TCC Training Seminar on
One-month Forecast Products

7 – 9 November 2011

Tokyo, Japan

Tokyo Climate Center
Japan Meteorological Agency

TCC Training Seminar on
One-month Forecast Products

7 – 9 November 2011
Tokyo, Japan

Contents

- Schedule of the Training Seminar **i**

(Item 3)

- How to use JMA's web-based application for climate analysis **1**

(Item 4)

- JMA’s Ensemble Prediction System (EPS) for Seasonal Forecasting **25**

- Using Grid Point Value Data Provided on the TCC Website **35**

(Item 6)

- Outline of Guidance and Related Evaluation **49**

- Producing Site-specific Guidance Using Domestic Data **57**

Schedule

TCC Training Seminar on One-Month Forecast Products

Tokyo, Japan, 7 – 9 November 2011

Draft Schedule

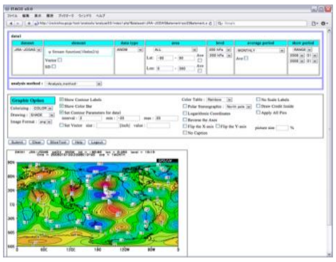
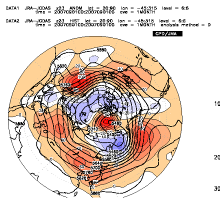
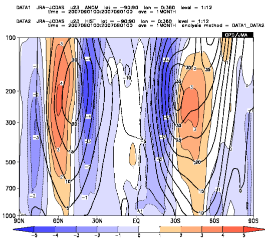
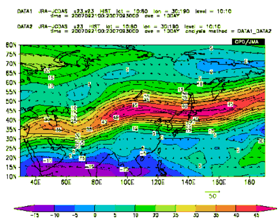
Day 1	Program	
10:00 – 10:30	1. Opening session Welcome address Self-introduction by participants Group photo Courtesy call on JMA's Director-General	
10:30 – 10:45	2. Introduction - Objectives and structure of the training seminar - Products available on the TCC website	Lecture (TCC)
10:45 – 13:00	3. Climate Analysis - Atmospheric Circulation Analysis for Seasonal Forecasting - How to use JMA's web-base application for climate analysis (ITACS: Interactive Tool for Analysis of the Climate System)	Lecture and Exercise (Analysis Unit)
13:00 – 14:30	Lunch	
14:30 – 16:00	4. JMA Ensemble Prediction System (EPS) for seasonal prediction - Outline of the JMA operational EPS - Model products on the TCC website - How to download and use the gridded data (grid point values; GPVs) of the JMA numerical models	Lecture (Numerical Prediction Unit)
16:15 – 18:15	5. Seasonal Forecasting - Predictability of the model - How to interpret model output - How to issue seasonal forecast in JMA	Lecture (Forecast Unit) and Exercise (guided by Forecast Unit)
18:45 – 20:00	Dinner	Invitation by JMA
Day 2		
9:30 – 12:30	6. Producing Guidance and Verification - What is seasonal forecast guidance - Verification of probabilistic forecasts Exercise :Producing site-specific guidance using domestic data	Lecture (Forecast Unit) and Exercise (guided by Forecast Unit)
12:30 – 13:45	Lunch	
13:45 – 15:30	6. Producing Guidance and Verification (exercise continued)	Exercise (guided by Forecast Unit)
15:45 – 18:00	7. Preparing one-month forecast Exercise: Preparing one-month forecast using the guidance prepared in session 6	Exercise (guided by Forecast Unit)
Day 3		
9:30 – 12:30	7. Preparing one-month forecast (exercise continued) Preparing presentations to be given in session 8	Exercise (guided by Forecast Unit)
12:30 – 14:00	Lunch	
14:00 – 17:45	8. Presentations and Discussion Presentations on one-month forecast for November 2011	(All the participants)
17:45 – 18:00	9. Closing session	

How to use JMA's web-based application for climate analysis

How to use JMA's web-based application for climate analysis

ITACS: Interactive Tool for Analysis of the Climate System

ITACS is the Interactive Tool for Analysis of Climate System since 2007. The aims are analyzing the causes of climate events and monitoring current climate status. And, the system consists of Web interface, programs, GrADS and data files on the web server.

