ELECTRONOTES

1016 Hanshaw Road

Ithaca, NY 14850

APPLICATION NOTE NO. 384

September 1, 2012

FUN WITH RED NOISE

Noise has color! Well, we have all heard of white noise. That is an audible signal with a flat spectrum, and likely deserves to be called white for that reason. There is also so-called pink noise often used in audio testing. It rolls off at 3db/octave, and is not all that easy to generate (it is approximated by a series of filterings of white noise). You can also find blue noise, a 3db roll-up. A grey noise is a white noise adjusted to partially compensate for the human hearing curve (which rolls-off on both high and low ends). And some people market recordings of background natural sounds such as ocean surf or wind through tree leaves as “green” noise. Then there is red noise, a 6db/octave roll-off. It is thus a sound with emphasized lower frequencies - not a bad analogy with red light, and we can better appreciate pink noise as an intermediate between white and red. Red noise is also called “brown” noise – but not because of the color. Rather it is named after Robert Brown who is famous for “Brownian Motion”. The connection between red (brown) noise and Brownian motion relates to the “random walk” nature of both.

The random-walk or “drunkard’s walk” is integrated white noise. It has memory. The next value of a red sequence begins with the current sample value, and then adds or subtracts an uncorrelated white sample to it, and so on. It is very similar to producing a signal by flipping a coin and totaling. If it’s heads, for example, you can increases the sequence by one, and if it’s tails, you subtract one. This means the resulting sequence is correlated. That is, the next sample can’t be very far from the current one. This random walk has some unexpected (at least counter-intuitive) properties. For example, we might expect it to hover around zero, with frequent crossings of zero. On the contrary, it often wonders far from zero for long stretches. Zero-crossings are rare. Yet, when we consider that it is composed of all frequencies, and that the lower frequencies are of larger amplitude (the 6db/octave roll-off), it is not surprising that no matter how far away, it will eventually come back. After all, you got far away, and the same sort of “luck” can just as well get you back. Put another way, in the scenario of a drunkard making random changes of direction in an attempt to get home, he will always get home eventually. Of importance later however: note that when we terminate our generation, there is generally a very good chance we will be a good distance from zero.

So we can generate red noise easily, digitally, by integrating a white noise sequence (a random number generator). All we need to do in add up the sequence. Fig. 1 shows the absurdly simple digital filter integrator that we will use here. So it is very similar to adding up heads and tails, except here there is no requirement that the white noise is binary (heads/tails being binary). We generally would use a uniform or Gaussian distribution for the input. So a white noise algorithm and a few lines of code do what we need.

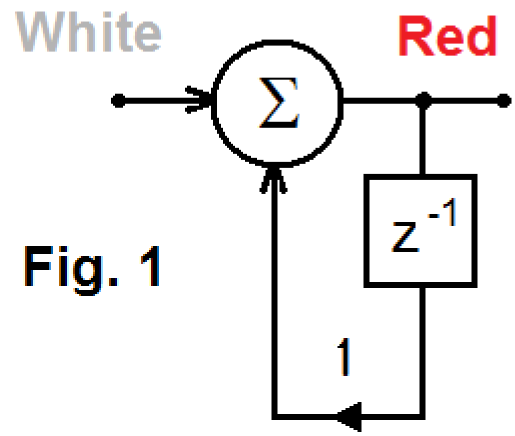


Fig. 1

```
% Red Noise Test   redtest.m
% m signals
for m=1:10000
 sw=2*(rand(1,1020)-0.5);  % white, uniform
 g=1  % values less than 1, like
    % 0.95  are sometimes used.
    % red filter - pole at z=+g,
    %     (integrator)
   sr(1)=sw(5);
     for k=2:1000
         sr(k)=sr(k-1)+g*sw(k+4);
     end

     %  FFT analysis and sum
     SR=abs(fft(sr));
     SR=SR(1:501);
     S=S+SR;
end

figure(1)
plot([0:500],S)
figure(2)
loglog([0:500],S)
grid
axis('square')
figure(2)
```

