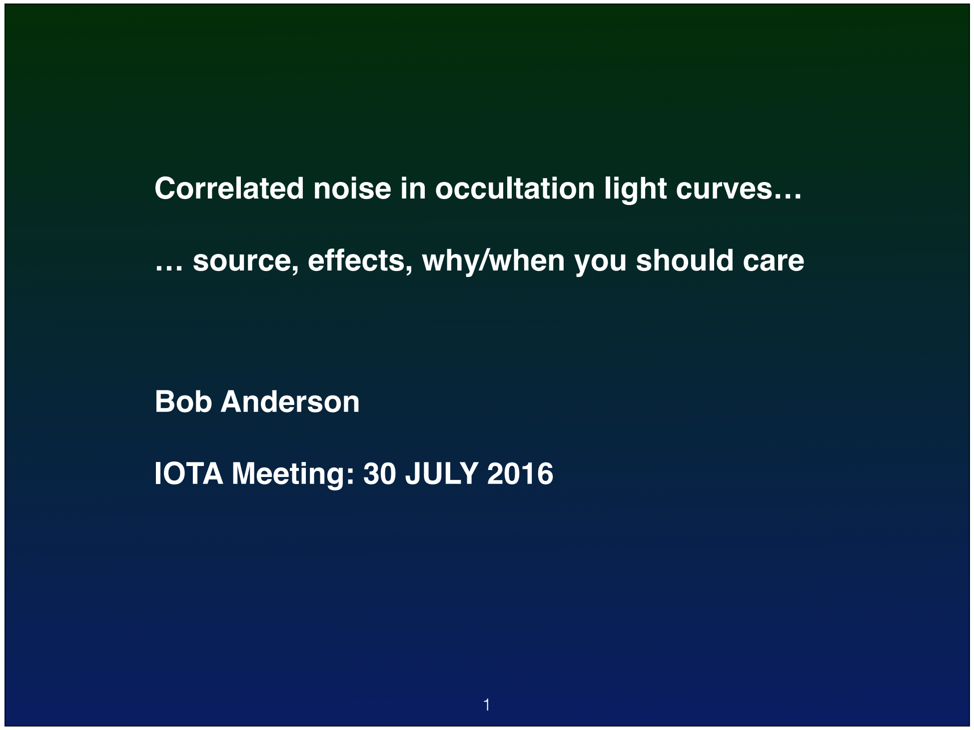
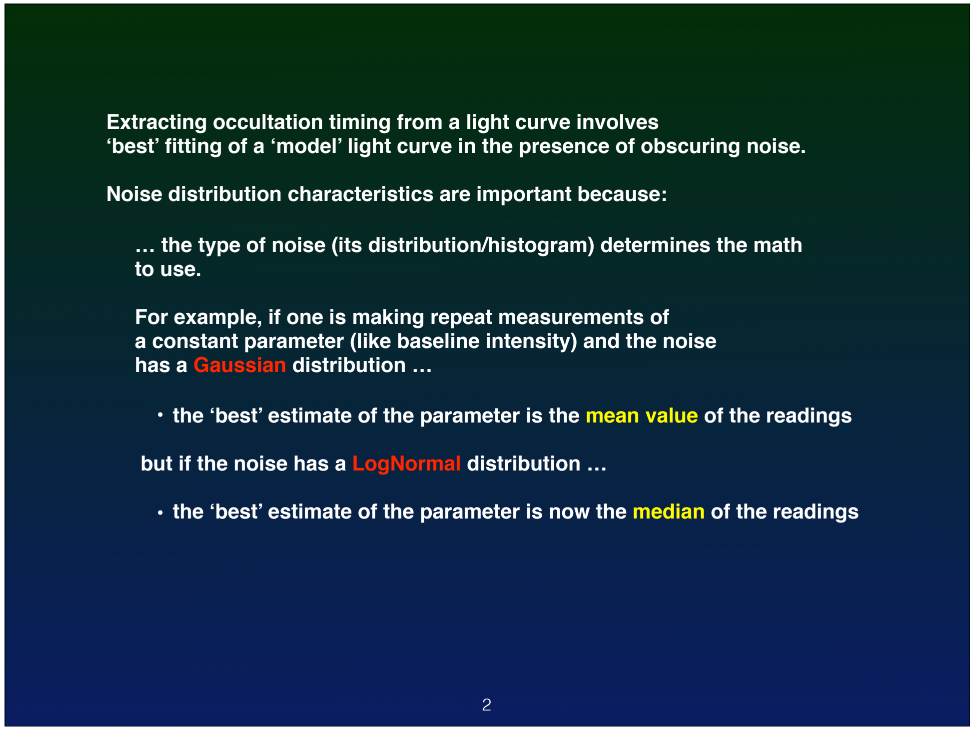


Good morning — I'm Bob Anderson. That having been said, I'm going to turn off my camera so that I can focus on the slides and my presentation notes.

TURN OFF CAMERA

I'll be giving a short talk about the effects of a special kind of noise that often appears in occultation light curves — temporally correlated noise.

NEXT



I'm the programmer behind R-OTE, so I have involved myself for a long time in the problems associated with extracting so-called square wave occultation timings from light curves — occultations where the star can be treated as a point source and diffraction effects are too fast to be captured by cameras operating at 30 or 25 frames per second..