Data sets

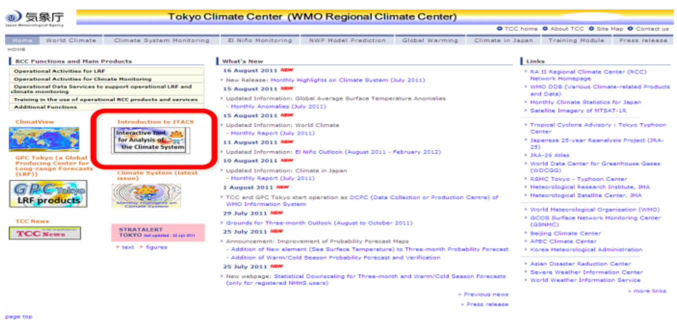
Atmospheric analysis data / Outgoing Longwave Radiation (OLR) data / Sea surface temperature analysis data / ocean analysis data / CLIMAT reports …

Application and contact

There is application form and introduction about ITACS on the homepage of Tokyo Climate Center.

http://ds.data.jma.go.jp/tcc/tcc/

tcc@met.kishou.go.jp



Tokyo Climate Center homepage

ITACS

Climate Prediction Division, JMA

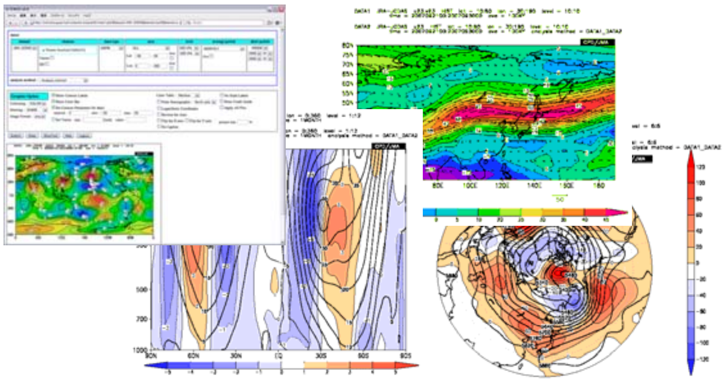
menu

- What's ITACS
- Data
- Application to use
- Exercise and learning by using ITACS

What's ITACS

ITACS is the Interactive Tool for Analysis of Climate System since 2007.

Aim:
Analyzing the causes of climate events and monitoring current climate status.
System:
Web interface + programs(Ruby, Gphys…) + GrADS + data files on the web server



data

- CLIMAT
  - Monthly world climate data derived from CLIMAT messages via the GTS line from WMO Members around the world.
- INDEX
  - El Nino Monitoring Indices consisting of monthly mean Sea Surface Temperature produced by COBE-SST.
- JRA-JCDAS
  - Atmospheric circulation data produced by JMA's Climate Data Assimilation System (JCDAS), which is consistent quality with Japanese 25-year reanalysis (JRA-25).
- MOVE-G
  - Oceanic assimilation produced by the system operated by JMA.
- SAT
  - Outgoing Longwave Radiation (OLR), which is derived from observations by NOAA's polar orbital satellites, and provided by Climate Prediction Center (CPC) in the National Centers for Environmental Prediction (NCEP) of the National Oceanic and Atmospheric Administration (NOAA).
- SST
  - Sea Surface Temperature produced by the system operated by JMA (COBE-SST).

Application to use

There is banner link about application to use ITACS in the TCC homepage:
http://ds.data.jma.go.jp/tcc/tcc/index.html

Exercise

Now, let's access and use ITACS. Using it will help you to understand ITACS.

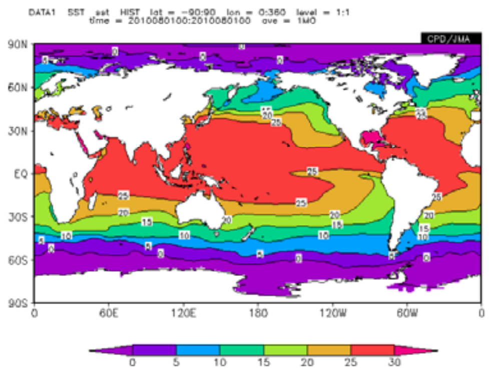
Sample image ⇄ Making image like sample image by yourself

