**R-Dial to N values**

As we discussed in the last chapter, we need to turn our R-dial readings from the Dobson into what are known as “N-values”. Let’s start at the dial readings and work towards getting our N-values.

The R-dial is the silver dial on top of the instrument that has a range of between zero and three hundred degrees. This dial is connected to an optical wedge, where the greater numbers correlate to a greater absorption of light.



Figure Cartoon diagram of the optical wedge. The white (optically thin) portion is quartz and the shaded (optically dense) portion is chromel, with 300 marking the maximum absorption and 0 marking near zero absorption. The deposited layer of chromel is in reality very thin.

The absorption of the wedge is a measured quantity and is quantified in an R-to-N table for each instrument, rather than assuming all instruments are the same. The R-to-N table can be found in cmet\OZONE\dobsons, in the same file as the constants α and β where it is located here:

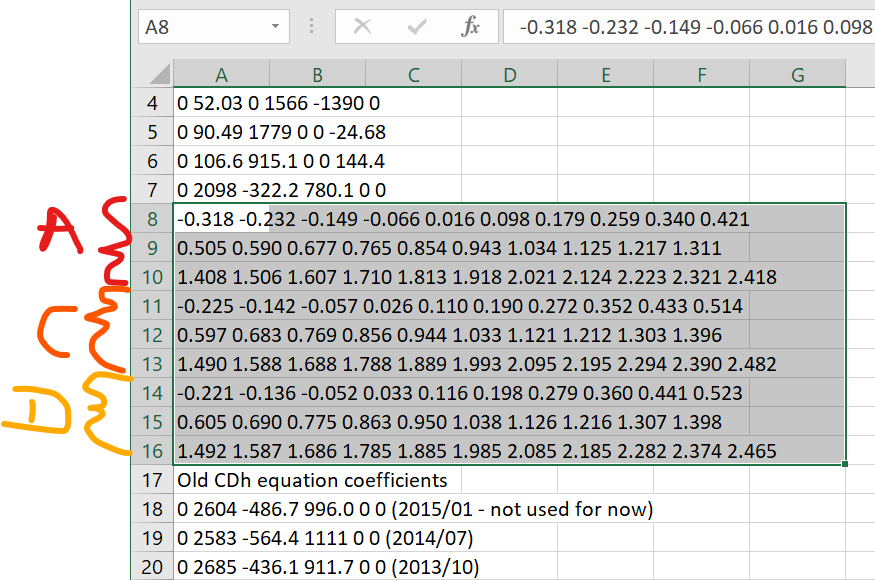


Figure 2: Highlighted portion of where the R-to-N tables can be found in the data file D031E.DAT

As with our α and β constants, the absorption varies with wavelength, so there are three different R-to-N tables corresponding to the A, C and D wavelength pairs that we measure with the Dobson. Each table goes from left to right giving the values that correspond with 0, 10, 20… 290, 300 degrees. A negative value *does not* mean that it emits light rather than absorbs, only that it absorbs less than the standard that it is compared to during calibration.

Whilst our table does not correspond to every single value on our dial, we use the values we are given (in our case, every 10 degrees) and interpolate to get the values we desire from our table. (Interpolation only works if the absorption layer is deposited sufficiently smoothly for the absorption to be slowly varying. If it should vary more rapidly the lookup table would need to be at higher resolution)

For example, say we have a dial value of 41.2 while measuring the A wavelength pair. Our dial reading is between the values for 40 (0.016 in the table) and 50 (0.098). Using the interpolation formula:

With 40, 41.2, 50 as x1, x, x2 respectively and 0.016 and 0.098 as y1 and y2, we get a value of 0.026 (to 2d.p.). In this way we can obtain values for any dial reading.