PARAPHRASE SLIDE

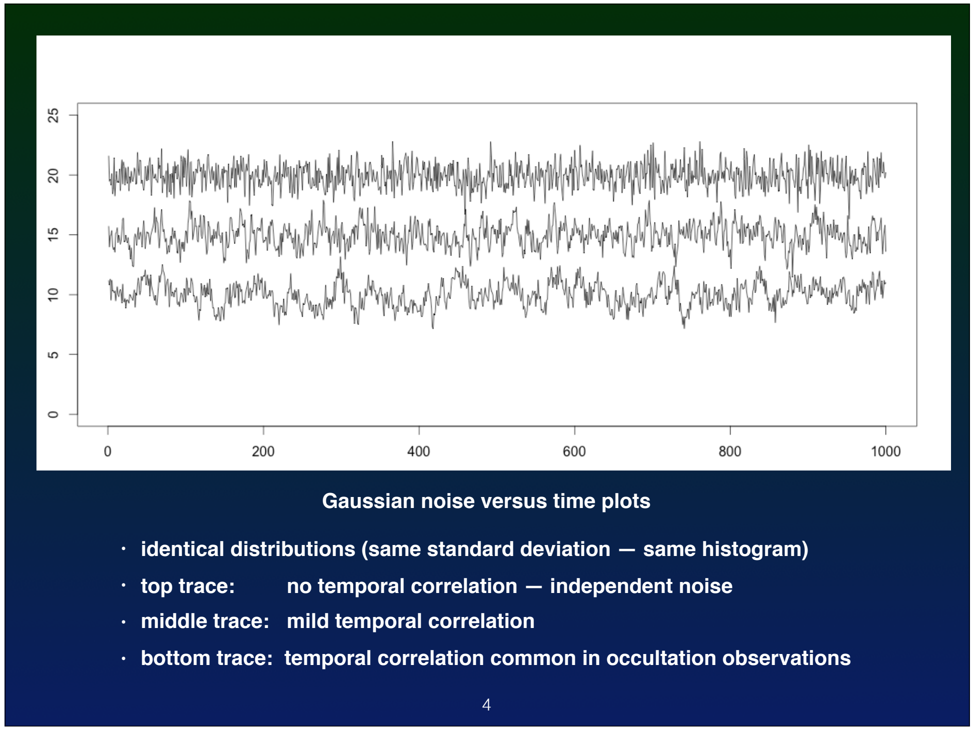
As it turns out, in spite of the fact that occultation light curve noise sources have a variety of distributions — Poisson, Gaussian, and LogNormal —in my experience to date, the net effect is always indistinguishable from Gaussian — and I have looked hard for deviations from Gaussian. But the noise distribution (its histogram) is only part of the story. The temporal characteristics of the noise must also be taken into consideration.

NEXT

READ SLIDE

And that is what this talk is about — the next slide shows graphically the effect to be discussed.

NEXT

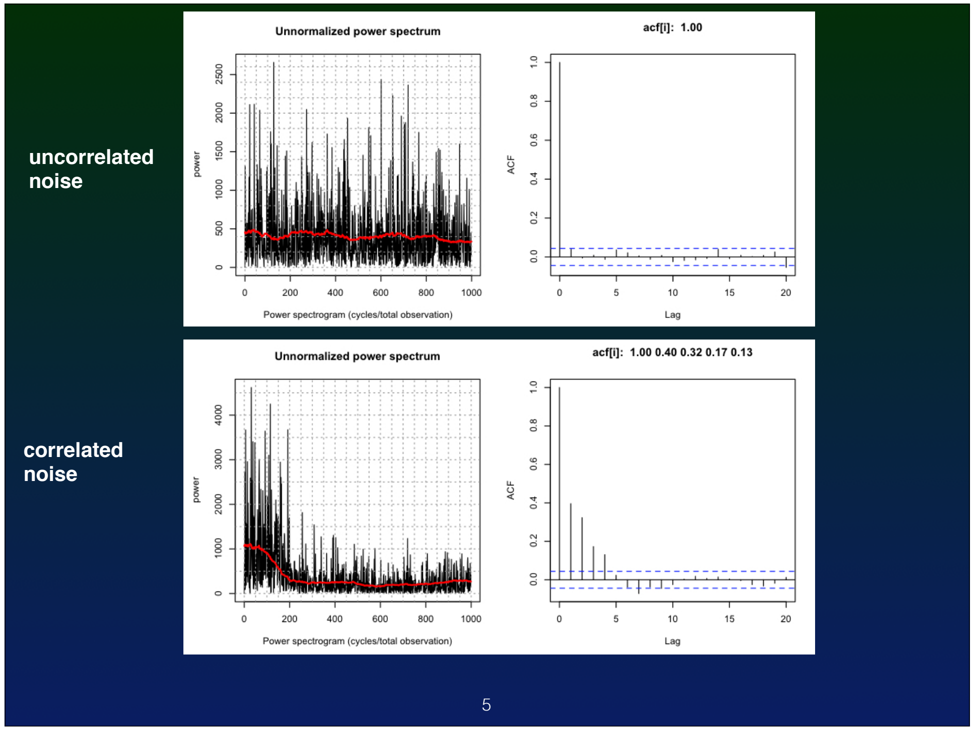


Here we see three traces that are visually different — but they have exactly the same distribution histogram — they are all examples of Gaussian noise with the same standard deviation— what sets them apart are their temporal characteristics — which means the degree to which reading n reflects what has happened to prior readings.

In statistics, one commonly encounters the term i.i.d (independent and identically distributed).
And many statistical analysis procedures require or assume the i.i.d. condition.

The top trace satisfies i.i.d — the other two traces do not. The observation values are identically distributed — they are drawn from a Gaussian distribution. But they are NOT independent — reading n depends on previous readings to some degree.

