Scientific Method

Steve Olney (VK2XV) – NRARAO

26 March 2015

The scientific method is discussed and why it is important to adhere to its principles

Wired for Optimism

Studies have shown that we are hard-wired to be optimistic. When a group of individuals are asked to estimate the likelihood of certain events occurring they are prone to over-estimate the likelihood of good things happening (winning the lottery) and under-estimate the bad things (getting cancer). Furthermore, when presented with the actual probability of these events, these same individuals, upon re-questioning, revised their estimates upwards if they were overly pessimistic originally. However, if they were overly optimistic originally they tended to keep their optimistic view.

In other words we tend to look at the world through rose-coloured glasses. It also explains why individuals who by nature have a realistic view of life tend to be labelled as pessimists.

While there are probably good evolutionary reasons for this optimistic view of the world (otherwise we might never have emerged from the caves for fear of the saber-toothed tiger), it makes us inherently poor estimators of the real world. That is, we naturally use what could be termed the Human Nature Method (HNM) rather than the scientific method.

Cherry-Picking

We all have our own particular view of the world and we garner support for that view by referring to 'evidence'. Unfortunately we are very much prone to selecting that evidence which supports our view, whilst rejecting the evidence which does not – cherry-picking. Proponents of particular ideologies select facts and figures which seem to confirm their ideas. Unless the listener has access to ALL the data it is difficult to detect cherry-picking. One can be reasonable suspicious when only a small amount of supporting data is supplied, especially when the topic is complex.

The bad news is that cherry-picking can occur on both input data and the presentation of results. The good news is that it can largely be avoided by adhering to the scientific method.

An Example

The following example is drawn from real data and outlines the effect of NOT adhering to scientific methodology.

Detection of the Vela Pulsar

Currently I am involving in a personal quest to build a system which can detect the signals from the Vela Pulsar. As part of that endeavour I have developed hardware and software customised for the purpose. During testing activities I have collected daily data runs of 4 hours duration for 28 days (when Vela is transiting). This principally was to do an RFI survey at the selected observational frequency of 400 MHz as well as as a stress test of the software. The antenna used was an existing 6M (50 MHz) dipole at a height of 3m. The LNA and BPF were located at the receiver indoors and there was a 20m run of RG-58 coaxial cable between the antenna and the first LNA.

Daily Analysis

While calculations predict that the Vela Pulsar signal should be more than 10 dB below the noise level, examination of daily data showed that on some days there *appeared* to be a signal right on the predicted frequency of the pulsar.

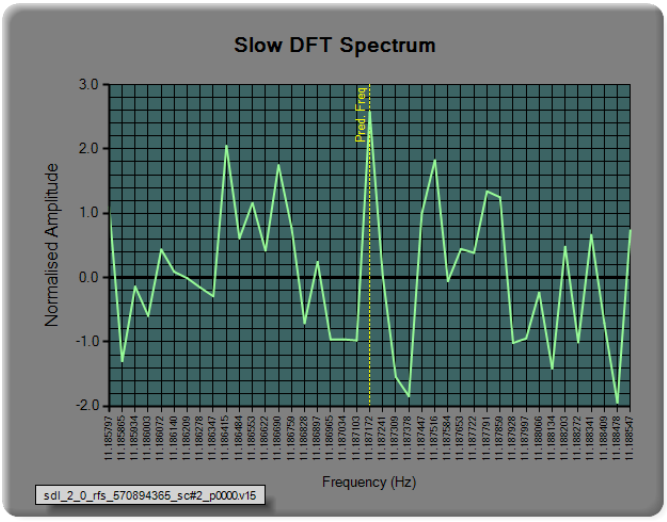

Figure 1: Example Daily Analysis

Out of the 28 daily data runs 9 days show a peak of varying amplitudes right on the predicted Vela Pulsar frequency for that day. Applying the HNM (Human Nature Method) principle it could be surmised that the days which do not show a significant peak at the predicted Velar Pulsar frequency could have been subject to cross-polarisation losses or RFI and so could be eliminated from the analysis.

The analysis software has the ability to compensate for the drift in topocentric frequency from day to day ('diurnal tracking') and analysing and summing the 9 daily data files with a peak > 1 standard deviation gives the result shown in Figure 2.

The sigma level for the peak at the predicted Vela Pulsar frequency is 5.7 which specifies that the probability of this being random is 1 in 26 million.

