Jason Byrne <jbyrne6@gmail.com>

Simulation of data

7 messages

David Long <long.daithi@gmail.com> **16 February 2009 19:33**
To: Peter Gallagher <peter.gallagher@tcd.ie>, James McAteer <rtjmca@gmail.com>, Jason <jbyrne6@gmail.com>, Shaun Bloomfield <shaun.bloomfield@tcd.ie>

Hi everybody

I took the 20070516 data from both 195A and 171A and added noise to it (the same amount of noise to each data set for each iteration) and plotted it to see how the data and fit would vary. The movies are now online at

http://www.maths.tcd.ie/~dlong/simulations/

I need to add chi-squared values to the plots for each iteration, and I hope to do this tomorrow morning.

While (by eye) the fits to the distance appear to be good, the velocity varies too much with the noise, which means that the fits are not as good (again by eye) in as many cases.

We should talk tomorrow about it and also about what to do with the paper. Since I won't be here next week, we should probably be finished or as close as possible to finishing by Friday.

Thanks

Dave

--
David Long
Astrophysics Research Group
School of Physics
Trinity College
College Green
Dublin 2
www.maths.tcd.ie/~dlong

Jason Byrne <jbyrne6@gmail.com> **16 February 2009 20:12**
To: David Long <long.daithi@gmail.com>
Cc: Peter Gallagher <peter.gallagher@tcd.ie>, James McAteer <rtjmca@gmail.com>, Shaun Bloomfield <shaun.bloomfield@tcd.ie>

It looks like the noise is increasing with the data?

I think an important consideration is always the error bars in use in the first place. They should be sufficiently large enough to account for any noise or other error in the measurements, which would mean that if we add noise in this manner we should still be getting fits which range within our errors.

2009/2/16 David Long <long.daithi@gmail.com>

[Quoted text hidden]

--
Jason P. Byrne,
Astrophysics Research Group,
School of Physics,
Trinity College,
Dublin 2,
Ireland.
Tel: +353-(0)1-8962157
Mob: +353-(0)87-6325173
www.physics.tcd.ie/Astrophysics

Peter Gallagher <peter.gallagher@tcd.ie> **16 February 2009 21:07**
To: David Long <long.daithi@gmail.com>
Cc: James McAteer <rtjmca@gmail.com>, Jason <jbyrne6@gmail.com>, Shaun Bloomfield <shaun.bloomfield@tcd.ie>

Hi David - that's a good start, but I see you've set the noise to scale ~linearly with height. This has given your final datapoints too much variance. For example, the last datapoints in the 171 plots vary in the range 190-290, which I suspect is too large (~50%). Could you fix the noise to <10% for one movie and then <20% in the next? I think these are more reasonable limits.

Talk tomorrow.

Peter.

Peter T Gallagher PhD
Astrophysics Research Group
School of Physics
Trinity College Dublin
Dublin 2
Ireland
W: www.physics.tcd.ie/astrophysics
T: +353 (0)1 896 1300
F: +353 (0)1 671 1759
Skype: petertgallagher

[Quoted text hidden]

Peter Gallagher <peter.gallagher@tcd.ie> **16 February 2009 21:10**
To: Jason Byrne <jbyrne6@gmail.com>
Cc: David Long <long.daithi@gmail.com>, James McAteer <rtjmca@gmail.com>, Shaun Bloomfield <shaun.bloomfield@tcd.ie>

Just saw Jason's email after I sent my last one.

> I think an important consideration is always the error bars in use in the first place. They should be sufficiently large enough to account for any noise or other error in the measurements, which would mean that if we add noise in this manner we should still be getting fits which range within our errors.

Absolutely - only then can we be confident that our results are robust.

[Quoted text hidden]

David Long <long.daithi@gmail.com> **17 February 2009 11:10**
To: Peter Gallagher <peter.gallagher@tcd.ie>
Cc: Jason Byrne <jbyrne6@gmail.com>, James McAteer <rtjmca@gmail.com>, Shaun Bloomfield <shaun.bloomfield@tcd.ie>

Hi

Ok, I've fixed the problem with the linear increase in the noise of the data and made movies for 10% and 20% noise added to the data. They are online in the same place as previously:

http://www.maths.tcd.ie/~dlong/simulations/

I'm not sure how to get the chi-squared values to print as the legend, and I'm going to try and fix that.

Let me know what you think.

Thanks,

Dave

2009/2/16 Peter Gallagher <peter.gallagher@tcd.ie>

[Quoted text hidden]

[Quoted text hidden]

Peter Gallagher <peter.gallagher@tcd.ie> **17 February 2009 12:00**
To: David Long <long.daithi@gmail.com>
Cc: Jason Byrne <jbyrne6@gmail.com>, James McAteer <rtjmca@gmail.com>, Shaun Bloomfield <shaun.bloomfield@tcd.ie>

OK, they look good. What is your opinion on the fits, uncertainties, etc ?