NEXT

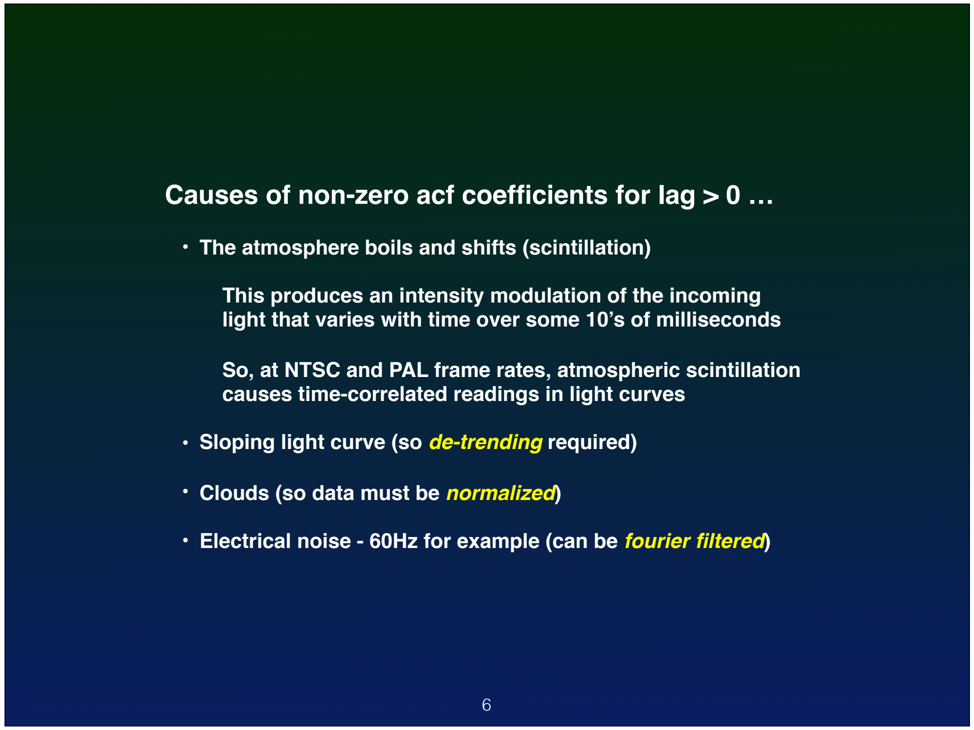


The primary tool for detecting and measuring the temporal characteristics of noise is the auto-correlation function.

What that function does is compare the noise with a copy that is shifted in time by one or more frame times and computes a correlation coefficient between the two at that time shift. The amount of the time shift is referred to as ‘lag’. At a lag of zero, the noise is being compared with itself, so the correlation coefficient is equal to unity. What is of more interest are the coefficients for lag greater than zero.

These plots are excerpts from the noise analysis panel in R-OTE. The right-hand plot in each panel is a plot of the auto-correlation coefficients out to lag=20. At the top of each panel is a list of the acf coefficients that meet a test for statistical significance.

NEXT



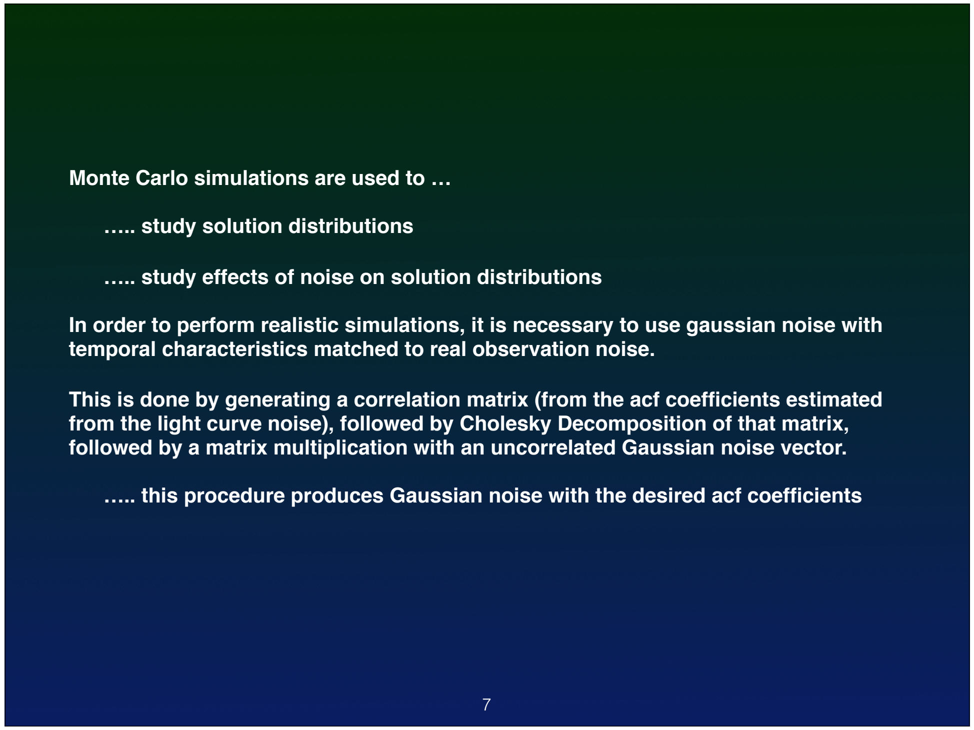
It’s worth spending a little time talking about what can cause the noise auto-correlation function to have statistically significant coefficients for lags greater than 0.

READ SLIDE

R-OTE provides for the determination of noise auto-correlation coefficients and then uses the values found during the estimation of error bars.

We are primarily interested in dealing with the effects of scintillation noise, so other observational conditions that can produce non-zero auto-correlation coefficients must be dealt with first. R-OTE provides tools for dealing with these.