OK – we know we have eliminated 'inconvenient' data, but we are reluctant to dismiss this result. Perhaps we can vindicate this result by making a prediction. As the result in Figure 2 was produced by summing topocentric corrected data, if we sum instead *without* correction we should see a smaller peak which also should be wider as the 'pulsar peak' drifts lower in frequency as each day passes. This result is shown in Figure 3.

Here we can see that the peak *has* broadened and dropped in amplitude and also the peak has drifted lower in frequency as predicted. As the days are summed we would expect the peak now to be averaged in frequency by an amount of half the end to end drift. The drift from the first day to the last day is calculated to be ~ -24ppm. The peak in Figure 3 is offset by half that amount at ~ -12ppm.

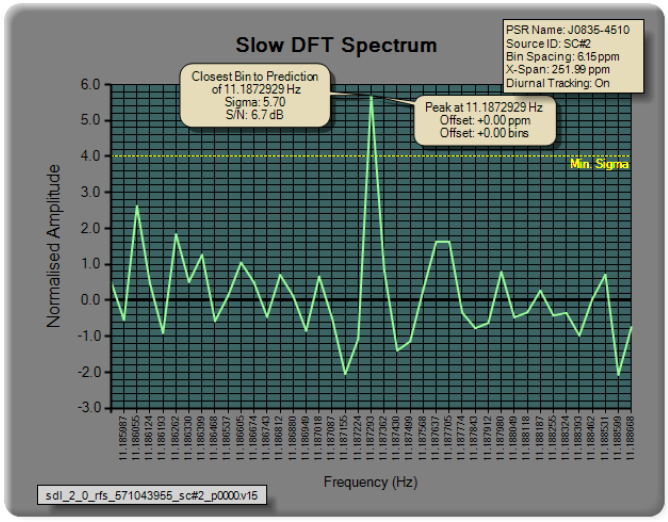

Figure 2: Incoherent Summing of Daily Data with SD > 1

Perhaps by now we are tempted to overlook the elimination of the 'bad data' days. Perhaps we can display the data in a more challenging way to support the result in Figure 2.

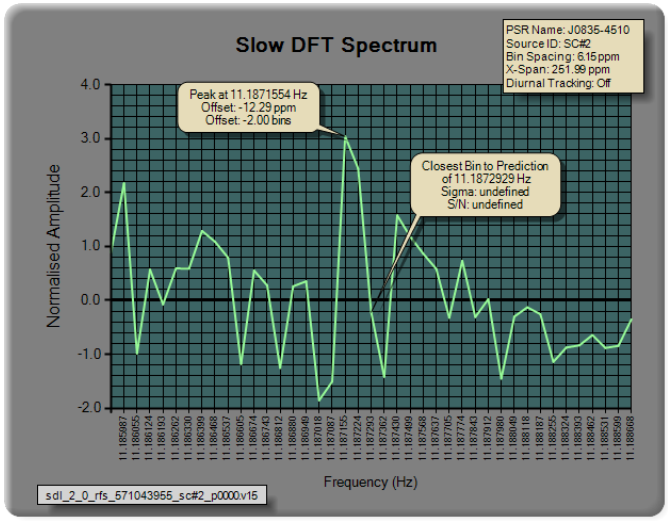

Figure 3: Incoherent Summing *Without* Diurnal Tracking

Let's widen the display to include more frequencies as shown in Figure 4.

Still we have a peak at the predicted frequency, but now sigma has dropped to 4.27. This is still a probability of being mere chance of 1 in 150,000. On face value fairly convincing.

Once again we repeat this with diurnal tracking turned off as shown in Figure 5. The peak at the predicted frequency has disappeared into the noise as expected.

