Re: Science research student

Benjamin Oren Goldman <bg502257@live.wpcsd.k12.ny.us>
Tue 8/4/2020 3:43 PM
To: John Fasullo <fasullos@gmail.com>

▌ 4 attachments (272 KB)
19.pdf; 21.pdf; 22.pdf; 23.pdf;

Hi John,

These are the correlation plots between the temperature cross sections and the Nino 3.4 variance timeseries for the ff, xaer, xghg ensemble means.

In all cross sections except for EEP, the xghg ensemble ocean temp. has increased (more positive) correlation with the Nino 3.4 variance on the surface layer, while in the xaer ensemble, it is reduced (more negative). While the subsurface layer correlations differ between ensembles, I cannot discern a coherent pattern for those differences. This might become clearer once I get the timeseries for each layer.

PS. All of the figures are now vector graphics and have font and size conforming to the AGU requirements.

> On Aug 2, 2020, at 8:01 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
>
> Those x labels are meant to be latitude. Sorry about that.
>
>> On Aug 2, 2020, at 7:29 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
>>
>> Should I look at the xlulc/xbmb/fixedO3 cross sections, or do we only care about the xghg and xaer?
>>
>>> On Jul 31, 2020, at 1:20 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
>>>
>>> Here's one more idea: I could find the correlation between Nino3.4 variance and EP SST. It could be overwhelmed by the warming trend, but it might be a good idea to check anyway.
>>>
>>> Here are plots of the correlation between the Nino 3.4 variance and ocean temperature. The fact that in the EEP the surface positive correlation survived detrending is what tells me that maybe I should look at the correlation with the surface. Its asymmetry surprises me. Also, interestingly, there are contrasting patterns in the subsurface correlation between the CEP and EEP plots, specifically the regions of lower (less negative) correlation around +- 4 latitude.
>>>
>>> In a couple of days I will have the timeseries for the surface and subsurface layers so we can check these plots. After that I will prepare the same 4 regions for the xghg and xaer regions. Then, I will do the pd and salinity data.

Would Dr. Gent like access to my figures and/or code?

<21.png>
<23.png>
<22.png>

On Jul 29, 2020, at 1:39 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

Ok. What dataset should I focus on first? Should I wait to do that until the SF cross sections are done or is it more urgent?

On Jul 29, 2020, at 12:31 PM, John Fasullo <fasullos@gmail.com> wrote:

Also it occurs to me that looking at potential density rather than temperature in our correlation plots will be helpful in diagnosing the cumulative influence of salinity and temperature anomalies. I've placed the corresponding PD fields on the ftp site.

On Jul 27, 2020, at 5:53 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

Yeah, that sounds great! What will his role be? Turns out I forgot to detrend the timeseries. I don't think that will drastically affect the correlation with the subsurface layer, as it's nonlinear. I will send the updated figures tomorrow.

On Jul 27, 2020, at 6:19 PM, John Fasullo <fasullos@gmail.com> wrote:

Note - on your ocean plots, use a logarithmic scale on the depth axis and plot actual depths (rather than the index). I'm going to invite an oceanographer friend, Peter Gent, to join us for some of the ocean analysis. I hope that is OK with you?

On Jul 27, 2020, at 8:30 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

I was able to log into the ftp server and can see the files. Can you please also add the single-forcing TEMPDT files? I think I should look at the correlation in those before I go to the meridional cross-sections. Thanks!

On Jul 26, 2020, at 5:16 PM, John Fasullo <fasullos@gmail.com> wrote:

Yes, they may be. I've also added the files form the single forcing runs. So the idea now is to explore explanations for the change in variance that we see across these ensembles.

I think the role of the forcing will be an interesting question and I have some ideas on that. One may simply have to do with timescale with which heat penetrates with depth in the central Pacific or changes in the curl of

the surface wind stress. Clearly there is no simple monotonic relationship between either a forcing agent or warming and ENSO variance. Rather, it seems to relate to the changes in the ocean base state over time and once we know what aspects of the structure are key (clearly the stratification in the central Pacific appears to be key) we can look at the drivers. We then might also look for evidence of these effects across models. I think this will make a nice manuscript. They key will be to fit it in the length of a short form manuscript, which may be tough.

