

Advanced Operation of iTacs

Hiroataka SATO,

Tokyo Climate Center, Japan Meteorological Agency (TCC/JMA)

5. Statistical Analysis

Following statistical analyses are available on iTacs.

- ◆ Regression/correlation analysis.
- ◆ Composite analysis.
- ◆ Single/multi EOF, Singular Value Decomposition (SVD) analysis.
- ◆ Fast Fourier Transform (FFT) analysis.
- ◆ Wavelet analysis

These methods can be powerful to consider and understand climate system. However, it should be noted that their statistical results don't always mean the existence of physical system or structures in targeted data, because statistics is just a matter of mathematics. We need physical interpretation after statistical analysis.

In this text, regression/correlation analysis and composite analysis are described. These methods are frequently used to analyze the relationships between two data like temperature in a region and ENSO.

5.1. Regression and Correlation analysis

Regression and correlation analysis are often used to examine the circulation pattern related to the focused one-dimensional time-series. Correlation coefficient means the degree of the correlation (that's to say, how close they have a linear relationship), and the regression coefficient means the gradient of the regression line.

As an example, let's make a regression coefficient map of sea level pressure (SLP) anomaly onto NINO.3 SST index for boreal winter (December to February) mean from 1958/1959 to 2016/2017. Set parameters on "Data1" box as shown below.

- ◆ Data1
 - Dataset: JRA-55.
 - Element: Surface → SLP (Sea Level Pressure).
 - Data type: ANOM.
 - Area: Lat: -90 – 90, Lon: 0 – 360.
 - Level: 1.
 - Time unit: MONTHLY, Check "Year-to-year" box in regression or correlation analysis.
 - Showing period: RANGE, 1958 – 12 for upper box and 2016 – 2 for lower box.

Be careful to parameters in "Showing period". Although the end of the range may appear to be February 2016, this setting means December to February mean from 1958/1959 (December 1958 to February 1959) to 2016/2017 (December 2016 to February 2017). Consider the setting for year and month separately.

"Data2" box will appear after selecting "REGRESSION_ANALYSIS" in "Analysis method" pull-down menu.

- ◆ Data2
 - Dataset: INDEX.
 - Element: NINO.3.

- Data type: ANOM.
- Time unit: MONTHLY, Check “Year-to-year” box in regression or correlation analysis.
- Lag: 0 YEAR. (In this example, simultaneous regression is calculated)
- Significance: 95% two side. (significance test using student’s t-test)

◆ **Graphic Options**

- Drawing: CONTOUR
- Color Table: Blue – Red.
- Set Contour Parameters for data1: interval:0.2, min:-1, max:1.

Data1

Dataset	Element	Data type	Area	Level	Time unit	Showing period
JRA-55	Surface SLP (Sea Level Presst)	ANOM	ALL Lat: -90 - 90 Ave <input type="checkbox"/> Lon: 0 - 360 Ave <input type="checkbox"/>	1	MONTHLY <input type="checkbox"/> Ave <input checked="" type="checkbox"/> Year-to-year <input type="checkbox"/> Time filter	RANGE 1958 - 2016 12 - 2

Analysis method: REGRESSION_ANALYSIS

Data2

Dataset	Element	Data type	Time unit	Lag	Significance
INDEX	NINO.3	ANOM	MONTHLY <input type="checkbox"/> Ave <input checked="" type="checkbox"/> Year-to-year <input type="checkbox"/> Time filter	0 YEAR	95%(two side)

Graphic Options

Colorizing: COLOR	<input checked="" type="checkbox"/> Show Contour Labels
Drawing: CONTOUR	<input checked="" type="checkbox"/> Show Color Bar
Image Format: png	<input checked="" type="checkbox"/> Set Contour Parameters for data1
Font: default	interval: 0.2 min: -1 max: 1
Color Table: Blue - Red	<input type="checkbox"/> Set Vector size: [inch] value: skip: 1

Fig. 22a Parameter setting and graphic options to draw regression coefficient and significance of SLP onto NINO.3 index for DJF 1958/1959 to 2016/2017

Finally, click “Analysis Data Submit” button and the regression map will be displayed.