Could you plot the fit parameters as a function of iteration for each simulation?
eg a_fit, v_0_fit and h_0_fit vs. iteration for <10% noise etc. I would like to see how sensitive the fits are in a quantitative manner. My impression is that the a=const fit is pretty robust in

http://www.maths.tcd.ie/~dlong/simulations/171_sim_noisy_data.html

I also like the way the uncertainties in position always overlap both fits in this simulation.

The 20% noise appears to be too large at first glance to me. There are a number of instances where the uncertainties do not lie on the fits, although that could just be because the uncertainties are too low.

From the scatter in the original data, 10% looks appropriate to me.

Which do you think is more appropriate, 10% or 20%? You can work our which is best by fitting a line to the data (h_fit = h0 + v0 * t) and then working out the distance of each datapoint from the fit (dh = abs(h_data - h_fit)). dh / h * 100 will then give you the percentage scatter for each datapoint, while average(dh / h * 100) will give you the mean scatter.

Peter.

P.S. Use

IDL> legend,'Chi-squared: ' + arr2str(chi, /trim), /bottom, /right

to display the chi-squared on the plots.

Peter T. Gallagher
Astrophysics Research Group
School of Physics
Trinity College Dublin
Dublin 2
Ireland
W: www.physics.tcd.ie/astrophysics
T: +353 (0)1 896 1300
F: +353 (0)1 671 1759
Skype: petertgallagher

[Quoted text hidden]

David Long <long.daithi@gmail.com> **17 February 2009 15:24**
To: Peter Gallagher <peter.gallagher@tcd.ie>
Cc: Jason Byrne <jbyrne6@gmail.com>, James McAteer <rtjmca@gmail.com>, Shaun Bloomfield <shaun.bloomfield@tcd.ie>

Hi

Attached 4 plots showing the variations of the different fit parameters with iterations for <10% noise. There is a lot of variation there, but usually about a mean value. I've also worked out the mean % scatter for the different passbands, fits and % noise, and these are as follows:

Mean % scatter (195A parabolic fit, 10% noise) = 7.75681
Mean % scatter (195A linear fit, 10% noise) = 7.79206
Mean % scatter (171A parabolic fit, 10% noise) = 14.2978
Mean % scatter (171A linear fit, 10% noise) = 14.3149

Mean % scatter (195A parabolic fit, 20% noise) = 14.2662
Mean % scatter (195A linear fit, 20% noise) = 14.3475
Mean % scatter (171A parabolic fit, 20% noise) = 26.9795
Mean % scatter (171A linear fit, 20% noise) = 26.2383

The scatter is less for the data with 10% noise as would be expected. However, you still cannot say anything about the velocity of the disturbances based on these fits and using deriv. In most cases also, both the linear and parabolic fits pass through the error bars for each data point.

I've also put the movies online in the same place - http://www.maths.tcd.ie/~dlong/simulations/

These movies have the chi-squared values added to give an indication of how good the fits are in each case.

Let me know what you think

Dave

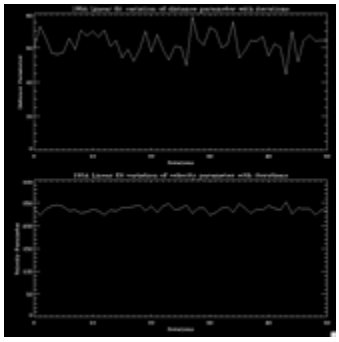
2009/2/17 Peter Gallagher <peter.gallagher@tcd.ie>

[Quoted text hidden]

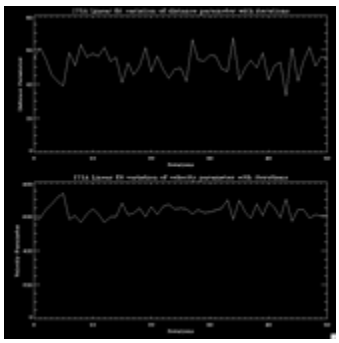
[Quoted text hidden]

---

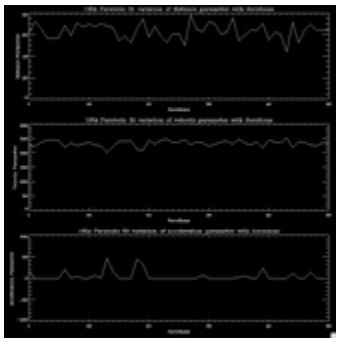
4 attachments



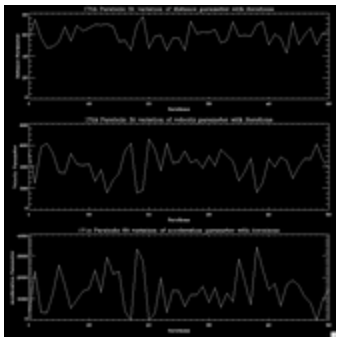
195_l_variations.png
16K



171_l_variations.png
17K



195_p_variations.png
19K



171_p_variations.png
22K