So, is this then our N-value? Nearly, but not quite. If the instrument had just been sent for calibration, then yes, this table would give us our N value. However, for our Antarctic instruments, the travel time is long enough and the instrument is exposed to such a range of temperatures and pressures, as it travels across hemispheres that we assume the instrument will fall out of calibration slightly. From the previous chapter, N is defined as:

L0 is assumed to be a constant and is defined as what the Dobson would measure outside the Earth’s atmosphere, for this reason it is sometimes referred to as the “Extra Terrestrial Constant”. If we had a perfect instrument, we would not need to muddy the waters further; however, what the Dobson would measure outside the Earth’s atmosphere would of course differ slightly from the real value by a small factor “k”, which we will call the instrument constant. Let us redefine L0 slightly with this in mind:

The value of L0 is fixed during calibration for each instrument – but it will drift with time. This could be due to the slow and subtle changes in the instrument or even through changes in the solar cycle affecting the ratio of the wavelengths slightly. Either way, our value of L0 from the calibration will become inaccurate with time and have an error associated with it. We must then redefine N slightly:

[1]

Where L0\* is the estimated value of L0 and S is the error in L0\*. We can simplify this a little (and to save us from referring to several different L’s) to become:

All this really means is that we assume our values of N\* from each of the R-to-N Tables (recall there is a table for each wavelength) needs a correction factor. How then, do we find these correction factors “S”? One such method is to have a “standard Dobson” by which everything else is compared to. This standard can then be ferried around and perform measurements next to another Dobson, thus an intercomparison can be made. However, this is impractical given the nature of working in the Antarctic, where logistics are difficult and expensive. So, without this option, there are two primary methods of deriving this correction factor, as well as a potential third method discussed at the end:

**Finding S: Standard Lamps**

The first (and simplest) is to use a standard lamp. Each Dobson gets assigned at least two standard lamps during its lifetime and the changes to the instrument can be tracked by seeing how the dial readings change in time for each of the three wavelengths (A, C and D) emitted by the lamp. A second lamp is used as a sense check to see if there have been any changes in the first. Any changes in the lamp readings can be converted to an N value and applied as a correction ΔN. [1]

This method is not outlined in great detail, as there is no guarantee that the lamps or the power supply are not changing in time, whereas the other methods listed rely only on the instrument itself.

**Finding S: Langley Method**

The other method for deriving S, known as the “Langley Method”, is quite a bit more complex and uses Direct Sun data to derive this correction factor without relying on other instruments as a reference. We will first go into how to derive this factor for the simplest case of a single wavelength with one day of data, before discussing how to derive it for multiple days and for wavelength pairs (e.g. AD or CD).

**Langley Method - Single wavelength, single day:**

The first step is to acquire the data from a day where there are at least 3 Direct Sun data points and where ozone is changing slowly (i.e. not when the ozone hole is forming or filling). These 3 points should ideally be across a morning or an afternoon, as this will give the greatest range of μ values (μ is symmetrical about solar noon).

Once you have the points for a given day, you take the dial readings from the Dobson and their associated μ values and plot N\*/ μ (remember N\* is the uncorrected N value) against 1/μ. The reasons for doing this are found in the derivation below, starting with our direct sun equation.

We then take the following steps:

* assume to be very small (this is true for areas with low pollution levels such as Antarctica)
* Call and terms and for simplicity
* Change N to “(N\* + S)”
* Assume p/p0  ~= 1 (this holds for stations near sea-level)

Assume m~ =

Assume X is roughly constant over the period of observation, therefore: can be called a constant K:

Which is of the form y = mx + c, where y = (N\*/μ), x = (1/μ), m = -S and c = K. Hence, we plot N\*/μ vs 1/μ and find the gradient of the graph to find -S. A small but easily missed point – our S value is the opposite sign to the gradient, if the gradient we find is *positive*, our value for S is *negative*.