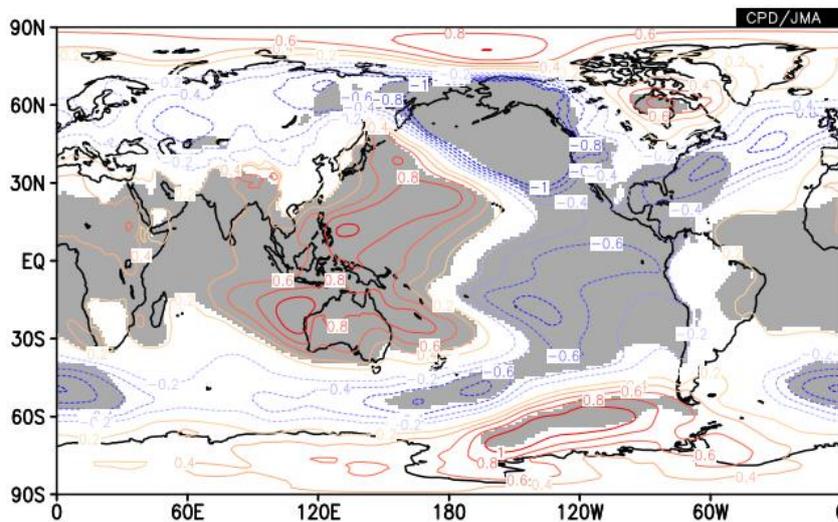


Fig. 22b Regression coefficient of SLP onto NINO.3 index for DJF 1958/1959 to 2016/2017 created by the setting shown in Fig. 22a. Statistical significance 95% or higher is also indicated by gray shading.

The regression coefficient will be mapped as contour, and the area where regression coefficient is significant at a given (in this case, 95%) confidence level is marked as gray shading.

Correlation analysis can be done by almost the same process but select “CORRELATION_COEFFICIENT” from the “Analysis method” box.

5.2. Composite analysis

Users can calculate and map the composite data by using “COMPOSITE” analysis method. Let’s make a composite map of 3-month mean zonal wind anomaly at 850hPa for December-January-February when SST anomaly in NINO.3 region was larger than +0.5. In this time, target period is from 1958 to 2016. Set parameters on “Data1” and “Data2” box as shown below (See also Fig.9a).

◆ Data1

- Dataset: JRA-55.
- Element: Pressure Levels → U (Zonal Wind).
- Data type: ANOM.
- Area: Lat: -90 – 90, Lon: 0 – 360 (90°S – 90°N, 0° – 360°E).
- Level: 850hPa.
- Time unit: MONTHLY, Check “Year-to-year” box.
- Showing period: RANGE, 1958 – 12 for upper box and 2016 – 2 for lower box.

“Data2” box will appear after selecting “COMPOSITE” in “Analysis method” pull-down menu.

◆ Data2

- Dataset: INDEX
- Element: NINO.3
- Data type: ANOM. > 0.5
- Area: Lat: -90 – 90, Lon: 0 – 360.
 - ✧ Check “Ave” box in both “Lat” and “Lon” fields to calculate area-averaged value. Actually, any area is OK in this case because INDEX depends on only time.
- Level: 1
- Time unit: MONTHLY, Check “Year-to-year” box.

◆ Graphic Options

- Drawing: SHADE
- Color Table: Blue – Red.
- Set Contour Parameters for data1: interval:0.5, min:-3, max:3.

Finally, click “Analysis Data Submit” button and the image will be displayed.

Data1

Dataset	Element	Data type	Area	Level	Time unit	Showing period
JRA-55	Pressure Levels U (Zonal Wind) [m/s]	ANOM	ALL Lat: -90 - 90 Ave <input type="checkbox"/> Lon: 0 - 360 Ave <input type="checkbox"/>	850hPa	MONTHLY <input type="checkbox"/> Ave <input checked="" type="checkbox"/> Year-to-year <input type="checkbox"/> Time filter	RANGE 1958 - 2016 12 - 2
<input type="checkbox"/> Vector <input type="checkbox"/> SD Derivative: <input type="checkbox"/> lon <input type="checkbox"/> lat						
Analysis method: COMPOSITE						