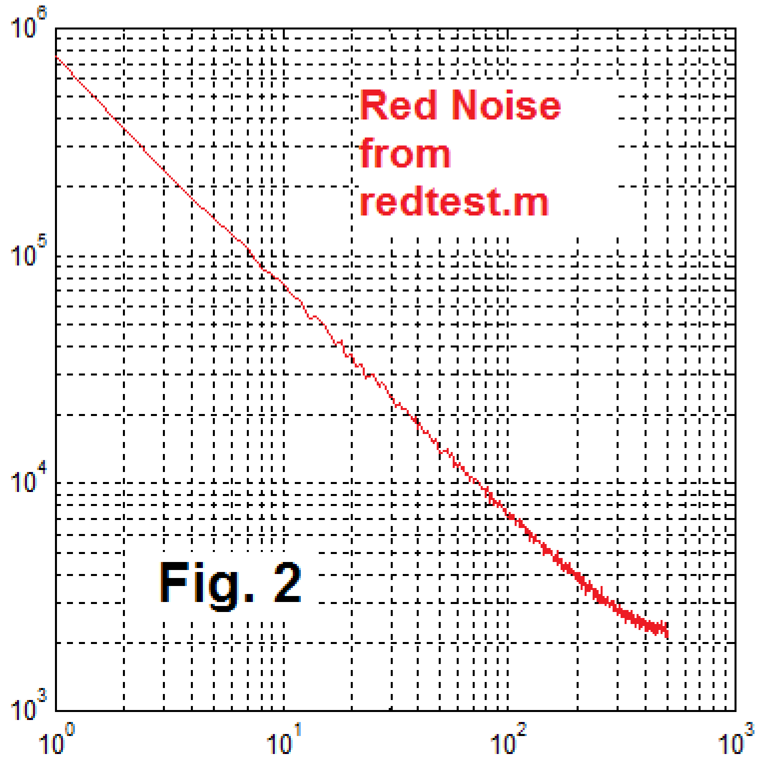
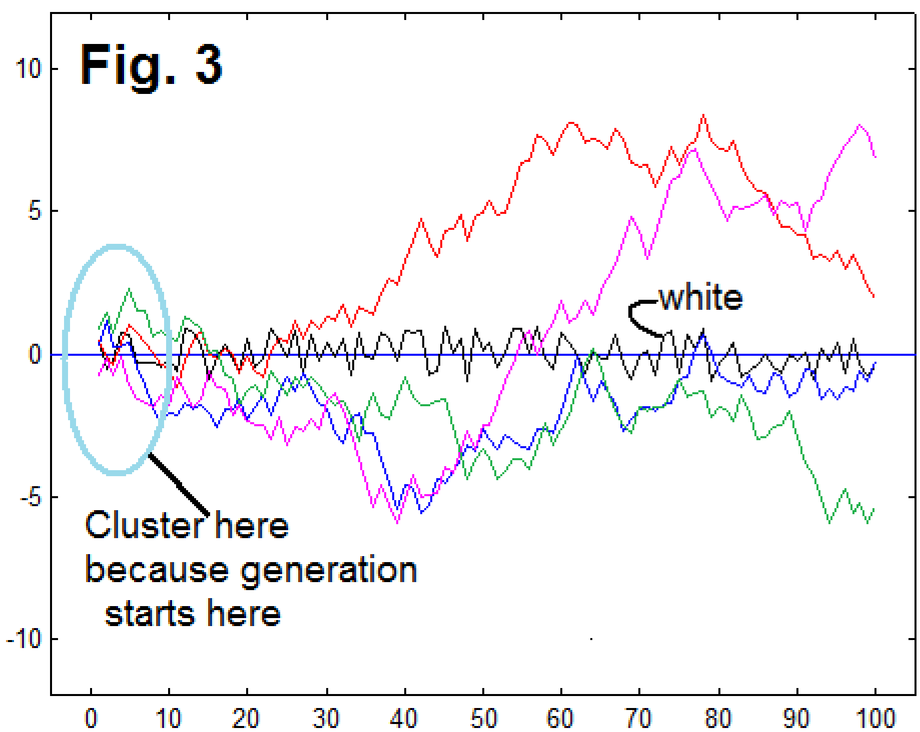
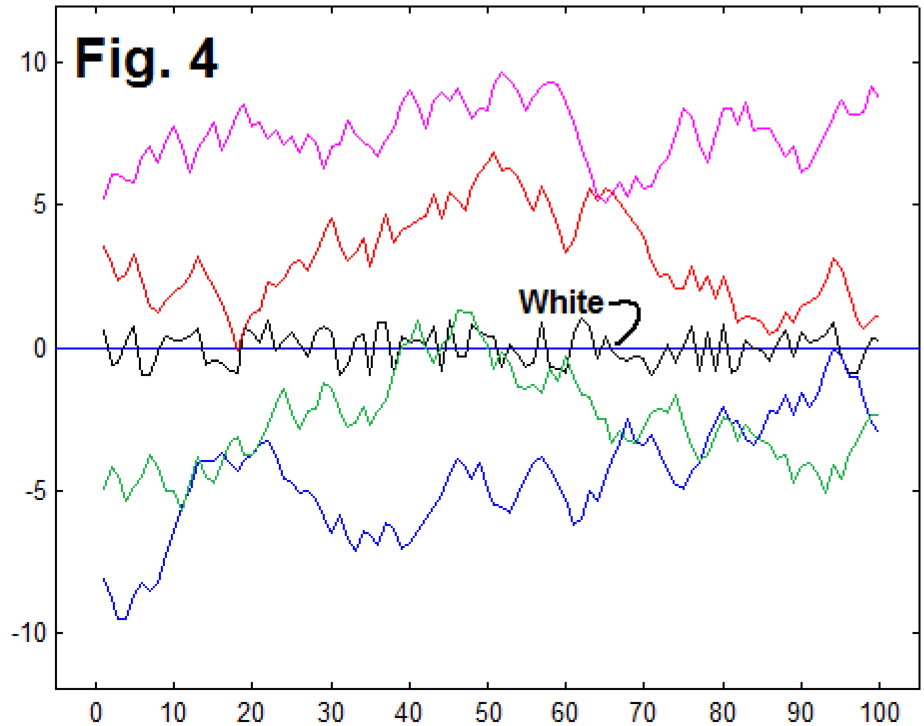