⇩ If we have enough time after every exercises…

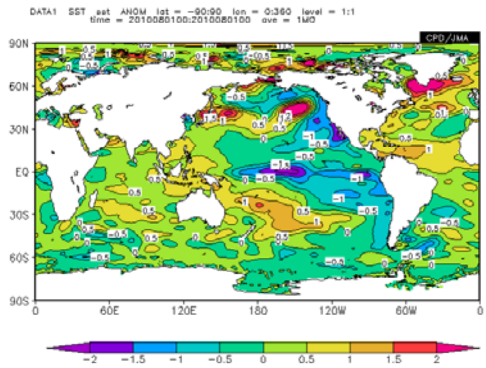
Let's try to change sample image as you like (self study)

Let's start exercise. Please access to following site. This site will be linked from TCC homepage soon:
http://extreme.kishou.go.jp/tool/itacs-tcc2011/

Sea surface temperature(SST)



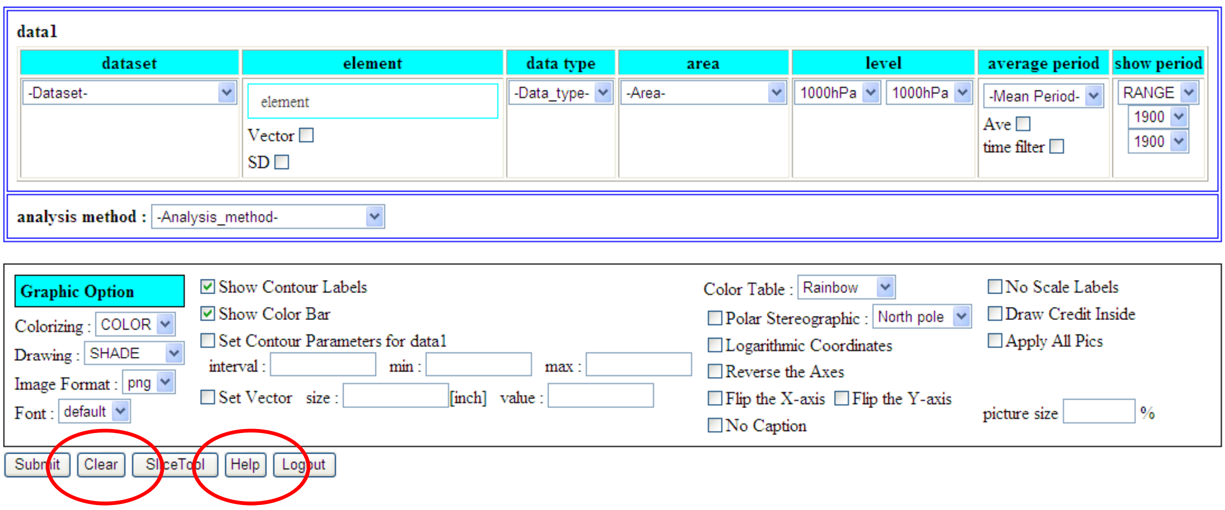
Sea Surface Temperature in August 2010



Sea Surface Temperature anomaly in August 2010