Data2

Dataset	Element	Data type	Area	Level	Time unit
INDEX	NINO.3 <input type="checkbox"/> SD	ANOM > 0.5	ALL Lat: -90 - 90 Ave <input checked="" type="checkbox"/> Lon: 0 - 360 Ave <input checked="" type="checkbox"/>	1	MONTHLY <input type="checkbox"/> Ave <input checked="" type="checkbox"/> Year-to-year <input type="checkbox"/> Time filter

Graphic Options

Colorizing: COLOR	<input checked="" type="checkbox"/> Show Contour Labels
Drawing: SHADE	<input checked="" type="checkbox"/> Show Color Bar
Image Format: png	<input checked="" type="checkbox"/> Set Contour Parameters for data1
Font: default	interval: 0.5 min: -3 max: 3
Color Table: Blue - Red	<input type="checkbox"/> Set Vector size: [inch] value: skip: 1

Fig. 23a Parameter setting and graphic options to draw the composite map of 3-month mean zonal wind anomaly for December-January-February when SST anomaly in NINO.3 region was larger than +0.5.

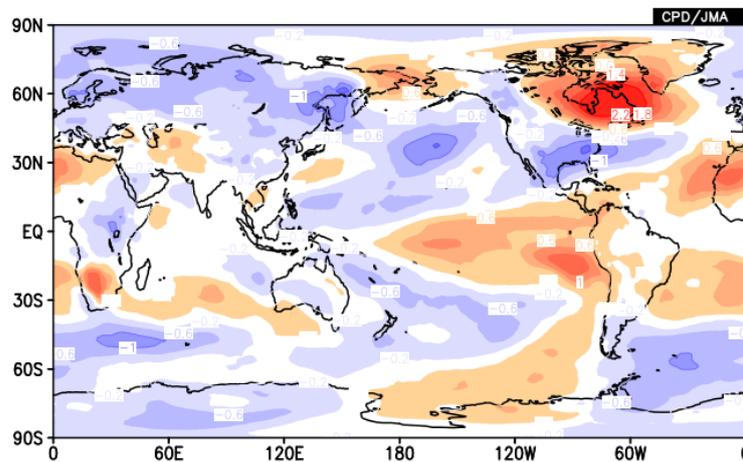


Fig. 23b The composite map of zonal wind anomaly at 850hPa in January created by the setting shown in Fig. 23a

If you know the compositing years already, you can make composite maps by another way. As an example, let's make the same figure as Fig. 23b again but by the different method.

◆ Data1

- Dataset: JRA-55.
- Element: Pressure Levels → U (Zonal Wind).
- Data type: ANOM.
- Area: Lat: -90 – 90, Lon: 0 – 360 (90°S – 90°N, 0° – 360°E).
- Level: 850hPa.
- Time unit: MONTHLY, Check “Ave” and “Year-to-year” box.
- Showing period: YEARS. Input following years in each small box or in large box with comma-separated

or space-separated format.

- ✧ 1965,1969,1972,1976,1982,1986,1987,1991,1994,1997,2002,2006,2009,2015
- ✧ Select 12 for the left box and 2 for the right box.

◆ **Graphic Options**

- Drawing: SHADE
- Color Table: Blue – Red.
- Set Contour Parameters for data1: interval:0.5, min:-3, max:3.

Finally, click “Analysis Data Submit” button and the image will be displayed.

Graphic Options

Fig. 24 Parameter setting and graphic options to draw the composite map of 3-month mean zonal wind anomaly for December-January-February when SST anomaly in NINO.3 region was larger than +0.5. This is another way for making Fig. 23b.

If you want to draw composite map using anomaly relative to default (1981-2010 mean) climatology, significance of the composite anomaly can be drawn using “SIGNIFICANCE_TEST” analysis.

◆ **Data1**

- Time unit: MONTHLY, Check “Year-to-year” box.
- Other parameters are the same as previous example.

“Data2” box will appear after selecting “SIGNIFICANCE_TEST” in “Analysis method” pull-down menu.

◆ **Data2**

- Showing period: RANGE, 1981 – 7 for upper box and 2010 – 7 for lower box.
- Significance: 95% (two side).
- Other parameters are the same as Data1.