Fig. 2

The Matlab code redtest.m and the program output in Fig. 2 shows that the generator does give a 6db/octave (45 degree on log-log paper) rolloff. This will be our basic tool when we want to generate red signals. Fig. 3 shows a typical white noise signal (the black curve) along with its corresponding red noise (red curve) and three other red noises (blue, green, and magenta) generated from white noise signals not shown. Note that the red noise curves are very different from the white noise. They move to the next samples gradually, so successive samples are correlated. Two additional things to note: First, the white noise has many zero crossings (about 53?) while the red noise curves have very few. In fact, the red curve seems to have a strong preference for positive, the magenta curve tried negative and then turned positive, the green favors the negative range, and the blue seems to be toying with the idea of returning to zero – but who can say? Every run will be different. Second, there is a clustering or convergence of the curves around the start. This is artificial - purely due to the fact that we started to generate the curves starting there. A more realistic idea of what the red noise curves should really look like is seen in Fig. 4, obtained by taking the samples much further in. This is much the same is just shifting the curves up or down at random. Where did that constant come from? Well, if it is an integration, recall the arbitrary constant.





So Fig. 4 is more typical, the curves doing their own thing. Note the scarcity of zero crossings, except for the one white noise curve. Corresponding to this is the observation that the portions shown here are not zero mean, or even close (the white noise segment may be near zero mean).

SIMULATING PROXY TEMPERATURE CURVES

So What? Well – this really is fun to play with red noise signals, but why am I doing it here. Actually, I played with this several years ago, and the result was ambiguous. Here is what it is about. (Still here, the discussion is not intended to be comprehensive.)

It is about the "Hockey Stick" of climate change fame. It is about the fact that there were no good thermometers or possibly even people who cared about keeping track of temperature in the past, at least until perhaps 1800 at best. It's hard (some say impossible) to determine, or even define, the "temperature of the earth" anyway. Yet by using temperature "Proxies" such as tree rings, one can attempt to recover some possible notion of historical temperatures – so called paleoclimatology.

So let's say we have a lot of "records" of something that may be related to climate, perhaps mainly to temperature. Perhaps tree rings, sediments, coral layers, etc. The data are all noisy, incomplete, and perhaps even totally suspect. Looking at the stack of data, we can't take it all in at a glance (calibrations difference, error bars, incomplete sequences, obvious conflicts) so we resort to statistics to try to dig out something meaningful. Things we do find may be coincidental (enough monkeys, enough typewriters…), or illusionary (like looping audio noise and hearing speech – especially if someone else tells us what we should be hearing). Of course, we are certainly also subject to confirmation bias (Hey – hey! – there it is – just what I was suspecting).

One overriding principle may perhaps remain useful. Garbage In – Garbage Out. But I am thinking here of that maxim in the reverse direction of the usual. I mean that if we put in garbage, we SHOULD get garbage out. Garbage in ⟶ Good Stuff Out is not acceptable. Specifically, if we put in random data, and think we see something valid coming out, we are fooling ourselves. If you get a result with your real data, that's one thing. If you get essentially the same result using random data, you just blew any credibility. And, you are supposed to try this!

To be clear here, there is not even one sample of real physical data below. It is all just random numbers. Neither have we selected signals that might resemble real proxy data. We took what came out, with few exceptions, and the code is all here. [One exception is when a plot ran off the chart. So a few things were redone.]

In the case of the "Hockey Stick" [1] a sharp upturn in the late 20th century "global temperature" was reported, along with an equally important flattening of supposed warm periods going back a thousand years. That is, it was said to be getting warmer (no not warming – alarmingly HOT) since about 1970, and it never was as hot in the past (at least for 1000 years = "forever" for many people). There are many many reasons why this claim is likely bogus. One of these reasons is that the statistical methods used to handle proxy data were apparently (for whatever reason) used incorrectly.

Specifically, the technique of "Principal Component Analysis" (PCA) was employed, but with some heroic detective work, was found to use a non-standard normalization [1,2]. It was claimed that the reported hockey stick (HS) shape was a result of the error (innovation!) in the PCA code. This is hard to argue with because the claim is that with random data (instead of tree rings, etc.) the method "mined" that data for hockey sticks.

Previously [3] we looked at a perhaps related example of how a "selection fallacy", as used by some researchers, would lead to a HS shape, from random data. This was basically – silly. Here, the faulty PCA is perhaps more subtle. The procedure is to work with red noise, using a standard PCA, and then using a PCA with faulty normalization. In this note, we are not using PCA (see AN-385 to appear), but just discussing how the non-standard normalization leads to signals of a different nature.

Why red noise – why a correlation. Local "weather" changes, quite unpredictably on an hourly (or minute by minute) time scale; as well as somewhat more predictably as daily and seasonal changes. "Climate" also varies always (far less predictably) on a time scale of perhaps decades and even much much longer. Knowing that a climate variable, let's say temperature averaged over a year might result in a more robust growth of a tree as reflected in tree ring characteristics, we might ask what else could lead to stronger or weaker tree ring growth. We would also suspect that a good year might be followed by another relatively good year. Perhaps decadal trends are at work. Or, perhaps a good year resulted in a storage of resources (like in the roots) so that the next year was also quite good, despite less favorable weather perhaps, and so on. It is likely clear that a redder noise is more appropriate than a whiter one.