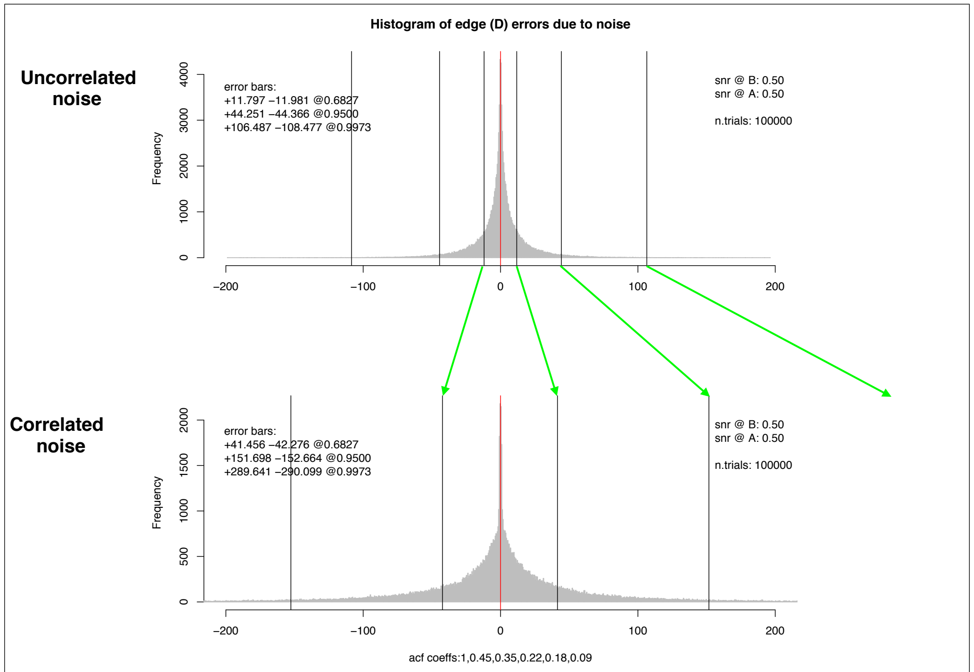
NEXT



Occultation observations are, by their very nature, ‘one-off’ — we can’t repeat an observation many times, so there is no direct way to study the statistical properties of our D and R time extractions.

The work-around that is commonly used is Monte Carlo simulation. It proceeds by taking a noise-free model light curve with the D, R, B and A parameters extracted from the original light curve, adding realistic noise, and re-finding the D and R values. This procedure is repeated thousands of times. Then, the results are summarized in a histogram that shows how many time a particular solution occurred.

NEXT



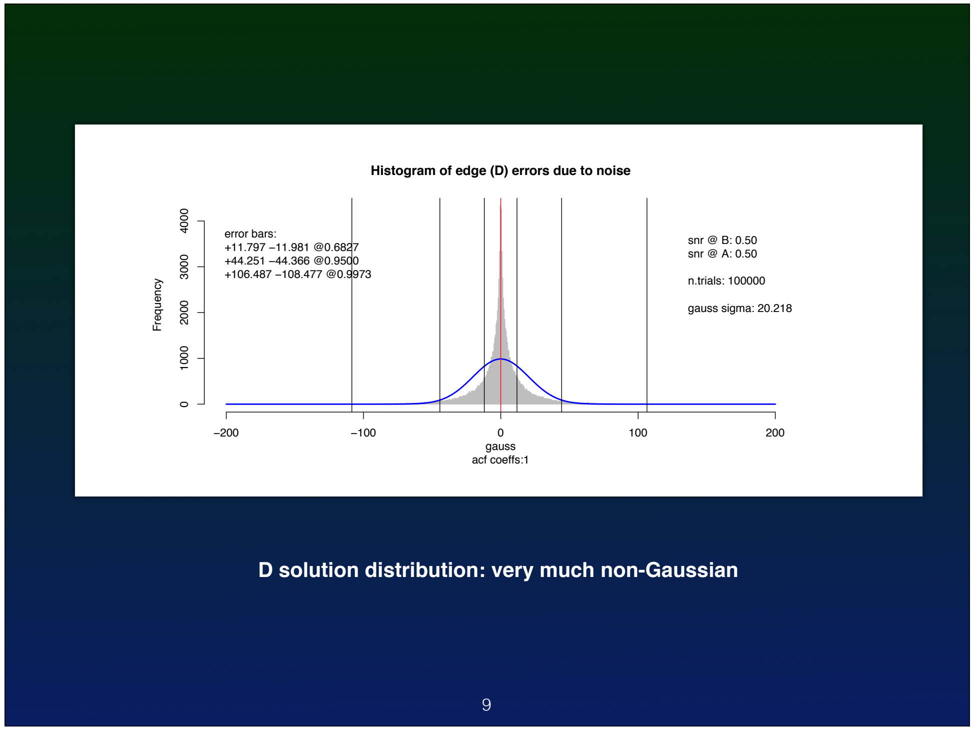
Correlated versus uncorrelated noise: effect on solution distributions (note vertical scale change)

This is from an R-OTE error bar estimation panel. The top plot gives shows the solution histogram when the observation noise has no temporal correlations — auto-correlation coefficients at lag=1 and greater are effectively zero. The lower plot gives the solution histogram with the same amount of noise — same standard deviation — but with significant auto-correlation coefficients at lag=1 and higher — in this case, all the way out to lag=6.