◆ **Graphic Options**

- Drawing: CONTOUR
- Color Table: Blue – Red.
- Set Contour Parameters for data1: interval:0.4, min:-2.2, max:2.2.

Finally, click “Analysis Data Submit” button and the image will be displayed.

Data1

Dataset	Element	Data type	Area	Level	Time unit	Showing period
JRA-55	Pressure Levels ψ (Stream Function)	ANOM	ALL Lat: -90 - 90 Ave <input type="checkbox"/> Lon: 0 - 360 Ave <input type="checkbox"/>	850hPa	MONTHLY <input type="checkbox"/> Ave <input checked="" type="checkbox"/> Year-to-year <input type="checkbox"/> Time filter	YEARS <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> input years directly (comma-separated or space-separated) 1964,1966,1969,1970,1973,1977,1983,1987,1988,1992,1998,2003,2010 7 - 7

Vector SD
Derivative: lon lat

Analysis method: SIGNIFICANCE_TEST

Data2

Dataset	Element	Data type	Level	Time unit	Showing period	Significance
JRA-55	Pressure Levels ψ (Stream Function)	ANOM	850hPa	MONTHLY <input type="checkbox"/> Ave <input checked="" type="checkbox"/> Year-to-year <input type="checkbox"/> Time filter	RANGE 1981 - 2010 7 - 7 remove data1 period <input type="checkbox"/>	95%(two side)

Graphic Options

Colorizing: COLOR	<input checked="" type="checkbox"/> Show Contour Labels
Drawing: CONTOUR	<input checked="" type="checkbox"/> Show Color Bar
Image Format: png	<input checked="" type="checkbox"/> Set Contour Parameters for data1
Font: default	interval: 0.4 min: -2.2 max: 2.2
Color Table: Blue - Red	<input type="checkbox"/> Set Vector size: [inch] value: skip: 1

Detailed Options for Image 1

Fig. 25a Parameter setting and graphic options to draw composite of stream function anomaly at 850hPa in July for “post-El Niño” years.

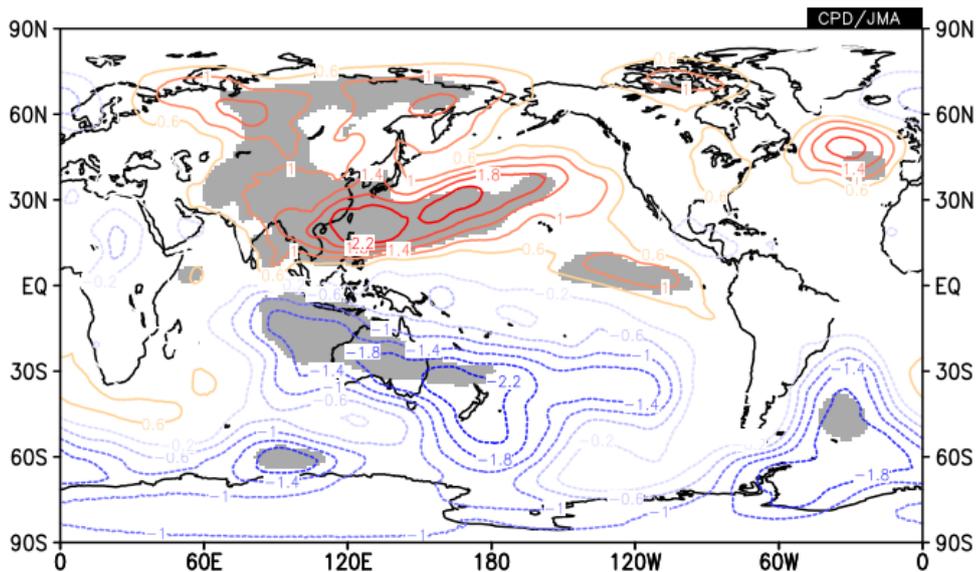


Fig. 25b Image of stream function anomaly at 850hPa composite for post-El Niño July with 95% confidence level marked by gray shading created by the setting shown in Fig. 25a

CAUTION: In “SIGNIFICANCE_TEST” analysis, “DATA1” is composited, not “difference between DATA1 and

DATA2”. Therefore, if you want to draw composite map of deviation relative to reference value except default climatology (1981-2010 average), this method cannot be used. For example, if you change the period of DATA2 to “1971-7 to 2000-7” in the above example, the contour would indicate the composite of anomaly (relative to 1981-2010 climatology) and gray shading would indicate the significance of deviation relative to 1971-2000 average, and the figure might be misleading.