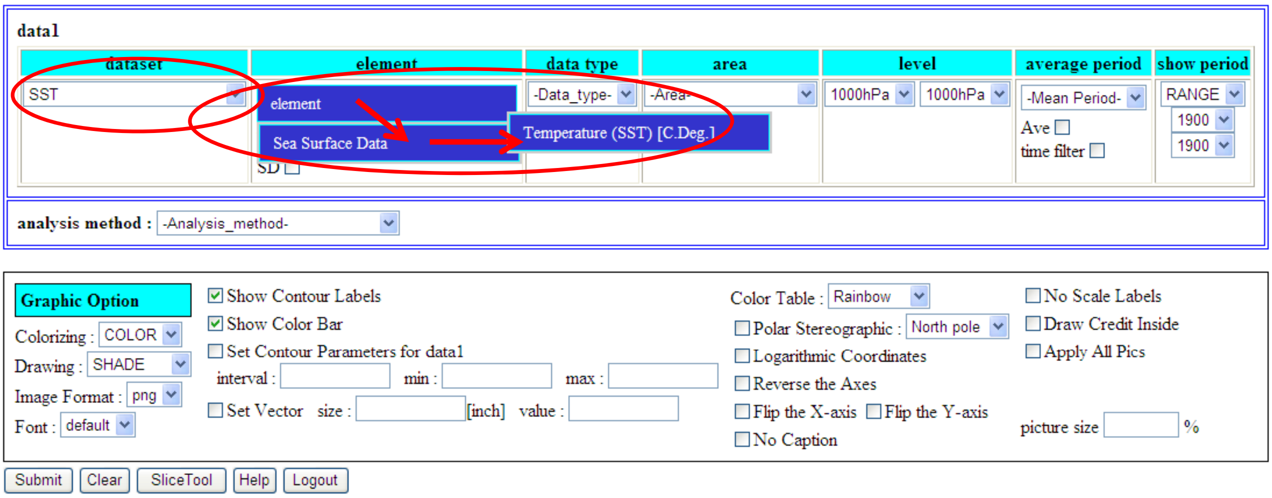
This is tutorial for making a map of Sea Surface Temperature(SST) and its anomaly.
Let's know basic use of ITACS.

Sea surface temperature(SST)



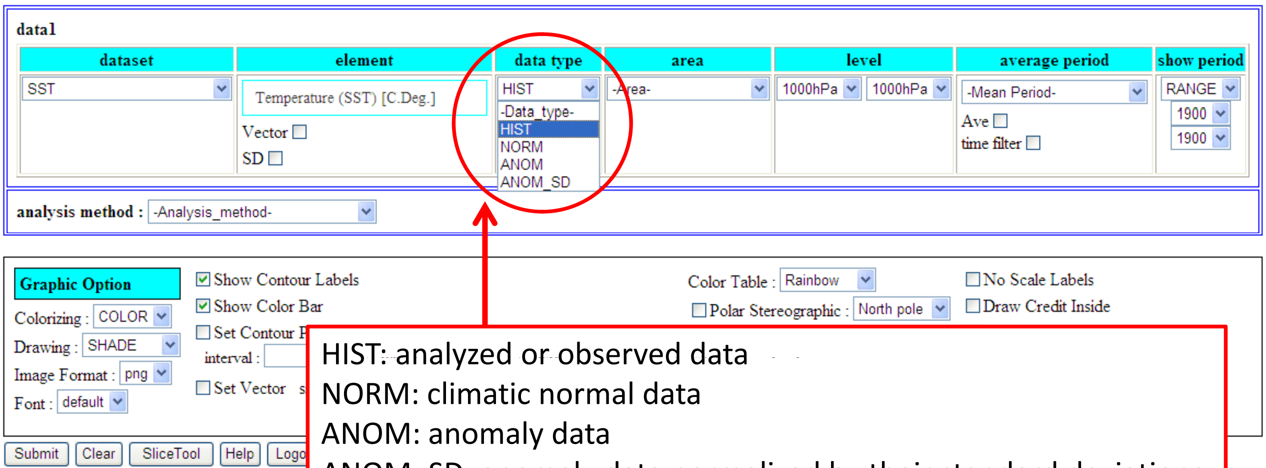
This is default screen of ITACS. Click "Clear" button if you need default screen.
"Help" button gives you help page.

Sea surface temperature(SST)



First, select “dataset” - “SST” and its “element” - “Temperature”.

Sea surface temperature(SST)