I should mention that R-OTE uses log likelihood for its 'best fit' metric. Using this metric, the 'best fit' is the set of parameters that maximizes the likelihood calculation. Shortly I will discuss other commonly used metrics. Meanwhile, I just want to point out a few things on this slide:

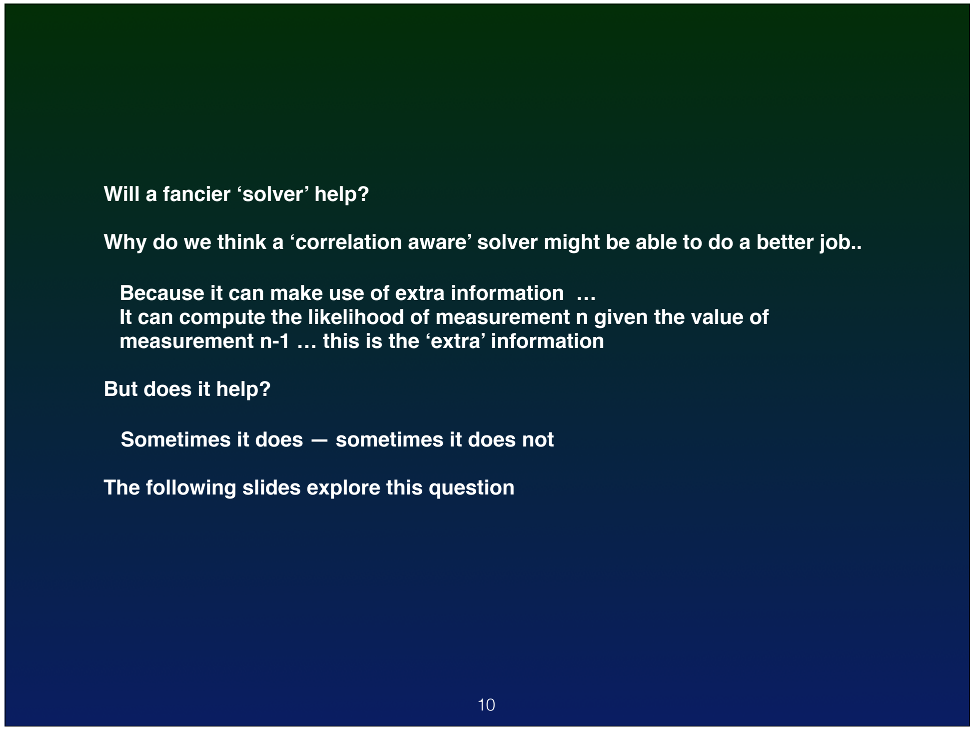
POINT OUT ERROR BAR LINES

The error bar expansion shows how correlated noise in an observation increases the uncertainty in D time extractions.
POINT OUT SCALE CHANGE
NEXT



We are more used to measurements that have a Gaussian distribution of ‘solutions’. That is obviously **<u>not</u>** the case for the distribution of occultation edge positions — as the present slide demonstrates. Not to belabor the point too much, but the blue curve on this plot is the best fit of a Gaussian distribution to the solution histogram of 100,000 Monte Carlo trials. Hopefully this also makes it clear that, for occultation edge position error bars, it would be extremely misleading to refer to them as 1 sigma, or 2 sigma, or 3 sigma — such terminology should be used only when the solution distribution is Gaussian.

NEXT



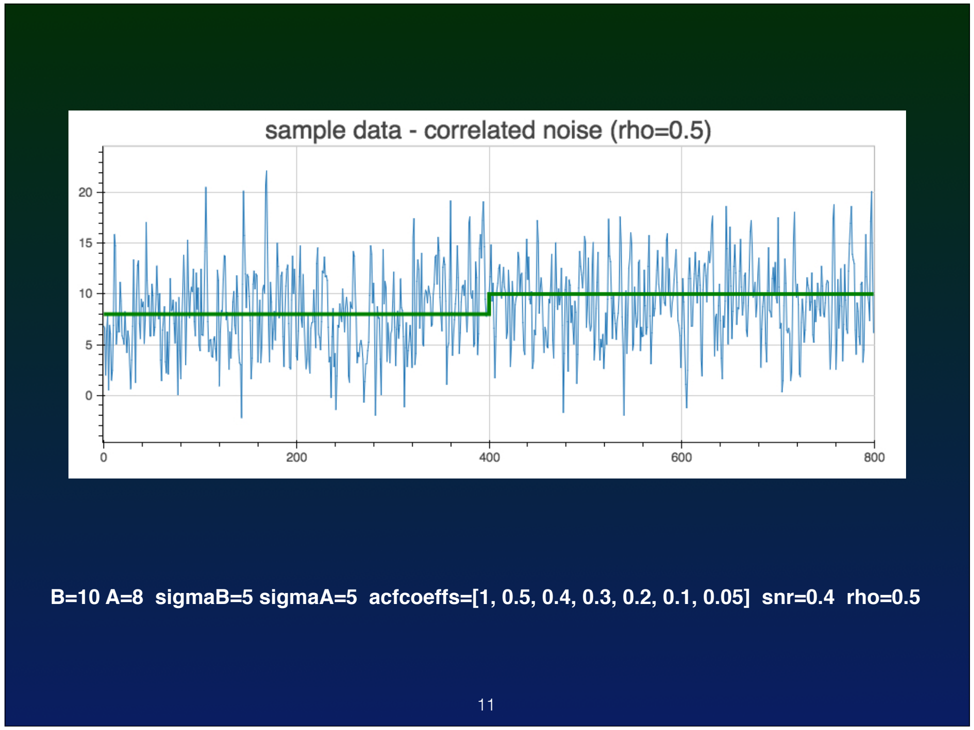
The standard metrics in common usage for curve fitting— like ordinary least squares for example — assume that the noise is i.i.d — that is, independent and identically distributed.

