**Data Exercise 2**

Assigned 2/13/2018

Due 2/22/2018

In this exercise we’re going to examine how climate varies differently on different timescales, and how this effect changes through space. To do this we will be calculating standard deviations of temperature and precipitation from the NCEP-NCAR Reanalysis v1 ([temperature](http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/.Diagnostic/.surface/.temp/) and [precip](http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/.Diagnostic/.surface/.prate/)).

Better yet, IRIDL doesn’t have standard deviation functions that work the way we need to so we get to calculate our own!

!!!!!!!!!!!!!!BEGIN INSTRUCTIONS!!!!!!!!!!!!!!!

So here’s a reminder about how to calculate standard deviation from scratch.

Once you have the data that you want calculate,

1. Subtract the mean from the dataset
2. Square it
3. Sum those squared values
4. divide that by the number of time steps minus 1 (N-1)
5. then take the square root

Easy peasy. Except that you have to do all of these operations on every cell. Thankfully, IRIDL has mathematical functions that apply to all cells.

1. to calculate the average in time, use *[T]average*
2. to subtract, use *sub*
3. to sum, use [T]sum
4. to multiply, use *mul*
5. to divide, use *div*
6. to square root, use *sqrt*

OK, that’s all well and good; but IRIDL needs very specific instructions. If you want to subtract, it needs to know to subtract what from what. IRIDL keeps tracks of the datasets by putting them in a stack. [Take a look at this example.](http://iridl.ldeo.columbia.edu/expert/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/.Diagnostic/.surface/.prate%5BT%5Daverage/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/.Diagnostic/.surface/.prate%5BT%5Dsum/) Here, I’ve made a “stack” of two datasets. View my code in expert mode. The stack grows downward, so the most recent dataset is at the bottom, and has a green pin on the left side, the ones above are previous stacks. If I give a command that needs one dataset, it will operate on the most recent dataset in the stack. If I give a command that uses two datasets, it will operate on the two most recent.

So [if I now use the *sub* function](http://iridl.ldeo.columbia.edu/expert/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/.Diagnostic/.surface/.prate%5BT%5Daverage/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/.Diagnostic/.surface/.prate%5BT%5Dsum/sub/), it knows to subtract the most recent dataset from the previous one. Note that there’s now only one dataset again, the difference between the other two. I can map it, or go onto further analysis.

Here’s an even simpler example. If I wanted to divide 80 by 10. I’d have to type

expert

80

10

div

[See it works.](http://iridl.ldeo.columbia.edu/expert/80/10/div/) This is not how we think. We say: 80 div 10. But that won’t work in IRIDL. All the data needs to be in there first, then you *div.* So when you divide by N-1, it will look something like this if N is 88:

expert

previousData (your data here)

87

div

[real example](http://iridl.ldeo.columbia.edu/expert/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.MONTHLY/.Diagnostic/.surface/.prate%5BT%5Daverage/98/div/).

OK, one more function you’ll find useful. *dup*. This will duplicate the dataset on the bottom of the stack. This is necessary if you want to subtract the mean… you’ll do something like this: (hashtags and following text are comments, you wouldn’t include them in your code)

data #the dataset of interest

yearlyAverage #average the data to the requested timescale

dup #duplicate it

[T]average #average the second data set across time

sub # subtract the average data from the original data

or to calculate the square, you duplicate then multiply

previousData #everything up to this point

dup #duplicate

mul #multiply those two duplicate sets together

!!!!!!!!!!!!!!END INSTRUCTIONS!!!!!!!!!!!!!!!

**Tasks:**

1. Create standard deviation maps of temperatures at
	1. monthly (original)
	2. yearly (use *yearlyAverage*)
	3. decadal (use *[T] 120 boxAverage* to average across 120 months)

Note: While you’re creating these three maps, you’ll need to figure out N {the number of observations in time} (and divide by N-1). You can see this if you scroll down (on the description page) after you average appropriately, e.g.:

*“Time* (time)

grid: /T (months since 1960-01-01) ordered (1949) to (2015) by 12.0 **N= 67 pts**”

1. Create two maps that are ratios of
	1. yearly / monthly standard deviations
	2. decadal / yearly standard deviations

[remember: you can just cut and paste your code from 1 to get the two datasets you need to ratio]

1. Repeat 1a-c and 2a-b for precipitation.

Now you should have a total of 10 maps. Use those maps to provide **short answers** to the following questions. If you’re not confident that you created your maps correctly, check with your classmates and/or Diane, since you’ll need good maps to answer these questions.

1. Where does **monthly temperature** vary the most? Where does it vary the least? What are some physical mechanisms that explain this pattern?
2. How about **interannual and decadal temperature** variability? Where does it vary the most and the least? How do these patterns compare to the monthly variability?
3. Where does **monthly precipitation** vary the most? Where does it vary the least? What are some physical mechanisms that explain this pattern?
4. How about **interannual and decadal precipitation** variability? Where does it vary the most and the least? How do these patterns compare to the monthly variability?
5. Now look at the **ratio of yearly to monthly temperature** variability.
	1. Do most regions have more yearly variability, or more monthly variability?
	2. What are the major features of the ratio pattern. Where are these timescales of variability most similar? Why? Where are they most different? Why?
6. Now look at the **ratio of decadal to yearly temperature** variability.
	1. Do most regions have more yearly variability, or more decadal variability?
	2. What are the major features of the ratio pattern. Where are these timescales of variability most similar? Why? Where are they most different? Why?
7. Now look at the **ratio of yearly to monthly precipitation** variability.
	1. Do most regions have more yearly variability, or more monthly variability?
	2. What are the major features of the ratio pattern. Where are these timescales of variability most similar? Why? Where are they most different? Why?
8. Now look at the **ratio of decadal to yearly precipitation** variability.
	1. Do most regions have more yearly variability, or more decadal variability?
	2. What are the major features of the ratio pattern. Where are these timescales of variability most similar? Why? Where are they most different? Why?

**Deliverables.**

**Turn this all in a single document. You should have**

1. **10 maps**
2. **10 legends for those maps**
3. **The answers to 4-11.**