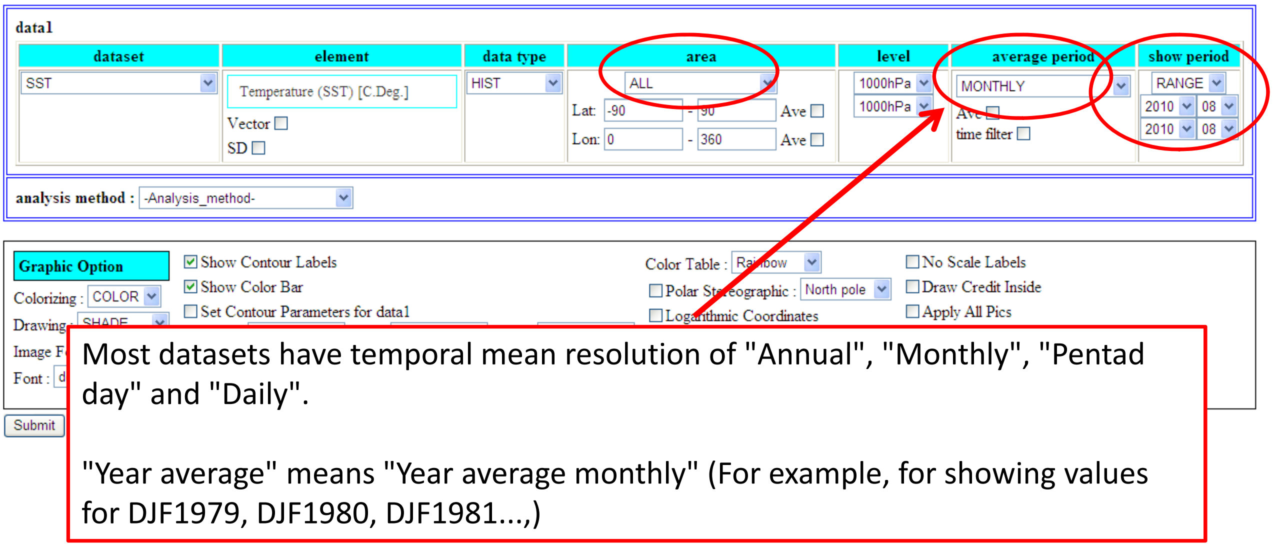


HIST: analyzed or observed data
NORM: climatic normal data
ANOM: anomaly data
ANOM_SD: anomaly data normalized by their standard deviations

Note:
"HIST" minus "NORM" is "ANOM"
"ANOM" divided by σ is "ANOM_SD"

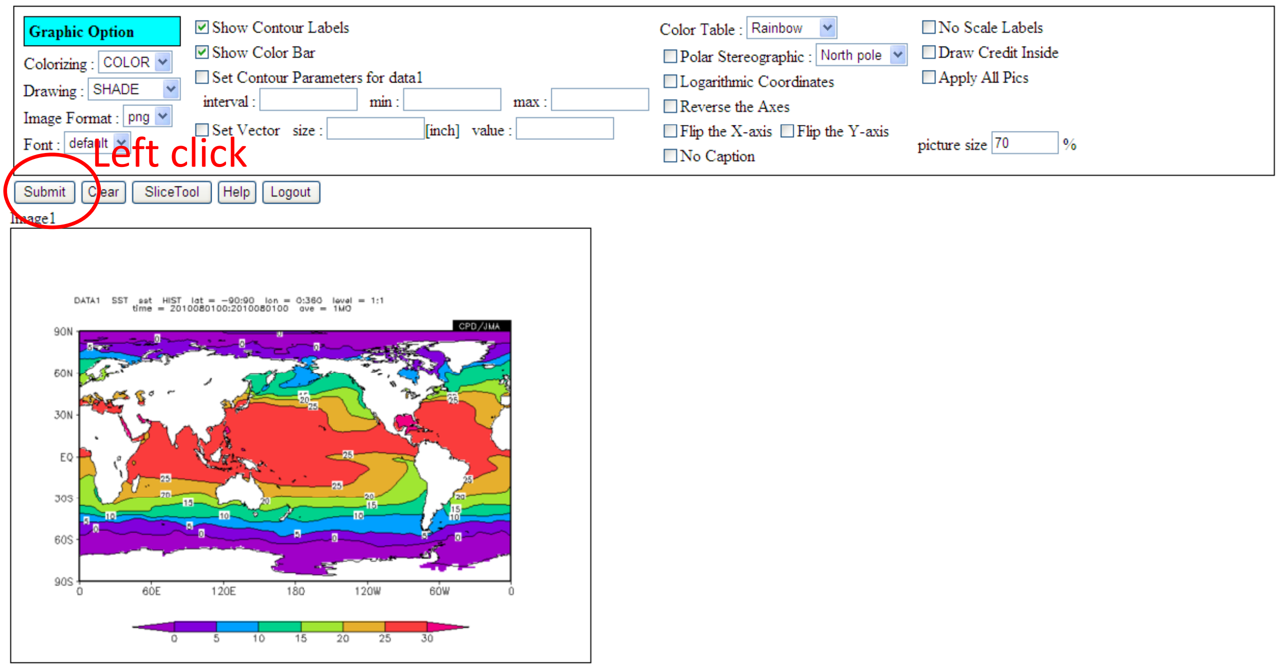
Secondly, select “data type” - “HIST”(historical data).
Please note there are some data type.