The i.i.d requirement can of course be ignored — the solvers will still work. But the obvious question is posed by this slide…

READ THE SLIDE

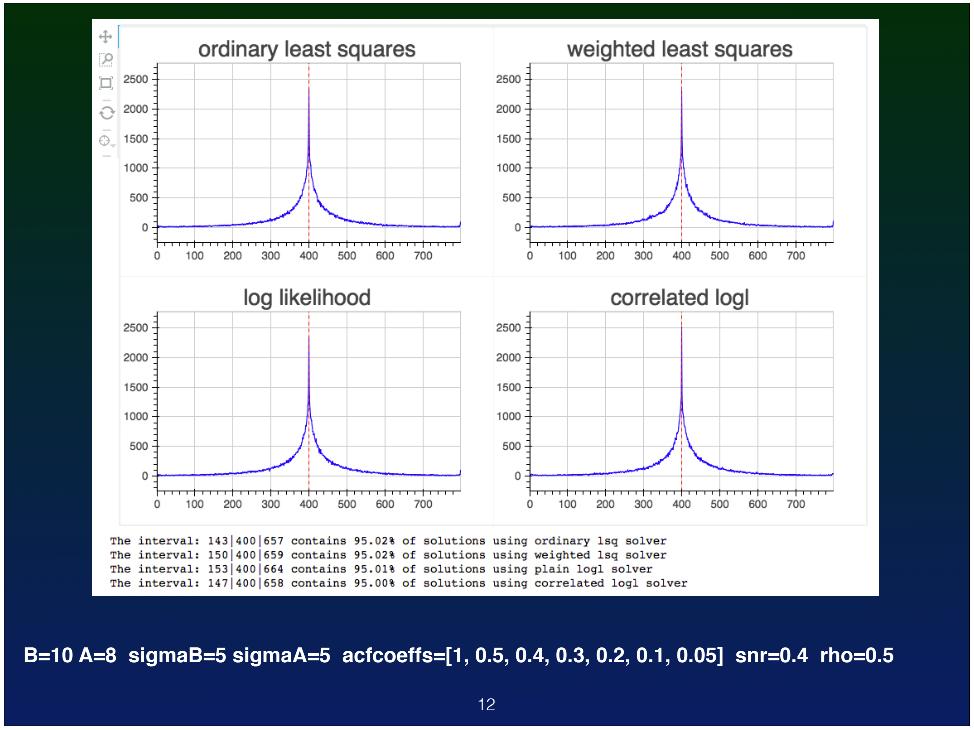
To shed some light on this question, I set up a series of Monte Carlo simulations that employed different solvers — that is, different metrics of 'best fit'. The next slide shows an example of one test case.

NEXT



Here we see a typical test case that is ‘solved’ for D as part of the Monte Carlo simulation. The rho=0.5 is the value of the acf coefficient at lag 1 — it reflects the dependence of reading n on reading n-1. The importance of this number is that it is the input to a special ‘correlation aware’ solver that was developed for this test. We”ll see this in the next slide.

NEXT



Here are the simulation results using four different 'solvers' — each 'solver' uses a different calculation of the metric used to determine the 'best fit'.

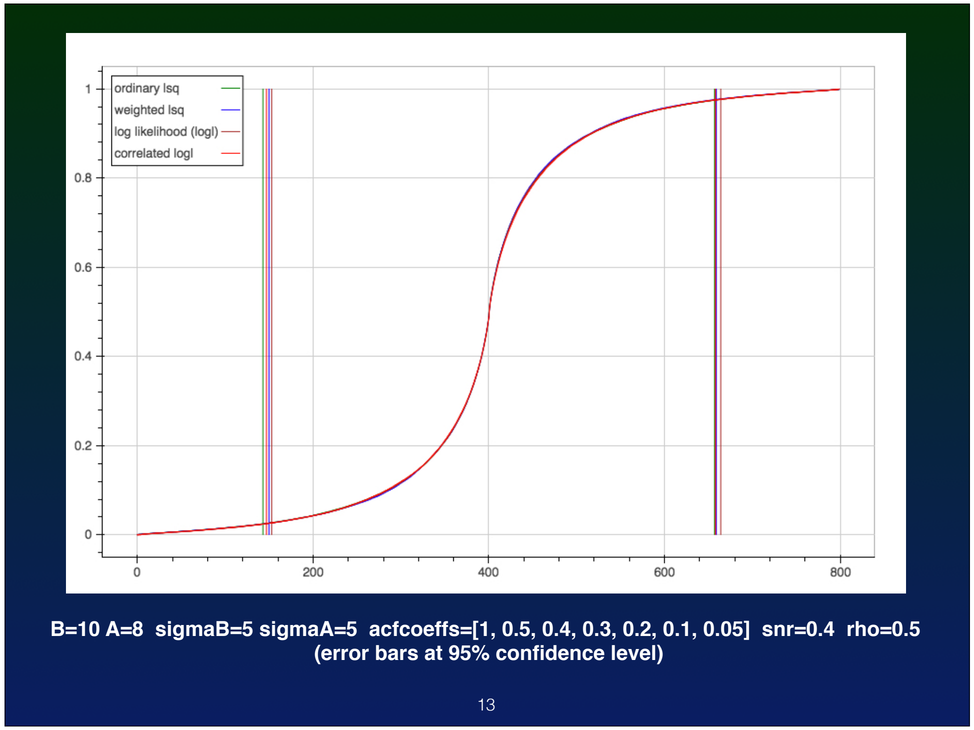
The metric for the ordinary least squares 'solver' is the usual sum of the squares of the deviations — a metric that is minimized to achieve 'best fit'.

The weighted least squares 'solver' uses a metric that also sums the squares of the deviations, but weights each point by the noise level in that region — this metric is minimized to achieve 'best fit'.

The log likelihood 'solver' is the one used in R-OTE and achieves 'best fit' by maximizing this metric.

The correlated logl 'solver' includes a likelihood calculation based on conditional probability, that is, it calculates the likelihood of reading n conditioned by the value of reading n-1 — this is the only 'correlated noise aware' solver in the experiment and it **only** takes into account the auto-correlation coefficient at lag=1.