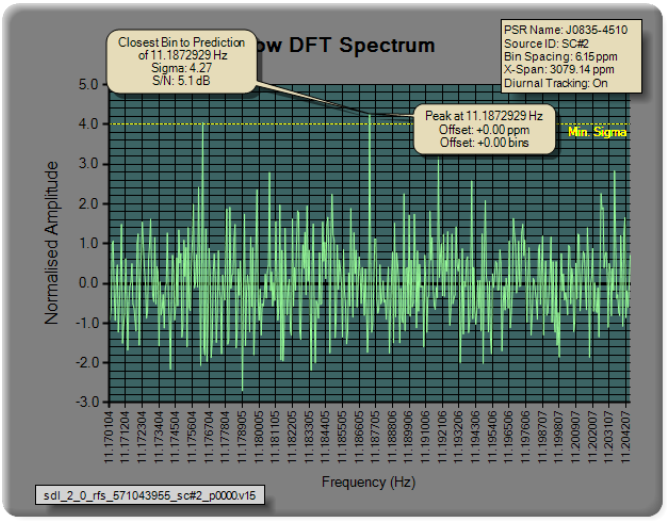

Figure 4: Wider Display

Comment

So we are left with a result where there appears to be a signal which behaves like a Vela Pulsar signal (tracks exactly as the expected daily drift in topocentric frequency) which holds over a time span of 28 days.

The question is:

Is this a valid result ?

The answer is:

NO !!!

Despite looking like a very convincing detection of the Vela Pulsar signal it is *invalid.*

Why ? Because it violates a basic principle of the scientific method which states data cannot be excluded on the basis that it makes the data look bad. Of course we can eliminate data which is obviously corrupted *before* analysis, but not *after* doing analysis which identifies data which is 'inconvenient'.

The only valid result for this whole exercise is shown in Figure 6 which is the result of incoherently summed ALL 28 days of data. This shows that a peak can be seen in the data at the predicted frequency, but at a level which is statistically insignificant. An 'encouraging' result – but not a valid Vela Pulsar signal detection.

Conclusions

There are good reasons for the development of the scientific method. It the best method for establishing the probability that a result is valid or otherwise.

It is not sufficient to simply present data as shown in Figure 2 which looks very convincing. The principle of eliminating data simply because it makes the result look bad violates a basic principle of the scientific method.

Another principle is 'peer review'. Results which are protected from peer review must be viewed as invalid. Details of how the data was captured and how, if any, data exclusion was done must be presented. The power of this technique is that the review is being done by other parties who are not emotionally attached to the result.

Leave the 'cherry-picking' to the politicians...

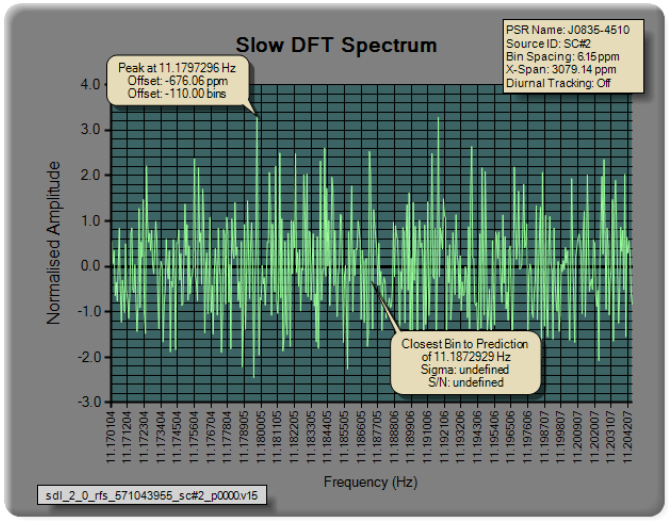

Figure 5: Wider Display *Without* Diurnal Tracking

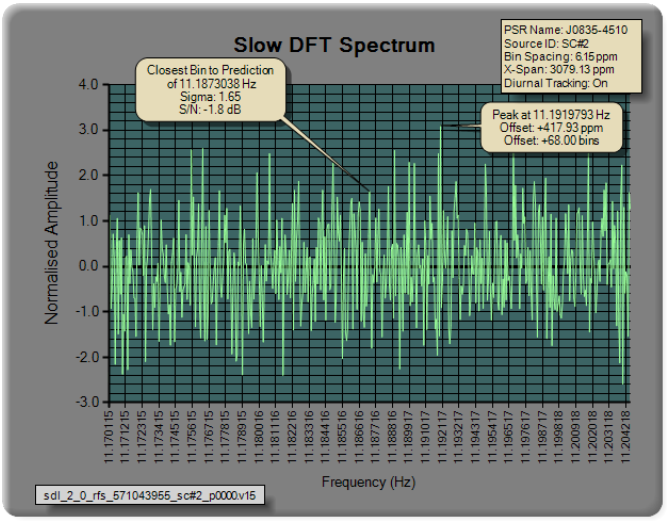

Figure 6: All 28 Days Summed Without 'Cherry-Picking'