Now that we understand where this comes from, let’s go through an example. The 2020/2021 season had relatively few observations due to a restricted season from the pandemic, but does have a few half days of Direct Sun obs. First we find a day that has sufficient Direct Suns – the codes for these are 00, 20 and 90. Looking through the file from the SAN with January’s data (data\momu\_data\halley\ozone\2021\data\jan\zozJan21.dat) we can see that there are many of these observations on Jan 19th:

Table Excerpt from zozJan21.dat. Note that Dial readings are multiplied by 10 e.g. 1636 = 163.6.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| year | month | day | hour | min | sec | ob type | Dobson ID | R-Dial 1 | R-Dial 2 | R-Dial 3 | R-Dial 4 | R-Dial 5 | R-Dial 6 |
| 2021 | 1 | 19 | 11 | 58 | 10 | 90 | 031E | 1636 | 1636 | 990 | 992 | 725 | 727 |
| 2021 | 1 | 19 | 12 | 0 | 18 | 92 | 031E | 1360 | 1363 | 654 | 660 | 362 | 364 |
| 2021 | 1 | 19 | 13 | 29 | 50 | 90 | 031E | 1582 | 1585 | 958 | 965 | 700 | 700 |
| 2021 | 1 | 19 | 13 | 31 | 47 | 92 | 031E | 1326 | 1319 | 635 | 634 | 349 | 345 |
| 2021 | 1 | 19 | 14 | 25 | 48 | 90 | 031E | 1593 | 1591 | 970 | 966 | 709 | 707 |
| 2021 | 1 | 19 | 14 | 28 | 3 | 92 | 031E | 1333 | 1332 | 638 | 640 | 346 | 348 |
| 2021 | 1 | 19 | 17 | 55 | 40 | 90 | 031E | 1873 | 1868 | 1129 | 1129 | 814 | 816 |
| 2021 | 1 | 19 | 17 | 57 | 51 | 92 | 031E | 1548 | 1545 | 752 | 758 | 408 | 409 |
| 2021 | 1 | 19 | 18 | 40 | 58 | 90 | 031E | 2031 | 2025 | 1223 | 1218 | 871 | 868 |
| 2021 | 1 | 19 | 18 | 42 | 55 | 92 | 031E | 1677 | 1674 | 820 | 823 | 436 | 436 |
| 2021 | 1 | 19 | 20 | 2 | 52 | 22 | 031E | 947 | 953 | 959 | 506 | 508 | 0 |
| 2021 | 1 | 19 | 20 | 46 | 48 | 20 | 031E | 1577 | 1583 | 1583 | 1105 | 1108 | 0 |
| 2021 | 1 | 19 | 20 | 48 | 59 | 22 | 031E | 1044 | 1047 | 1044 | 555 | 558 | 0 |

From here, we can take the C and D readings for all the highlighted obs, and the A readings for the first 5 highlighted obs (AD observations can be taken over a smaller range of sun angles, here code 90 is a “CDA” ob and code 20 is a “CD” ob i.e. A wavelengths are not measured in these) and convert the R-dials to N\* readings (using the file from Z:\cmet\OZONE\dobsons\D031E.DAT). Firstly we can neaten up the formatting and divide the R-dial readings by 10 to put them back into their original values:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Date | Time | A1 | A2 | C1 | C2 | C3 | D1 | D2 |
| 19/01/2021 | 11:58:10 | 163.6 | 163.6 | 99.0 | 99.2 | 0.0 | 72.5 | 72.7 |
| 19/01/2021 | 13:29:50 | 158.2 | 158.5 | 95.8 | 96.5 | 0.0 | 70.0 | 70.0 |
| 19/01/2021 | 14:25:48 | 159.3 | 159.1 | 97.0 | 96.6 | 0.0 | 70.9 | 70.7 |
| 19/01/2021 | 17:55:40 | 187.3 | 186.8 | 112.9 | 112.9 | 0.0 | 81.4 | 81.6 |
| 19/01/2021 | 18:40:58 | 203.1 | 202.5 | 122.3 | 121.8 | 0.0 | 87.1 | 86.8 |
| 19/01/2021 | 20:46:48 | 0.0 | 0.0 | 157.7 | 158.3 | 158.3 | 110.5 | 110.8 |