Sea surface temperature(SST)



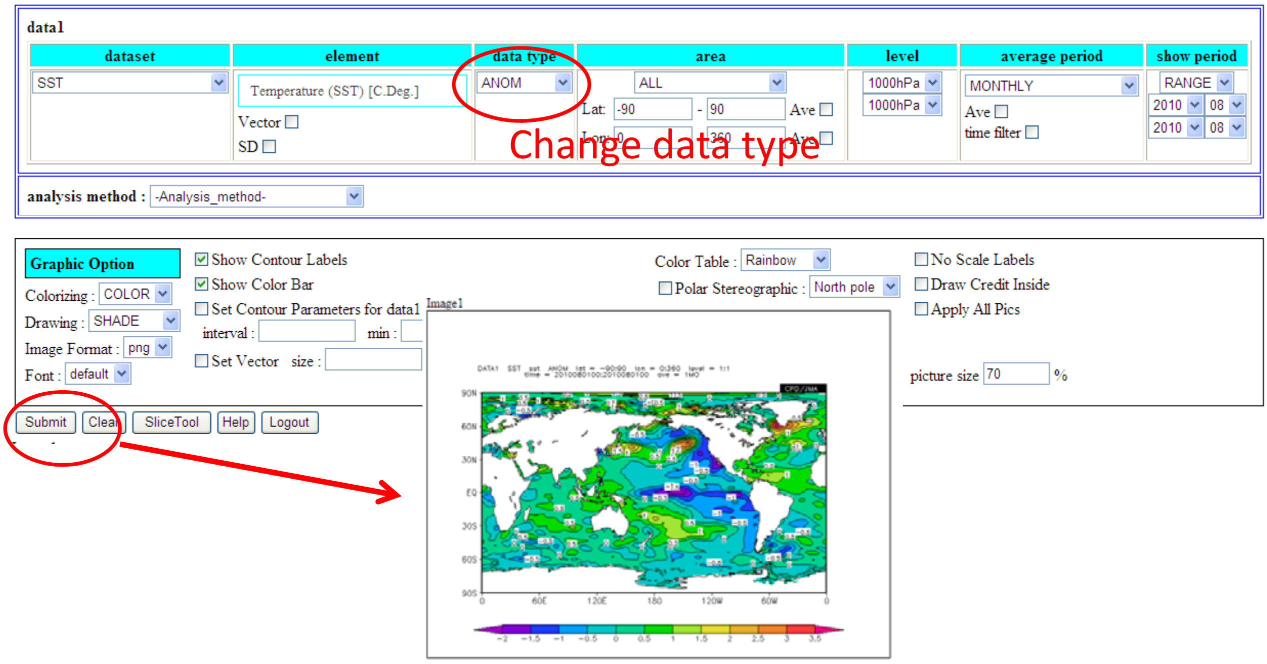
Next, select “area”, “average period” and “show period”.

Sea surface temperature(SST)