> On Jul 26, 2020, at 1:23 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
>
> Sounds good! Are the meridional cross sections the same as the files I downloaded earlier from the google drive? If they are, then I already have those files.
>
> So we still have yet to explain the physical link between forcing, the temperature of the subsurface layer, and ENSO amplitude. Is there previous research explaining the trend in the subsurface layer? Also, you mentioned kelvin waves earlier, which I am reading about.
>
> I know how to use LaTeX, which I thought might help with the final paper. Do you or the journal have any special formatting requirements I should keep in mind? Latex produces a paper formatted similar to APA.
>
> > On Jul 26, 2020, at 2:51 PM, John Fasullo <fasullos@gmail.com> wrote:
> >
> > This is promising. Let's see now what the relationships look like with meridional cross sections. I've placed files for the cross sections of the western, central, and eastern Pacific on the ftp site (TEMPWEP, TEMPCEP, TEMPEEP).
> >
> > After that, we should see if these relationships are also suggested in the single forcing runs. If they are, I think we are getting close to starting to write things up.
> >
> > > On Jul 21, 2020, at 1:07 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
> > >
> > > I added a few color/margin adjustments to save space.
> > >
> > > <20.png>
> > >
> > > > On Jul 21, 2020, at 2:50 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
> > > >
> > > > Ok, here is the updated figure. C is the subsurface layer, d is the surface layer, and e is the difference between the 2.
> > > > <20.png>

On Jul 21, 2020, at 1:00 PM, John Fasullo
<fasullos@gmail.com> wrote:

Yes, I'd try both.

> On Jul 21, 2020, at 10:50 AM, Ben Goldman
> <bg502257@live.wpcsd.k12.ny.us> wrote:
>
> Ok. Which method (or both) should I follow through on? I
> could easily modify the the correlation masking to use the
> longitude bounds you mentioned.
>
>> On Jul 20, 2020, at 12:18 PM, John Fasullo
>> <fasullos@gmail.com> wrote:
>>
>> Great - I think. I was actually thinking of something simpler
>> where you just create a timeseries for the subsurface layer
>> (indexes 15-20, lon 120-220) and do the same you've done
>> here. Or instead you could create a difference of the
>> timeseries for indexes 0-10 lon 120-220 and the indexes
>> 15-20, lon 120-220 already mentioned, to see if the
>> correlation is higher. I think that stability is arguably very
>> relevant to the efficacy of Kelvin waves to influence
>> nino3.4 SST and is easily justified for playing a key role.
>>
>>> On Jul 20, 2020, at 10:10 AM, Ben Goldman
>>> <bg502257@live.wpcsd.k12.ny.us> wrote:
>>>
>>> Hi,
>>>
>>> I calculated the mean TEMPDT timeseries for regions
>>> with high and low correlations with the Nino 3.4 variance
>>> timeseries. I excluded the deeper ocean (index < 30).
>>> The timeseries are smoothed with a 5-year running
>>> mean, to remove the strong annual variability.
>>>
>>> <PastedGraphic-1.tiff>
>>>
>>> The region that is negatively correlated has a timeseries
>>> mirroring that of the Nino 3.4 variance, while the
>>> positively correlated region is dominated by a warming
>>> trend. This data suggests that there may be a
>>> meaningful connection between the structure of the
>>> equatorial Pacific and ENSO variance. Do you agree?
>>>
>>>> On Jul 16, 2020, at 4:30 PM, John Fasullo
>>>> <fasullos@gmail.com> wrote:

I think you're on the right track. These look impressive – but too impressive really. I think the trend is playing a major role in the correlation, so best to remove it if you haven't. One way to test this is to create timeseries for averages across areas where the correlation is coherent and strong, and look at the timeseries. I don't think the deep ocean is playing a role but rather that the trends tend to align...but perhaps I'm wrong.

> On Jul 16, 2020, at 11:54 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
>
> Hi Dr. Fasullo,
>
> I finished calculating the correlation and found some interesting results. Here is a contour plot of the correlation between the smoothed equatorial sea temperature and the smoothed Nino 3.4 variance in the full forcing ensemble. Sorry I forgot the units, depth is in feet.<19.png>
>
> Here is the same plot but with the y-axis as the array index, rather than depth, showing more detail closer to the surface. <19a.png>
>
> This is the same plot, but with no smoothing. Same pattern, with decreased significance.<19b.png>
>
> The region with negative correlation near the surface is the thermocline, correct? This data suggests that the change in ENSO amplitude is accompanied (and maybe caused) by increasing stratification. I am curious what the role of the very high coefficients in the lower levels is?
>
>> On Jul 7, 2020, at 11:52 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:
>>
>> Ok, thanks for explaining. One last thing, can you please clarify what you mean about the degrees of freedom? I know that they are the distinct ways in which a system can vary, but how does the correlation affect it, and how does that affect finding the significance? Thank you so much for spending the time to help me with small things like this.
>>
>>> On Jul 7, 2020, at 11:42 AM, John Fasullo <fasullos@gmail.com> wrote:

That's essentially what it is. Get the timeseries of the ocean temperature data, smooth it with a ~30-yr smoothing, and correlate at each lon/depth against the variance timeseries (smoothed with the same smoother). Note that the # of degrees of freedom will be significantly reduced in estimating significance.