Then using our interpolation formula and the table from D031E.DAT we can get the N\* values. As well, we can get mu for each reading by converting the datetime, latitude and longitude to a solar zenith angle (this is not trivial, however there are solar calculators available, either as software or online) [3] and converting the zenith angle to mu using the formula described in the previous section. [4]

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Date | Time | Mu | N\*(A1) | N\*(A2) | N\*(C1) | N\*(C2) | N\*(C3) | N\*(D1) | N\*(D2) |
| 19/01/2021 | 11:58:10 | 1.839 | 1.06676 | 1.06676 | 0.589 | 0.59 |  | 0.38025 | 0.38187 |
| 19/01/2021 | 13:29:50 | 1.752 | 1.01762 | 1.02035 | 0.562 | 0.568 |  | 0.36 | 0.36 |
| 19/01/2021 | 14:25:48 | 1.757 | 1.02763 | 1.02581 | 0.572 | 0.569 |  | 0.36729 | 0.36567 |
| 19/01/2021 | 17:55:40 | 2.206 | 1.28562 | 1.28092 | 0.708 | 0.708 |  | 0.45248 | 0.45412 |
| 19/01/2021 | 18:40:58 | 2.429 | 1.43838 | 1.4325 | 0.789 | 0.785 |  | 0.49922 | 0.49676 |
| 19/01/2021 | 20:46:48 | 3.471 |  |  | 1.101 | 1.106 | 1.10604 | 0.69425 | 0.6968 |

From here we can take the mean values of each wavelength and calculate 1/mu and N\*/mu for each wavelength:

|  |  |  |  |
| --- | --- | --- | --- |
| 1/mu | N\*(A)/mu | N\*(C)/mu | N\*(D)/mu |
| 0.543684 | 0.579981 | 0.320518 | 0.207176 |
| 0.570621 | 0.581454 | 0.322426 | 0.205424 |
| 0.569265 | 0.584476 | 0.324732 | 0.208624 |
| 0.453366 | 0.581791 | 0.320956 | 0.205511 |
| 0.411737 | 0.591024 | 0.323969 | 0.205041 |
| 0.288112 |  | 0.317649 | 0.200389 |

We then plot these values and find a value for S from the gradient of these graphs:

Figure 3: graphs showing a Langley plot for the three wavelengths on one day

Which give us S values of 0.039, -0.015 and -0.022 for the A, C and D wavelengths. As we can see, the fit of the lines is not perfect and many factors could confound our calculation, a few observations could be unreliable, our assumption of ozone staying constant through the day could be wrong, or increasing haze or thin cloud could affect the readings. Therefore, this process must be repeated for multiple days, so that an average value of S can be computed for the season.

**Langley Method - single wavelength, multiple days:**

As a worked example of this, let’s look at a different year with more days of useable data (the 2020/21 season only has 5 such days). We’ll examine the record for 2015, just before Halley moved to its Halley 6A location and we’ll look at just a single wavelength – the A wavelength (though this method is applicable to any of the three).

Through a program that automates much of the work of the previous section [5], the S values for all available days of the 2015 have been calculated, alongside their standard error [6]. These are displayed below:

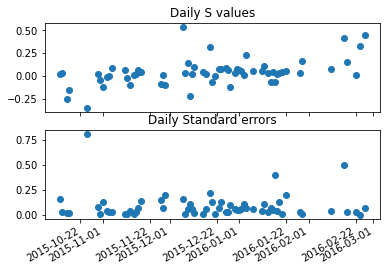


Figure 4: S values, calculated from daily Langley plots, plotted over the course of the 2015 season, alongside their associated standard errors.

