
Subject: Re: "Correct" Data Philosophy
Posted by [Giorgio](#) on Thu, 17 Dec 2009 22:31:12 GMT
[View Forum Message](#) <> [Reply to Message](#)

On Dec 17, 1:56 pm, David Fanning <n...@dfanning.com> wrote:

> Kenneth P. Bowman writes:

>> The problem of estimating values where you have no data is
>> very common and often very difficult. The best approach depends
>> on the character of the data, the size of the gaps, the methods used,
>> and the purpose of the analysis.

>

>> It is very important to not mislead yourself or your readers.
>> My first recommendation is **not** to fill gaps whenever possible --
>> instead, adapt your analysis and display methods to the data.
>> If you are displaying an image or contour, for example, show
>> the viewer where the data is missing with a special color
>> and don't display contours where there is no data.

>

>> If I am plotting global maps of 5 deg x 5 deg data, it should
>> look chunky (pixelated), not smooth. That reminds the viewer
>> what the actual resolution of the data is.

>

>> If you need to do a Fourier transform, consider using
>> least-squares estimation rather than interpolating
>> and using an FFT.

>

>> If the data is smooth and the gaps are small, interpolation
>> will probably work well. If the data is noisy and the gaps are
>> large, it is possible that nothing will work well.

>

>> If you do fill gaps, always test the impact on your results.
>> Does it matter whether you use linear or cubic interpolation,
>> for example?

>

>> In the end, you need to be confident that your results do not
>> depend significantly on how you chose to estimate the missing
>> data.

>

> OK, here is my problem: I don't have any idea what you
> people are talking about. And neither do the folks asking
> me questions. :-(

>

> This, in particular, is opaque to me:

>

> If you need to do a Fourier transform, consider using
> least-squares estimation rather than interpolating
> and using an FFT.

>

> OK, I will, but *how*!?

>

>> Is it similar to "interpolation" or "approximation" or "estimation"?

>

> Yeah, it's similar to all of those, I guess. But, how

> would you do it in IDL?

>

>> How about linear/bilinear/trilinear interpolation? Or minimum

>> curvature surface or thin-plate-spline? It also depends on how many

>> values are available and/or missing. There are other fitting/

>> interpolation functions too.

>

> Does IDL even *do* these things!? Or do I have to go learn

> Matlab?

>

> I guess I was hoping for a couple of examples. I really don't

> have the time or energy to open up a whole new research area

> here, although I can see that it might occupy my time quite

> fruitfully for a number of years. :-(

>

> Cheers,

>

> David

>

> --

> David Fanning, Ph.D.

> Fanning Software Consulting, Inc.

> Coyote's Guide to IDL Programming:<http://www.dfanning.com/>

> Sepore ma de ni thui. ("Perhaps thou speakest truth.")

My 2second thought:

I think it depends of the case. One example I can imagine is the removal of hot pixels from a CCD camera. Since you know that systematically your CCD camera is giving you a hot pixel at the same position, you can estimate its value from its nearest neighbours. However if you are not sure, its value can have a more profound meaning than your instrument has a different response function for that point. You could be missing something then.

I agree with Kenneth, you must always present the raw data and then the data treated so people can judge the difference. Or at least give the option about it.