On Jul 7, 2020, at 9:38 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

Ok. Would you mind explaining how correlation works between multi-dimensional data? I only really know how to correlate 2 1-d arrays.

On Jul 7, 2020, at 11:33 AM, John Fasullo <fasullos@gmail.com> wrote:

I'm thinking of starting perhaps with a simple correlation analysis between the variance timeseries with the detrended temperature trends as a f(lon,depth).

On Jul 7, 2020, at 9:23 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

Ok, great! So now I will try to make the same modification to all the files in the dataset and then I can start calculating. After concatenating the files into full time series, are there any specific calculations I should do? What specific things are we looking for? How can we calculate the statistical significance of any changes we find?

On Jul 7, 2020, at 11:18 AM, John Fasullo <fasullos@gmail.com> wrote:

Yes, that looks about right – depends on the color scale – and its upside down. This is the vertical structure of ocean temperature averaged 10N-10S and South America and Africa are the missing regions.

On Jul 7, 2020, at 9:16 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

I was able to produce this figure after deleting only the valid_range field: It is the TEMPDT variable at time=0<PastedGraphic-5.png> Does it look correct to you?

On Jul 7, 2020, at 11:11 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

Hi Dr. Fasullo,

I think I may have found out a way to solve the problem. Here is a screenshot of one of the files' attributes: <PastedGraphic-4.png>

The netCDF reader doesn't like the fields with "--" in them, especially in the variables, and where they should be numbers (such as valid_range). I figured out how to change or delete these fields in python. Could you please tell me which of them you think are necessary, and which values I should insert for the ones I keep? Are any of them necessary for calculations (except for the actual data)? Thanks!

On Jul 6, 2020, at 12:32 PM, John Fasullo <fasullos@gmail.com> wrote:

Ah – that could be. I use IDL for these sorts of changes though I imagine python can as well.

On Jul 6, 2020, at 10:03 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

My guess is that the nc reader doesn't like dashes in the fields such as "valid range". I can manually fix all of those in my text editor, I think. Do you use any particular software for making changes like that to many files at a time?

On Jul 6, 2020, at 11:50 AM, John Fasullo <fasullos@gmail.com> wrote:

Yes - ignore the .n files, I've deleted them. When I open the file you mention it all seems to work for me (info below). Perhaps '-' corresponds to missing values (due to bathymetry)? I'm not a python user so have limited insight.

```
IDL>
fi=ncdf_open('b.e11.B20TRC5CNBDRD.f09_g16.
001.pop.h.TEMPDT.185001-200512.nc')
% Loaded DLM: NCDF.
ncIDL> ncdf_varget,fi,'TEMPDT',in
hlp,in
IDL> hlp,in
    1   INTARR(180,60,1872), MIN(7,0,0)= -32767,
MAX(53,0,1685)= 18824
IDL>
```

On Jul 6, 2020, at 9:22 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

I tried it on LE/TEMPDT/b.e11.B20TRC5CNBDRD.f09_g16.001.pop.h.TEMPDT.185001-200512.nc and a few others. When I used the same process on other datasets, such as TREFHT, there was no error.

Also, in the TEMPCEP,

TEMPWEP, and
TEMPEEP folders, there
were many files with the
extension .n . Are those
part of the dataset?

On Jul 6, 2020, at
11:16 AM, John Fasullo
<fasullos@gmail.com>
wrote:

in the TEMPDT
variable? I don't see
that but go ahead and
send me the file name
in question and I'll
have a look.

On Jul 6, 2020, at
9:03 AM, Ben
Goldman
<bg502257@live.wp
csd.k12.ny.us>
wrote:

Oh yeah, sorry, I see
them now. The issue
was that when
opening the variable
in python, I got the
following error:

ValueError: invalid
literal for int() with
base 10: '-'

It looks like some of
the numbers in the
program were
replaced by dashes.
I tried opening a file
after downloading it
individually, and got
the same message.
Is this or something
similar happening
with you?

On Jul 6, 2020, at
10:58 AM, John

Fasullo <fasullos@gmail.com> wrote:

I show them as still being there?

On Jul 6, 2020, at 8:45 AM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

Hi, sorry to bother you again. I was wondering if you could please put the tempdt files back in the google drive? When I downloaded them the first time I think they became corrupted, so I need to re-download them. Thanks!

On Jul 5, 2020, at 4:44 PM, Ben Goldman <bg502257@live.wpcsd.k12.ny.us> wrote:

o