the data set was filtered accordingly. Additionally, weather data was used to interpret the weather conditions, such as fog, which was excluded from the data.

The filtering procedure was as follows:

1) Exclude data with windspeed $> 5$ m/s
2) Exclude data with demod values outside of $\pm 10\%$ of the average
3) Exclude data if its first generation scintillometer has $\text{PuCn}_2 > 1500$
4) Exclude data if its second generation scintillometer has $\text{PuCn}_2 > 2000$

This filtering of the data should eliminate all atmospherically undesirable conditions, retaining the data that cleanly compare the first and second generation scintillometers.

In order to find a good calibration for the newly filtered data, numerous regression techniques were considered - sum of squares, root-mean-square-percent-predictive error (RMSPE), and finally mean absolute percentage error (MAPE). Although initially attempted, the sum of squares technique is inadequate as it gives undue weight to the higher measurements, whereas all measurements should be equally weighted. This led to fitting via cross-validation with the RMSPE - a technique that gives equal weights to all measurements.
Although the RMSPE fits seemed good for the most part, two of the seven scintillometer data sets contained irregular data, which resulted in poor fits. Scintillometer 50016 data containing the most extreme irregularities was analyzed. Further investigation revealed that the ratio between the first and second generation scintillometer measurements spiked in approximately ten places for a period of time, causing unusually high differences in measurements. For example, from 10:20 to 10:46 on April 9th, the ratio between the measurements exceeded 175,000 at every single minute of measurement, whereas the second generation had typical readings, meaning the first generation instrument was reading over 175,000 times the second generation instrument. Clearly these data were incorrect, and were excluded. At two other of these spikes, the difference between the first and second generation scintillometer measurements varied by a factor of over 250, with unusually high measurements in between the peaks. These times of nonsensically high ratios were also removed, resulting for a total removal of 783 data points, roughly 13 hours from the eight days of data. It is important to note that all of the other data sets did not contain spikes comparable to the ones found in this data set, suggesting that there may have been further issues, possibly with this particular scintillometer, that may need to be addressed separately.

However, in order to remain cautious, no data was deleted without justification, regardless of if it was somewhat outlandish. These outliers lead to the RMSPE’s downfall, as it is sensitive to outliers, which heavily penalize the

![Figure 3: RMSPE and MAPE fits to Scintillometer 30005.](image)
result (Figure 4).

Figure 4 shows how the RMSPE can fail when challenged with a high number of outliers (mainly located near 0). Finally, using the MAPE seemed the correct technique, as it is more robust and less affected by outliers - and allows us to minimize the percent error, essentially accuracy of measurements.

Mathematically,

$$\min_{\beta} \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\text{Predicted}_i - \text{Measured}_i}{\text{Measured}_i} \right|$$

where $\beta$ are the parameters for the Predicted values: intercept and slope.

MAPE was able to overcome the outlier issues, as seen in Figure 4. The MAPE fits were computed and tested with cross-validation. This is performed by parsing the data randomly into two groups, the first consisting of 90% of the data, and the second containing the remainder (90-10 is a tolerable split when there are many thousands of data points). The prediction is then calculated for each of the first groups, and tested against the second groups.
3 Results

As aforementioned, the results of fitting the entirety of the data were poor, but after having separated the data by desirable weather conditions, the results drastically improved.

The results of the MAPE and cross-validation were as follows:

<table>
<thead>
<tr>
<th>Scintillometer</th>
<th>Calibrated MAPE (%)</th>
<th>Intercept</th>
<th>Slope</th>
<th>Uncalibrated MAPE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60032</td>
<td>12.9</td>
<td>-1.38</td>
<td>0.94</td>
<td>18.4</td>
</tr>
<tr>
<td>50024</td>
<td>22.5</td>
<td>-0.30</td>
<td>0.93</td>
<td>25.9</td>
</tr>
<tr>
<td>60031</td>
<td>12.5</td>
<td>-0.38</td>
<td>0.95</td>
<td>14.5</td>
</tr>
<tr>
<td>70045</td>
<td>11.3</td>
<td>-0.74</td>
<td>0.79</td>
<td>35.1</td>
</tr>
<tr>
<td>50016</td>
<td>22.0</td>
<td>-0.37</td>
<td>0.83</td>
<td>31.8</td>
</tr>
<tr>
<td>30005</td>
<td>11.9</td>
<td>0.00</td>
<td>1.03</td>
<td>12.3</td>
</tr>
<tr>
<td>50015</td>
<td>12.4</td>
<td>-0.31</td>
<td>0.96</td>
<td>14.3</td>
</tr>
</tbody>
</table>

The above MAPE values are percents - which were decreased sometimes drastically by the calibration.

4 Discussion

The sum of squares approach is an inappropriate regression technique for this calibration as it does not give equal weight to the small and large values and doesn’t quite minimize the practical quantity. RMSPE is less inappropriate as it begins to minimize the practical quantity, but it is not applicable because an outlier can blow up the procedure. The choice of the robust technique MAPE not only is penalized less by outliers, but also minimizes the practical quantity - percent error, as the calibration should aim to minimize the percentage difference between the first and second generation scintillometers. The results of the MAPE were in some cases remarkably better than their uncalibrated counterparts, with calibration slopes around 0.8, suggesting that these particular first generation scintillometers consistently measure 80% the value of the second generation scintillometer. However, in other cases the calibrated MAPE was only slightly better than the uncalibrated MAPE. In these cases, the slopes of the calibrated MAPE were closer to one, suggesting that these scintillometers already measure similarly to the second generation scintimeter.
5 Conclusion

The filtering procedure and MAPE cross-validation fitting proved to radically improve the results for some of the scintillometers; however, other scintillometers tested to be near on par with the second generation scintillometer. In order to improve these results in the future, we suggest that further filtering be justified and applied - especially to the weird data occurrences - two of the scintillometers produced unusual data containing strange, likely incorrect spikes. We don’t currently know why these instruments entered these brief periods of ridiculous measurements, but it required that a robust regression technique be used, as we would recommend be utilized in future work.
6 Bibliography

1) Brown, R. and Hendrickx, J.M., Test of the new LAS MkII Scintillometer for validation of statewide New Mexico evapotranspiration maps.