It's hard to see much difference in these plots. It's easier to see distribution differences by looking at the cumulative probability plots. These are shown on the next slide. NEXT

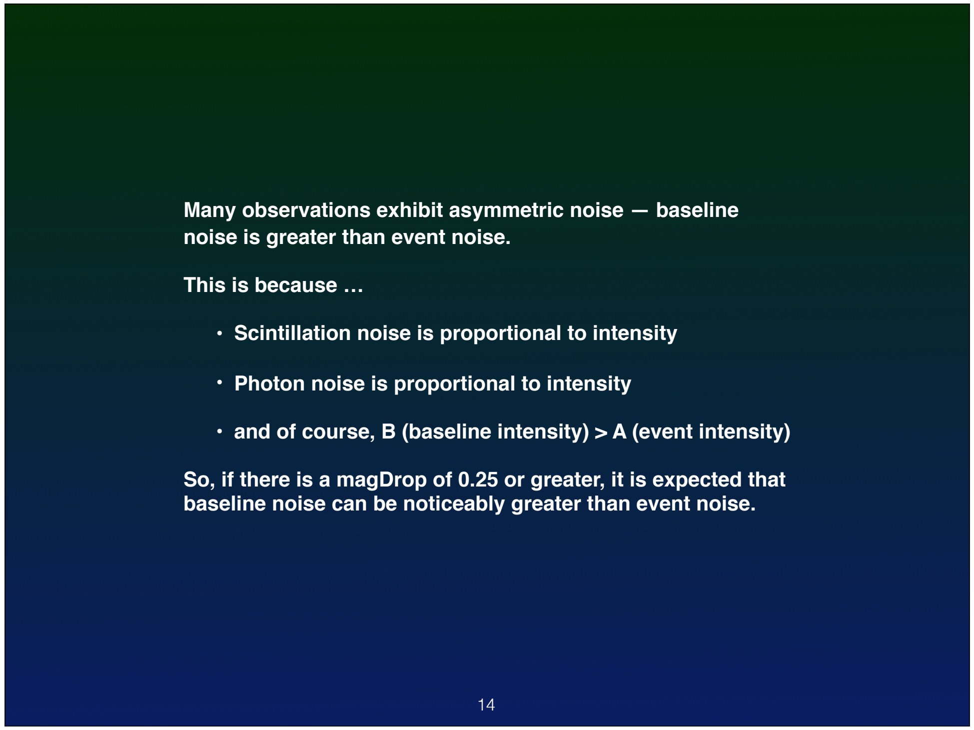


Clearly, there isn't much difference between the 'solvers' — they all produce essentially identical solution histograms — it's likely that if a larger number of trials were employed, the agreement would improve.

I want to point out that the baseline noise sigma (sigmaB=5) is equal to the event noise sigma (sigmaA=5) — I emphasize this because the situation when the baseline and event noise are NOT equal is rather different.

The error bars shown here are at the 95% confidence level. I conclude from this plot that any of the 'solvers' can be used when the noise at A is the same as the noise at B. The fact that the i.i.d requirement is not satisfied has no effect on the solutions obtained.

NEXT

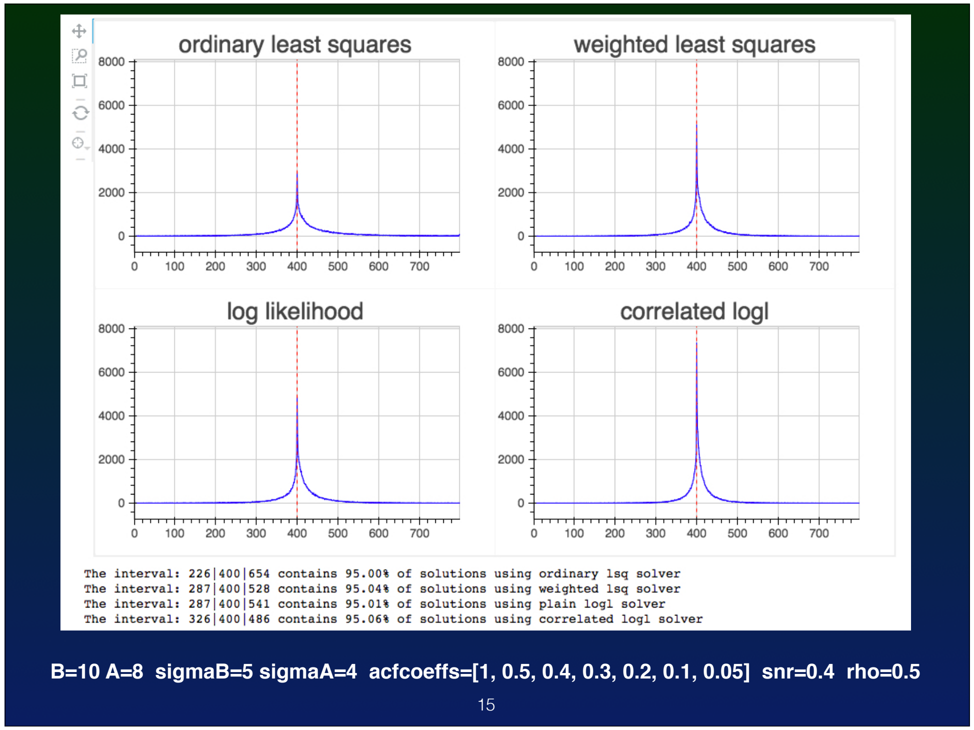


In the previous study, baseline and event noise where selected to be equal. However, it is often the case that there is a difference in baseline noise levels and event noise levels as the bullet points on this slide explain.