Let's begin with the generation of a set of 40 length 50 red noise signals. These are shown in Fig. 5a. This is basically an extension of the result in Fig. 3 where there were only four signals. Here we just consider the overall pattern (note that one of the 40 goes off scale). Yes, it's a mess. The exact details are of no importance. The clustering around zero is again observed, and we know why this happens (we start there). We likely guess that it is not the best place to choose our signals. We need to choose signals further in (like Fig. 4) and it seems we could just take the signals from the start but add a random starting point (lift or lower randomly).

It is obvious that these signals in Fig. 5a are not zero-mean. This is a consequence of the fact that they are truncated. Even a truncated white signal is not zero mean, and the red signals are even more problematic. Why do we care about zero mean? Because the PCA theory demands that. Further, it should be interesting to observe what happens when we remove the means on a signal-by-signal basis, for which we will obviously have to use smaller (less cluttered) sets of signals. But for now, the large set.

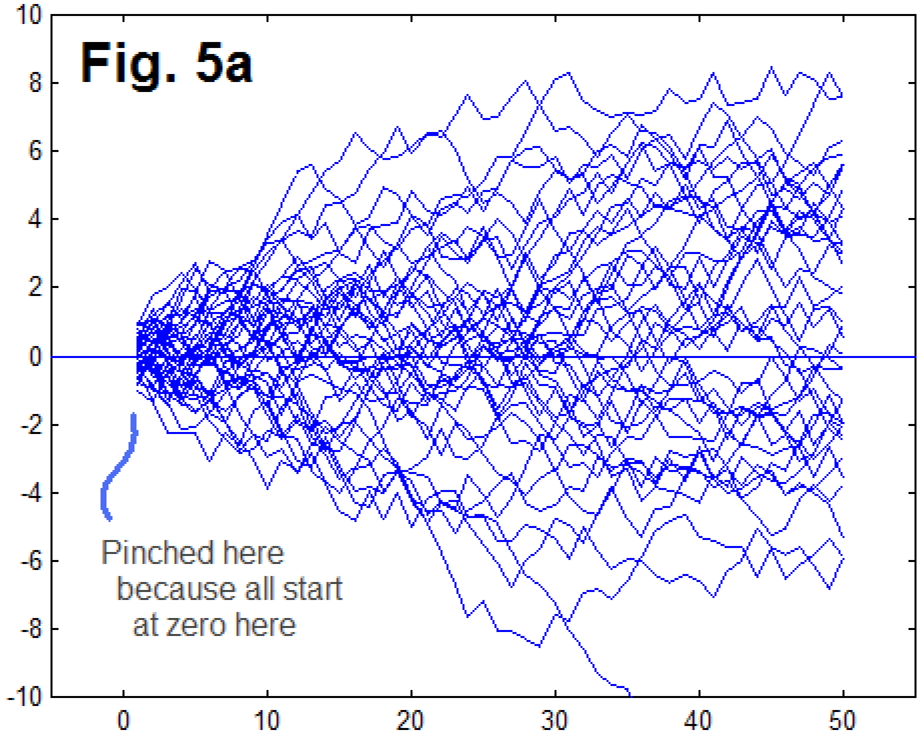


Fig. 5a
Pinched here
because all start
at zero here

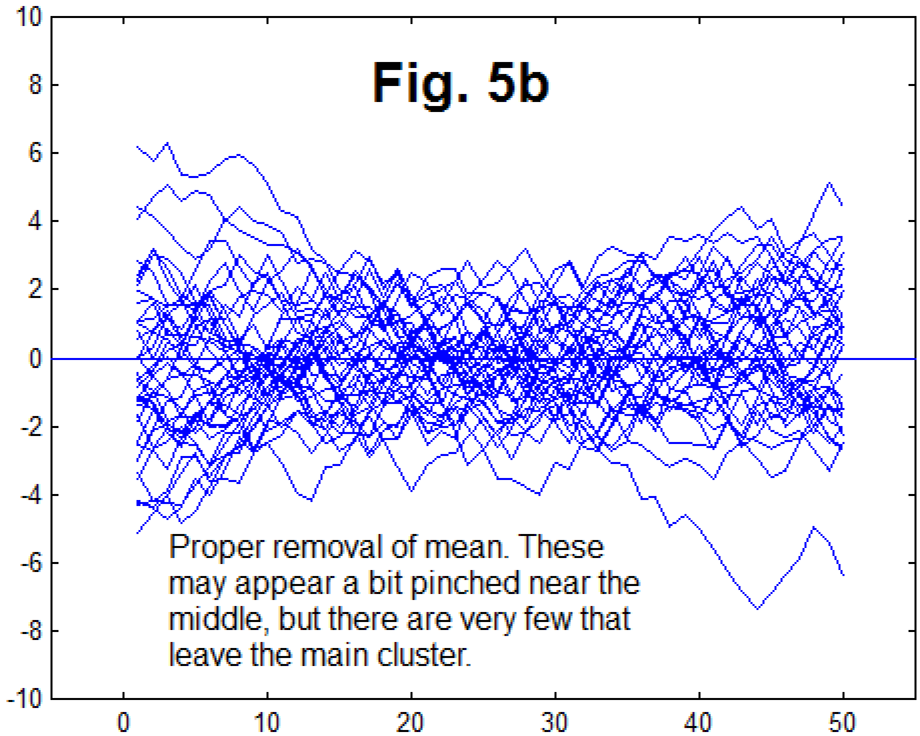
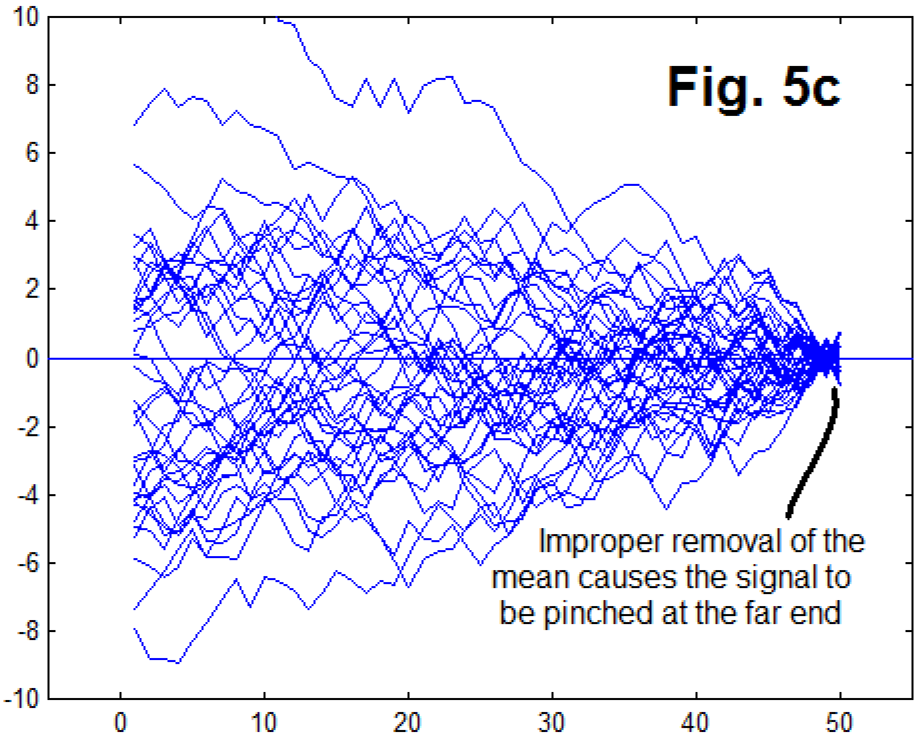


Fig. 5b
Proper removal of mean. These
may appear a bit pinched near the
middle, but there are very few that
leave the main cluster.

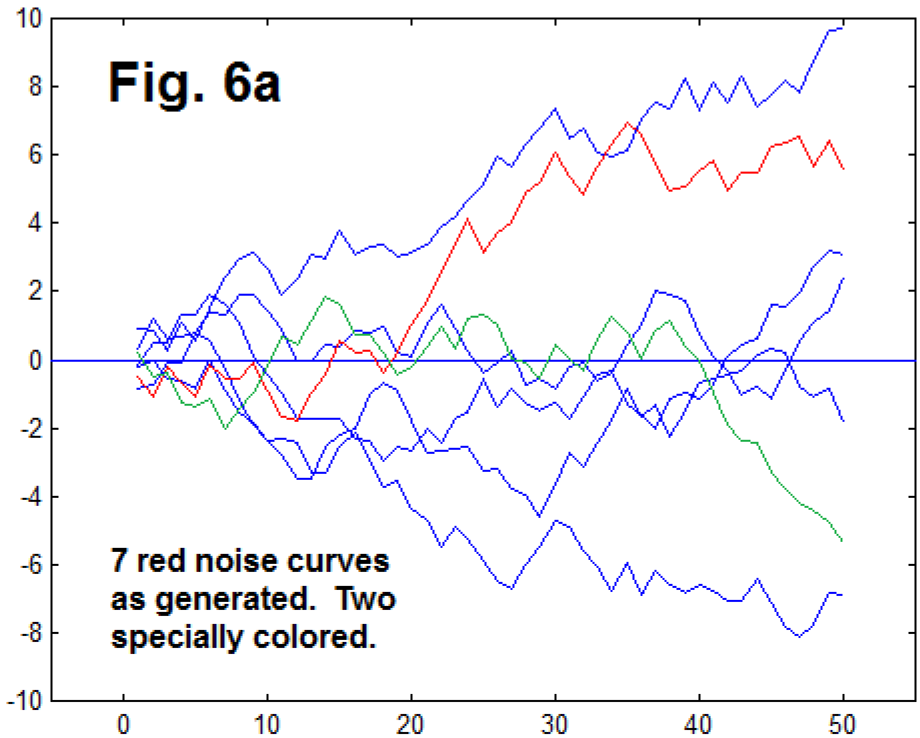
Removing the mean we see that the clustering around zero will go away (Fig. 5b). Recall that that was an artificial clustering – just a matter of where we started the integration. We noted that the clustering could have been eliminated by adding a suitable random offset, but here the clustering goes away because of the removal of the specific calculated means of each signal. This is a special treatment. So all the signals have been shifted up or down, although it is virtually impossible to see this in these plots (redone with fewer signals in Fig. 6 below).