As you can see, the S values are relatively stable over the course of a season, but they are quite noisy. Because of this, we should do a little quality control. There are a couple of ways to do this: we could give additional weight to days that have many observations (particularly morning and afternoon); exclude any value of S outside three sigma of the mean or we could exclude anything above a certain standard error. As with any quality control, there is no “right” way to do it, but it is important to be consistent with how you do it across the whole dataset.

In this example we’ll exclude anything outside 3 sigma:

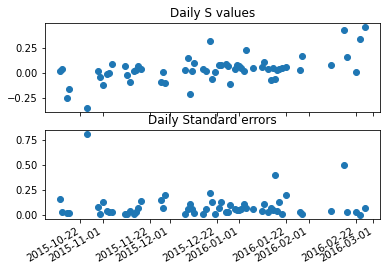


Figure 5: Same plot as Figure 4, after a 3 sigma filter has been applied

This has simply removed the value of S around +0.5. For this example, we shall leave the quality control here, but we could always do more (one data point with an especially high standard error and low S value seems ripe for exclusion). We now take the median value of the S values across a season, as this is less affected by extremes than the mean and we now have a plausible value for S. In this case, that value is -**0.036** (for the A wavelength). Using the same method, we obtain values of -**0.024** for the C and -**0.016** for the D wavelength.

**Langley Method – wavelength pairs, multiple days:**

You may recall from the previous chapter that we can use a pair of wavelengths for our ozone calculations than a single wavelength. This is done because using a pair of wavelengths means that some aerosol scattering effects tend to cancel out ~~when doing this~~.

Therefore, we should look at what the S value for a wavelength pair is across a season, rather than just for one wavelength. We could just subtract one S value from another – e.g. the S value for the AD pair would be:

S(AD) = S(A) – S(D)

Or we could subtract the N\* values from one another before we plot the graphs of N\*/mu vs 1/mu. In this case we would derive our S(AD) value explicitly. Let’s try both and see if they produce the same answer.

For 2015, the S values for a single wavelength are (derived using the method above): S(A) = -0.036, S(C) = -0.024, S(D) = -0.016. So **S(AD) = -0.02** and **S(CD) = -0.008**.

Now to derive S(AD) and S(CD), using the method in the previous section, but from the gradients of plotting N\*(A)-N\*(D)/mu vs 1/mu and N\*(C)-N\*(D)/mu vs 1/mu. Here we find that **S(AD) = -0.013** and **S(CD) = -0.006**

We know that our second method has produced lower values than the first, so now what? We must test our S values on real data and see what the effect is.

A good test of how accurate our S values are is to see what spread of ozone values it gives us. The worse an S value is the more variability there will be across the course of a day or so (a timescale when column ozone is unlikely to change very quickly). Obviously there will be some good days and some bad days for observations, so we need to do this across a whole season so that it evens out. What we can do then, is to take the standard deviation of column ozone for each day of a season, and take the average of all of these values. This then gives us a metric of how good our S value is, we’ll call this metric “P”.

In order to assess which of our two S’s are best then, we do the following:

* Compute what the standard deviation is for each day’s direct sun measurements
* Average this across a season (P)
* Repeat this for the second S value
* Compare the two – whichever has the lowest P is taken to be correct.

To continue our worked example, we can find P for the S-values we found previously:

|  |  |
| --- | --- |
| S-value (AD) | P |
| Method 1: -0.02 | 3.80691 |
| Method 2: -0.013 | 3.78333 |
| S-value (CD) | P |
| Method 1: -0.08 | 6.78969 |
| Method 2: -0.06 | 6.78813 |

From this table, we can see that **-0.013** for AD and **-0.06** for CD are the better values, though the difference in P between our two choices is fairly small.

Now that we have fixed our S value for AD, we assume that the D value calculated in the single wavelength method is correct and use S(A) = S(AD) + S(D) to find S(A). In our 2015 season example: S(A) = -0.013 + -0.016 = **-0.029**.