6. Other Functions

6.1. Data download

Users can download data as a zip compressed file with a plain text and GrADS format (control file and data file) to create a map. If you need to download data, please click “Download text zip file” or “Download data (ctl file and 4-byte-binary data)” at the bottom of the iTacs window.

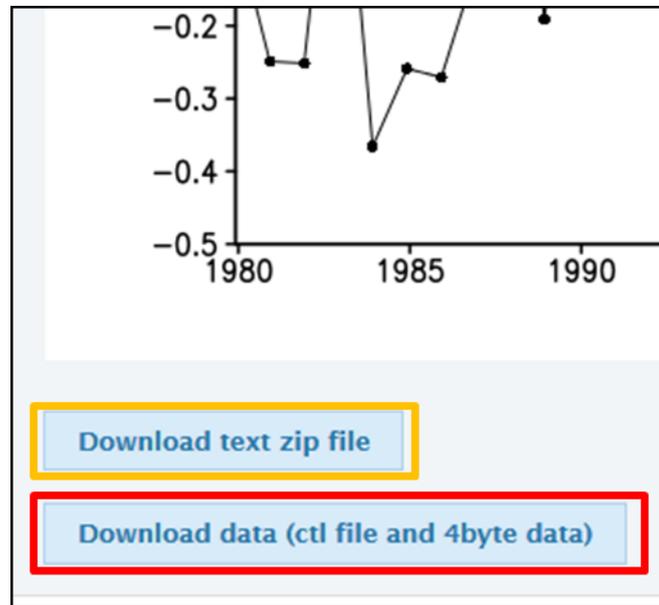


Fig. 26 A screen image of banners for download data from iTacs.

6.2. Using user input data

Time series data made by individual users can be used in a dataset name “USER_INPUT”. Using this function, a correlation or regression coefficient map between single station data or user’s original index and another dataset like JRA-55 can be created, for example.

There are two ways to set data, which can be selected in element name.

- UPLOAD_TXT : Data are given by an uploaded text file.
- INPUT_DATA : Data are directly input to the box.

Input data must be consists of one element, separated by comma and given by specified format as “year, month, day, value”. In case of monthly data, “day” is always given as “1”. Similarly, in case of annual data, “month” and “day” are always given as “1”.

Sentences beginning with “#” have special meanings as below. Other sentences beginning with “#” is taken as comments and not used in analysis.

- “#undef =” : Definition of missing data (default is -9999). For example, “#undef=-9999”
- “#element =” : Data name used to save them on the server. For example, “#element=Daily Ts”

Data1

Dataset	Element	Input_txt	Time unit	Showing period
USER_INPUT	INPUT_DATA	1979, 6, 1, 3.282446 1979, 7, 1, -5.284460 1979, 8, 1, 0.555721 1979, 9, 1, -2.454476 1979, 10, 1, 1.289361 1979, 11, 1, -7.725315 1979, 12, 1, 0.010303 1980, 1, 1, 4.623410 1980, 2, 1, -5.407787 1980, 3, 1, -8.023340 1980, 4, 1, -13.504246 1980, 5, 1, -5.164176 1980, 6, 1, -6.316853 1980, 7, 1, -6.181804 1980, 8, 1, -0.870708 1980, 9, 1, -1.255426 1980, 10, 1, -1.512890 1980, 11, 1, -0.324971 1980, 12, 1, -8.659753	MONTHLY <input type="checkbox"/> Ave <input type="checkbox"/> Year-to-year <input type="checkbox"/> Time filter	RANGE 1980 1 1990 12

Vector SD
 Derivative: lon lat

Fig. 27 A screen image when monthly time series data is input with "INPUT_DATA".