In the controversial alternative procedure, instead of removing the means of the full sequence, only the means of a small region at the end was used. Initially we might supposed that the result would look similar to Fig. 5b. But it's very different, and this is the point. The actual result is seen in Fig. 5c. There is a strong resemblance to Fig. 5a, except it is time-reversed. How did that happen? It looks as though the signals all decided to head for zero. The pinching in Fig. 5c here is completely artificial in much the same sense that the pinching in Fig. 5a was a matter of choosing (artificially) a zero starting value prior to integrating the white noise. What did we do here – specifically?



We subtracted off a mean. NOT the mean of the entire sequence, as PCA requires, but the mean of only the last four samples of the 50 total. Because the samples are highly correlated, the last four samples are somewhat alike (as are any four consecutive samples), so tend toward one number. Now, taking this as a valid notion of the correct mean to subtract off, we necessarily force all the red noise curves close to zero for those end samples. It could not be otherwise. This is all that happens in Fig. 5c. It is in fact a wider pinch (four samples) than the pinch in Fig. 5a at zero. In the actual climate data, the mean was taken not over the entire proxy data record (about 1000 years), but only of the most recent hundred or so years [2].

So, the question as to whether or not the faulty (at best, alternative) method makes a difference is clearly answered yes. Well, at least it is yes in the case where the mean of different sub-sequences are very different. Thus it is likely to make a difference in a red noise case, where, for some, perhaps most of the curves we generate, long sections wander far away from zero for a while. It is far less important if the sequence had no correlation or change of statistical parameters (obviously). But what do we get in a relatively large collection? Well, we get Fig. 5c, but that's not much help if we want to consider the influence of any one curve. So let's look at fewer curves, keeping in mind that the word "typical" is not really applicable here.



In Fig. 6a we have plotted seven red signals just as they are generated from zero. Two of them were specifically colored (red and green respectively) to help recognize them later in Fig. 6b and Fig. 6c. This was the impossible task with Fig. 5. It is possible to recognize the other 5 in Fig. 6 if you don't mind a little eyestrain.

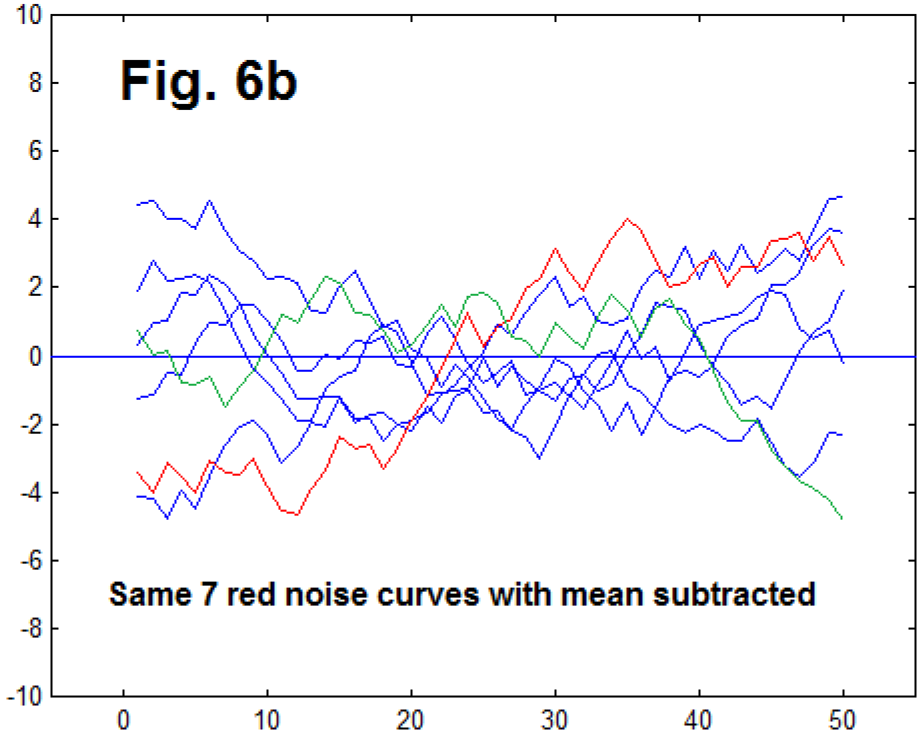


Fig. 6b
Same 7 red noise curves with mean subtracted
10
8
6
4
2
0
-2
-4
-6
-8
-10
0
10
20
30
40
50