Then with our S(CD) and S(D) value, we use S(C) = S(CD) + S(D) to find the S value for the C wavelength: S(C) = -0.06 + -0.016 = **-0.022**.

These can then be placed into the annual constants file.

I hope from the above that it is clear that this is not a trivial process - it has certainly taken some time to get my head around – however it is the process that is currently used at BAS. I would encourage you to contact myself ([jevesonjr@hotmail.co.uk](mailto:jevesonjr@hotmail.co.uk)) or Jon Shanklin if any of the above is unclear.

To summarise the process in a concise way:

1. Filter observations so that you only have Direct Sun data, where there are >=3 points and ozone is changing slowly.
2. Plot N\*/mu against 1/mu for these days, for individual wavelengths and then for AD and CD.
3. Take the *negative* of the gradient for these graphs, these are your S values across the season.
4. Perform quality control on any suspect points.
5. Take the median of these quality controlled points to give S values for A, C, D and AD, CD.
6. Use the standard deviation of column ozone across a day, averaged across a season (P) to assess whether (S(A)-S(D)) or S(AD) provide a better S(AD) value. Do the same for S(CD).
7. If S(AD) has a lower P than (S(A) – S(D)), assume S(D) is correct and derive S(A) = S(AD) + S(D) then derive S(C) = S(CD) + S(D)
8. Check the derived S values against standard lamp values
9. Write the S values (A, C and D) into the annual constant file.

Notes:

Not sure where to fit this in, but this is a note from Jon:

It’s worth checking that the range of mu values used is not so great as to affect our derivation – there are likely some DS measurements that are at greater mu values that should not be used for deriving S. This depends on the atmospheric conditions and should not be considered a hard cutoff necessarily, as it will depend on atmospheric scattering, which varies with station, time of year etc.

**A New Method? – the Eveson Method?**

When Jonathan Shanklin recommended that I look at the spread of ozone values that an S value gives, it occurred to me that you could simply take a suitable range of S (based on historic values) and iterate through this range to see how this affects the data.

I have used S values at intervals of 0.002 between -0.1 and +0.1 to calculate column ozone values for each observation across the 2015 season. Then for each calendar day I have calculated the standard deviation for the direct sun observations. To get a measure of how good an S value is across a season I have taken the mean of all the daily standard deviations, in the same way as the end of the Langley method, discussed in the previous section. Then I have plotted the average season standard deviation against the value of S used to produce the following graphs:

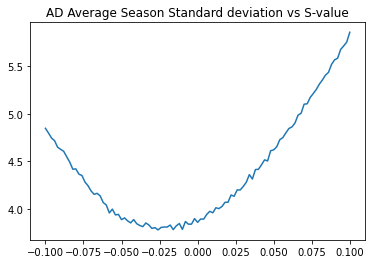


Figure 6: Plot showing how "P" (y axis): the daily standard deviation of column ozone, averaged across a season, changes with different S values (x axis) for the AD wavelengths.

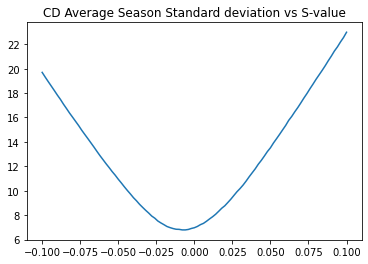


Figure 7: Plot showing how "P" (y axis): the daily standard deviation of column ozone, averaged across a season, changes with different S values (x axis) for the CD wavelengths.

It is unclear to me why the AD graph is so much more jagged than the CD, but, with a bit of smoothing, such as fitting a polynomial curve to the graph, it is quite simple to derive S based on the minimum of these graphs. Whilst these S values have not been tested to a great degree, it seems like this is be a simple and intuitive method for deriving S.