REVIEW THE SLIDE BULLET POINTS

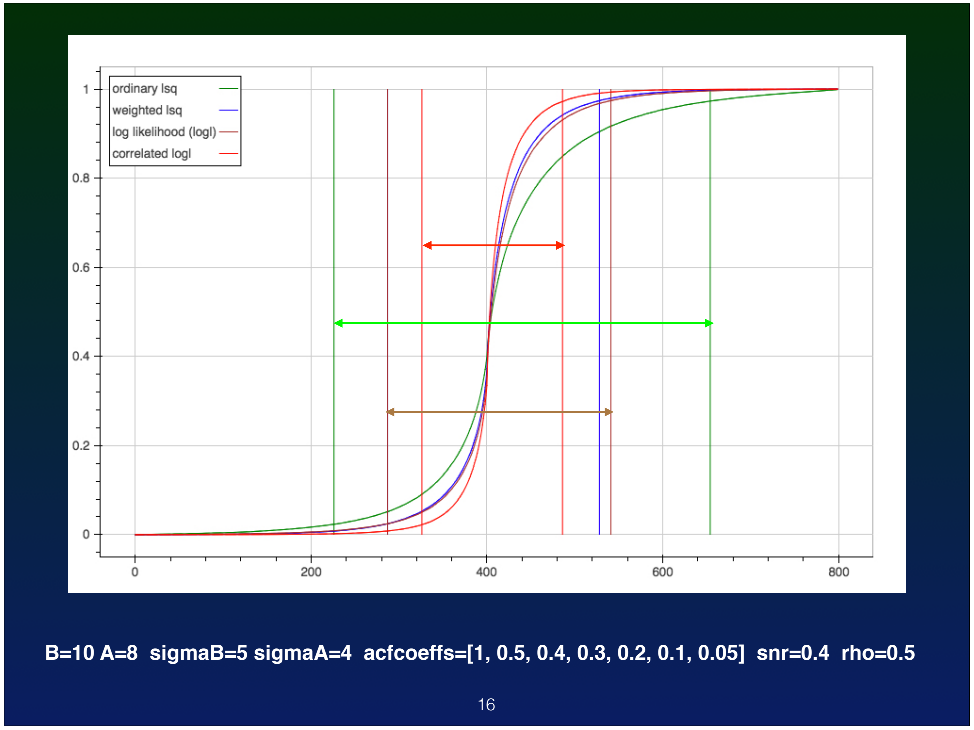
When this is the case, the choice of ‘solver’ becomes important. This is shown in the next slide.

NEXT



Now there **<u>are</u>** visually apparent differences in the solution distributions. The correlated logl solver (the one whose metric takes into account that the value at frame n is influenced by the value at frame n-1) is three times more likely to 'nail it' than the ordinary least squares solver. It's a little less obvious that the distributions are narrower, but the cumulative probability plots are easier to use for such comparisons.

NEXT



And now we can see that under some conditions, the choice of ‘solver’ becomes an issue to be considered.

The error bars are shown here at the 95% confidence level for all four ‘solvers’.

The red arrow points out the error bars achieved with the ‘correlation aware’ solver — it has tighter error bars.

The green arrow points out the error bars resulting from standard least squares based ‘solver’

The brown arrow shows the error bars that are reported by the R-OTE log likelihood ‘solver’ — that particular ‘solver’ is oblivious to the temporal characteristics of the noise but does deal with asymmetric noise.

NEXT

Do we need a ‘correlation aware solver’ for occultation timing extraction?

The previous slide suggests that narrower solution histograms can result from the use of a metric that takes into account temporal correlation.

But the conditions were rather special — to see significant differences, a long duration, noisy observation with detectable noise asymmetry is required.

In any case, none of the ‘solvers’ introduce a bias — the most frequent solution (the peak of the histogram) is always at zero.

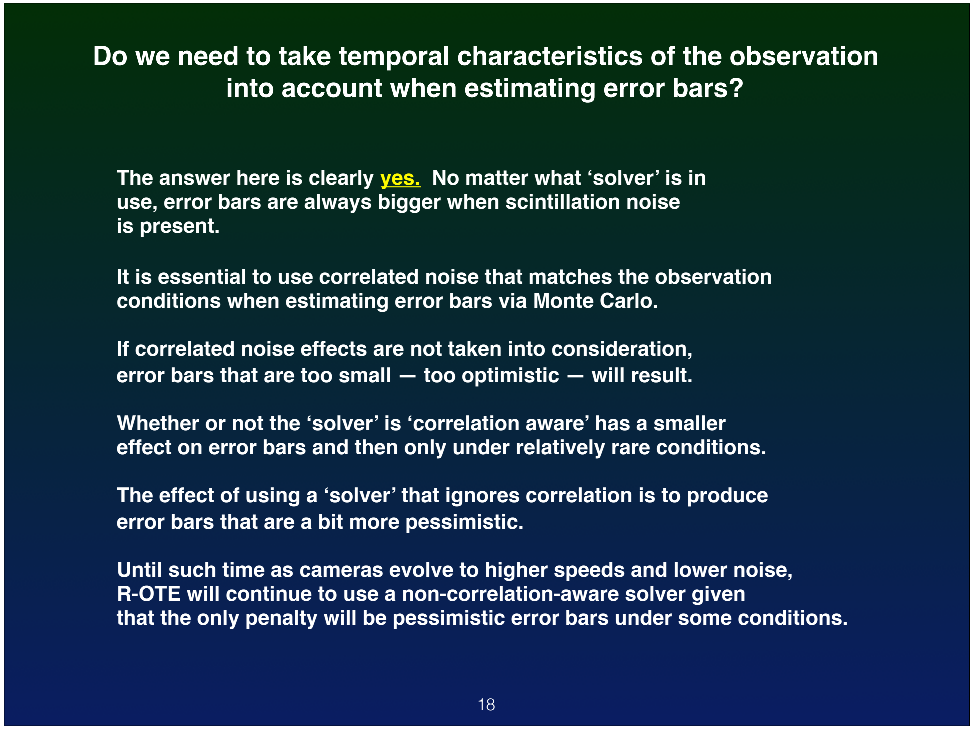
So a ‘correlation aware solver’ is not required.

17

Here I’m just going to read the slide.

READ THE SLIDE

NEXT



Again, I'm just going to read the slide.

READ THE SLIDE

And that concludes my talk.