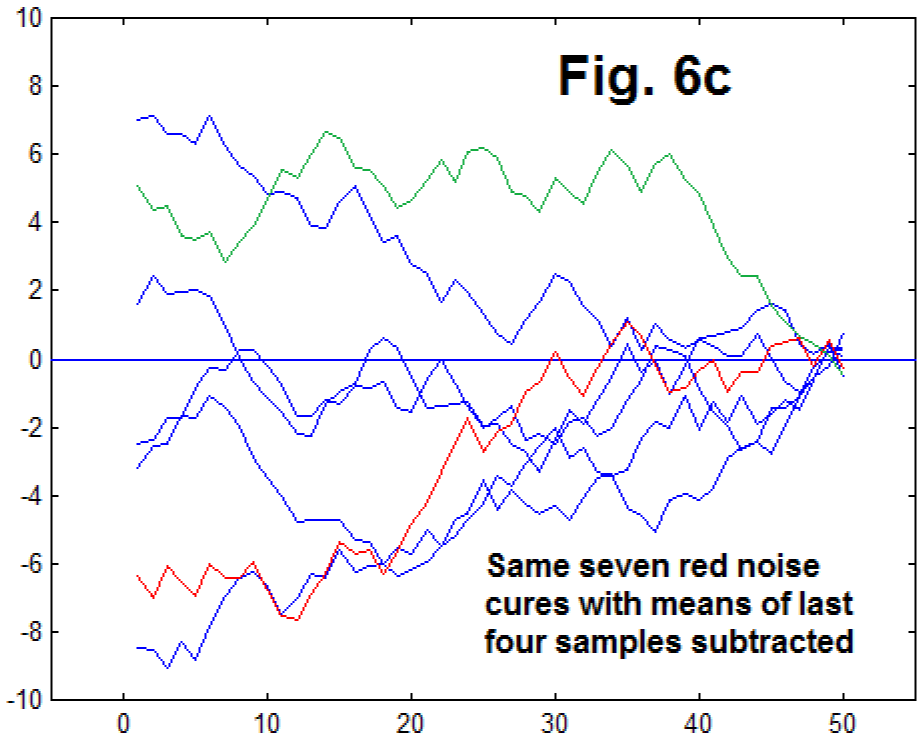
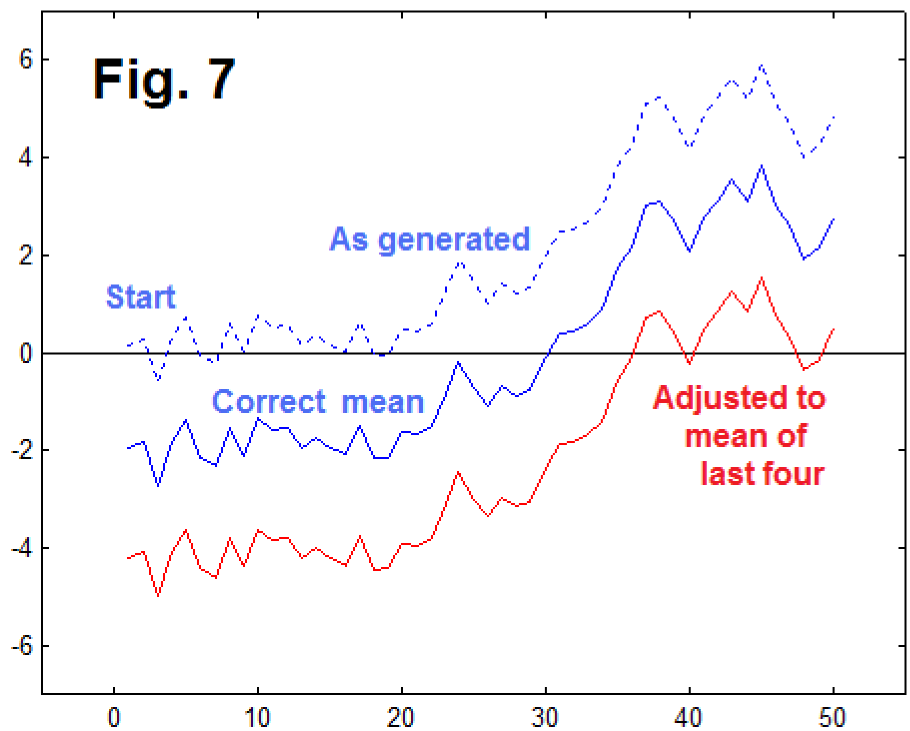


Fig. 6c
Same seven red noise
cures with means of last
four samples subtracted
10
8
6
4
2
0
-2
-4
-6
-8
-10
0
10
20
30
40
50

Happily the red curve is a good example, as it has a diagonal sweep to it. Thus we see that it is highest in Fig. 6a, and lowest in Fig. 6c. Specifically, in Fig. 6a the left end is caught in the start pinch, and in Fig. 6c, the right side is caught in the pinch caused by the irregular normalization. The middle case (Fig. 6b) shows what we can well expect has equal extent above and below the horizontal axis – the usual mean. Note the difference between Fig. 6b (removal of means) and Fig. 4 (generation away from the starting point). Fig. 4 does not have zero mean; Fig. 6b does for each curve.



Now, in Fig. 7, we do show just one curve, the original as generated, the zero mean, and the mean of the last four. Above we said we took the curves as they randomly came out, with a few exceptions. This is one of the exceptions. The program was iterated until we got a good example (probably a dozen tries), specifically to get the upward swing (as occurred naturally in the red-colored curve in Fig. 6).

The curve is an example of a hockey stick, but this is not the point. If we ran enough examples, we could get about any curve we wanted. Curves such as the example of Fig. 7 with the HS shape probably occur in 5% of the examples. But, this is not the point. Here we have shown what happens under the different normalizations.

The dotted blue curve is the original as generated, and note that its left side is in a region of the axis. The blue version is the correct normalization. The red curve is the unconventional normalization, and note that its right side is in the region of the axis. So the curve is shifted so that it changes from the dotted blue to the red – effectively pined at its different ends. But that is not the point either, except we agree that the first curve (dotted blue) is not normalized at all, and it was not acceptable for PCA. Who would have even though to starting there?