Jonathan Shanklin has the following comments:

*“Hi Josh,*

*That looks pretty neat.  Probably a far better way of determining Lo (S)\* than using gradient plots.  I presume you can output the minimum point on the curve and give a standard error to it.  Next step would be to see how that changes from one year to the next, and how that compares to the lamp tests.*

*For Halley I don’t think it makes much difference whether you use calendar days or 24 hour periods about local noon.  For Vernadsky it might, so it would be best practice to use local time for both.*

*Regards,*

*Jonathan Shanklin”*

\*Jonathan refers to S as Lo, this is something I’ve changed in my notes to be more consistent with the wider literature.

To test these values (or indeed any S values that are obtained) we should check what we produce against the lamp tests. The lamps are not entirely definitive, as we do not have any guarantees the lamps are not changing their output with time, but they provide a good sanity check on our S values.

**Conclusion – Calculating Ozone Values**

Once a satisfactory Sis obtained by any one of these methods, we can finally compute a value for ozone using one of the direct sun equations:

Text, letter

Description automatically generated

Where N = N\* + S.

However in performing direct sun observations, we are at the mercy of the weather, as we are unable to perform them when the sun is obscured by cloud, fog or blowing snow; these conditions could last for days at a time. In such situations, we have to resort to zenith observations. The problem with these however, is there is no single, well defined equation for turning these into column ozone values. Instead, we must perform a regression analysis, linking each zenith observation to a nearby direct sun and then creating an equation (often some kind of polynomial) that fits this pattern. This is what we will discuss in the next chapter.

[1] A derivation for the Langley method was taken from “Dobson Spectrophotometer 83: A Standard for Total Ozone Measurements, 1962-1987” Komhyr, Grass and Leonard (1989)

[2] A guide to performing corrections based on standard lamp measurements is found in Annex E of the WMO Operations Handbook for Dobson Spectrophotometer (<https://library.wmo.int/doc_num.php?explnum_id=9405>)

[3] To find the Sun’s position for a given position and datetime, <https://www.esrl.noaa.gov/gmd/grad/solcalc/azel.html> is a trusted resource from NOAA, or, as I have done, you can use “\_\_get\_zenith” from the autodobson processing repo (this function relies on astropy and numpy libraries from python): <https://gitlab.data.bas.ac.uk/joson/autodobson/-/blob/master/autodobson/ozone.py>

[4] For these steps “\_\_get\_mu” from the autodobson processing repo was used. <https://gitlab.data.bas.ac.uk/joson/autodobson/-/blob/master/autodobson/ozone.py>.

[5] Get\_S.ipynb can be found in a folder “Jupyter Notebooks” with these handover notes.

[6] Standard Error of the regression: <https://statisticsbyjim.com/glossary/standard-error-regression/> or <https://en.wikipedia.org/wiki/Reduced_chi-squared_statistic>

Notes:

I have created a program using a Jupyter Notebook: Get\_S.ipynb.

Get\_S performs the following actions:

1. Given a path with the Dobson data files in, it finds the data points that could be used for an S calculation (three or more over the course of a day) and writes these points to a file
2. It then calculates the mu value, N\* value and 1/mu.
3. Plots 1/mu vs N\*/mu for each day's S observations and stores the gradient and the standard error.
4. Allows the user the option of doing some quality control. Are any points anomalous, or do any of them have too high of a standard error?
5. Produces a suggested S for the season using the Langley Method
6. Puts this S value in the context of previous S values to see if it is sensible
7. Compares the changes in S to the changes in standard lamp tests (these should provide a sense check and be very similar)

This should go some way to making the process less laborious than what has been described in the above section, drawing graphs out individually and manually searching for days where there are enough direct sun observations to perform an S calculation.

Once this has been done, the value of S for the season can be compared to previous seasons in order to see if it is a plausible value or if any sudden changes have taken place in the instrument. One further sense check on the value of S is to compare it to values from the standard lamp test. Both of these provide a check on the state of the instrument and should, when plotted against one another, be a straight line.