So what is the point? In pinning one end (either end) of the curve at zero, the contribution of that particular curve to the total variance is greatly enhanced relative to the properly centered curve. This means that in the PCA calculation, this curve has much greater influence than it deserves if it were properly centered, and its shape is enhanced in the first principal component. In pinning all the curves about zero at the right side, those with greater "slope" will be given more weight than they deserve. This is true to a certain extent for most of the curves, although generally to a lesser extent than in the example of Fig. 7. It all adds up to a HS. Incidentally, the downslopes and upslopes contribute to the same result in PCA.

Another way to look at it that is more familiar is perhaps to ask what the "energy" in the three curves of Fig. 7 is. It would seem that the dotted blue and red curves would have more energy than the blue curve, and this is true. If we sum the squares of the samples for each curve, we get:

dotted blue (as generated)	428.03
blue (correct mean)	206.89
red (non-standard mean)	466.06

SUMMARY:

Here we hope to have provided a basic introduction to red noise and why we are using it. Further, we have attempted to show why improper normalization of the noise leads to a biasing of the signals. In particular, even before we consider PCA and its requirements, we see that the signals are drastically shifted and "pinched" to pass into a region near the axis near the end of the signal. The consequences of this, along with PCA examples, will appear in the next AN in this series.

REFERENCES:

[1] Montford, Andrew W., *The Hockey Stick Illusion, Climategate and the Corruption of Science,* Stacey International (2010).

[2] McIntyre, Stephen and Ross McKitrick, "Hockey sticks, principal components, and spurious significance", *Geophysical Research Letters*, Vol. 32, L03710, doi:10.1029/2004GL021750, (2005):

http://climateaudit.files.wordpress.com/2009/12/mcintyre-grl-2005.pdf

See also an introductory summary by Ross McKitrick "What is the Hockey Stick Debate About?", APEC Study Group, Australia April 4, 2005:

http://www.uoguelph.ca/~rmckitri/research/McKitrick-hockeystick.pdf

[3] Hutchins, B., "Models – Good, and Bad: and the (Mis-)Use of Engineering Ideas in Them," *Electronotes* Volume 22, Number 211 July 2012

http://electronotes.netfirms.com/EN211.pdf

SOME MORE MATLAB CODE USED HERE:

```
%  redsignals
%  Compare Standard vs. Custom Mean Removal
%  This program generates some of the figures in AN-384 as
%    noted in comments.  The generation portion is common
%    to all the figures 3-7 of AN-384.  The display code
%    is similar to that here for all of Fig. 3-7.
%
% GENERATE SIGNALS
% 40 length 50 signals - RED Noise
R=40    % number of signals
C=50    % length of signals
xsorig=2*(rand(R,C)-.5);  % White
% Now Redden
xs=xsorig;
for kk=1:R
   for mm=2:C
      xs(kk,mm)=xs(kk,mm-1) +xs(kk,mm);
   end
end
% Transpose and save copies zs and ys
xs=xs';   % to be mean-adjusted standard
zs=xs ;   % reserve copy
ys=xs ;   % to be mean adjusted non-standard
%
% ****************************************
% Standard Removal of Mean - xs
mn = mean(xs,1);
for k=1:R
   xs(:,k)=xs(:,k)-mn(k);
end
%
% Now Plot/Display Standard - First 5
figure(1)
plot([-100 500],[0 0])
hold on
plot(xs(:,1),'b')
plot(xs(:,2),'r')
plot(xs(:,3),'k')
plot(xs(:,4),'g')
plot(xs(:,5),'m')
axis([-5 55 -7 7])
hold off
figure(1)
```

```
%
% ****************************************
% Modified Removal of Mean  - ys
for k=1:R
   rowk=ys(:,k);
   mn(k)=mean(rowk(48:50));
   ys(:,k)=ys(:,k)-mn(k);
end
% Now Plot/Display Modified - First 5
figure(2)
plot([-100 500],[0 0])
hold on
plot(ys(:,1),'b')
plot(ys(:,2),'r')
plot(ys(:,3),'k')
plot(ys(:,4),'g')
plot(ys(:,5),'m')
axis([-5 C+5 -7 7])
hold off
figure(2)
%
%
% PLOTS OF FULL SETS (Fig. 5 of AN-384)
% Plots of Red Signals
% Originals   zs
figure(3)
plot([-5 500],[0 0])
hold on
for k=1:R
   plot(zs(:,k))
end
hold off
axis([-5 C+5 -10 10])
%
% Standard Removal of Mean  xs
figure(4)
plot([-5 500],[0 0])
hold on
for k=1:R
   plot(xs(:,k))
end
hold off
axis([-5 C+5 -10 10])
%
```

```
% Modified Removal of Mean  ys
figure(5)
plot([-5 500],[0 0])
hold on
for k=1:R
   plot(ys(:,k))
end
hold off
axis([-5 C+5 -10 10])
figure(5)

% Single Signal (Fig. 7 of AN-384)
figure(6)
plot([-5 500],[0 0],'k')
hold on
plot(zs(:,k),'b:')
plot(xs(:,k),'b')
plot(ys(:,k),'r')

hold off
axis([-5 C+5 -7 7])
figure(6)
```