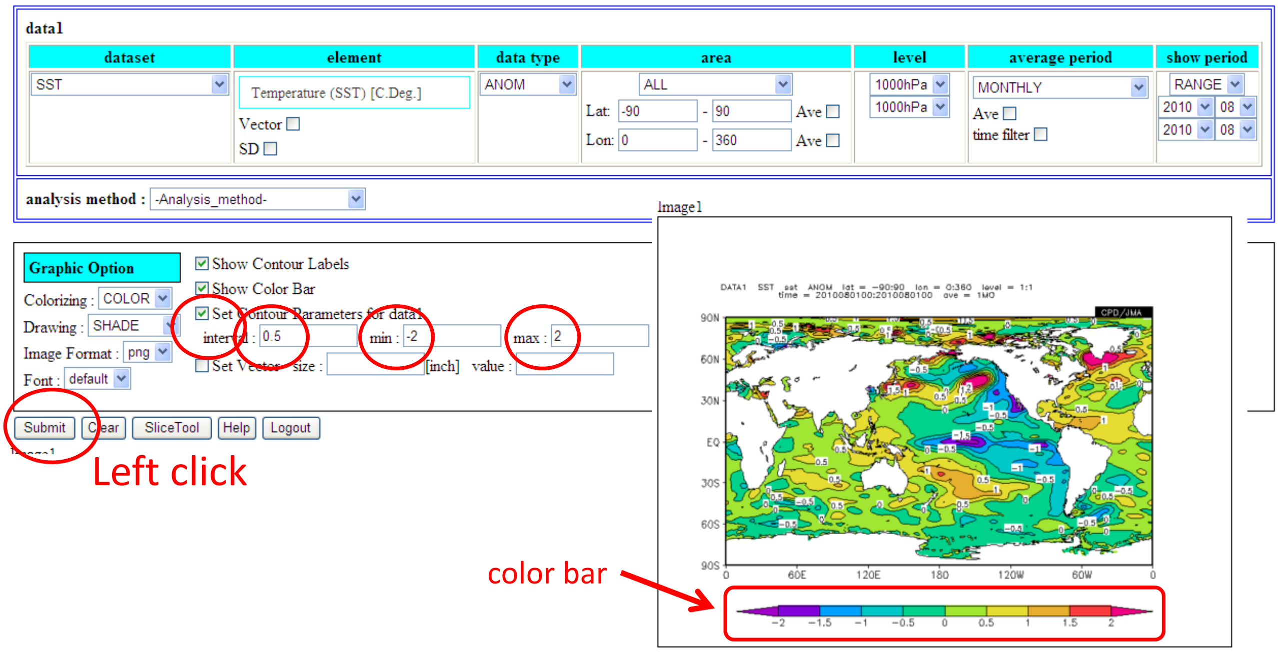
Finally, click “Submit” button. A map of Sea Surface Temperature(SST) will be made.

Sea surface temperature(SST) anomaly



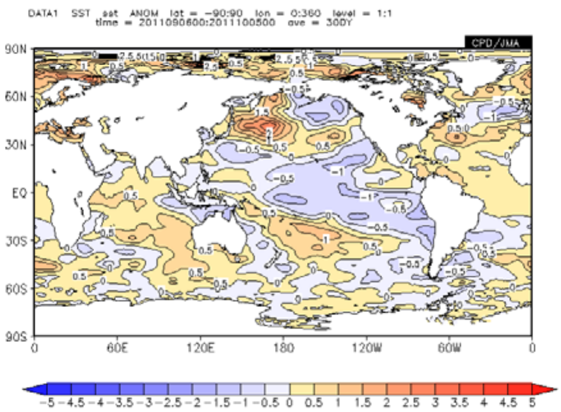
Let's change "data type" – "ANOM" to make map of SST anomaly and click "Submit".

Sea surface temperature(SST) anomaly



If you want to change the range of colors in the color bar, please use "Graphic Options". Check "Set Contour Parameters for data1" and input parameters for interval, min and max of values.

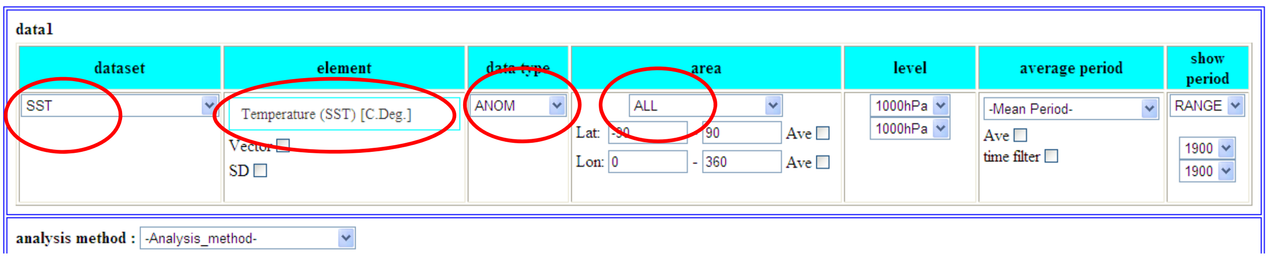
Average of SST anomaly



Average of SST anomaly <u>between 6 September and 5 October</u>

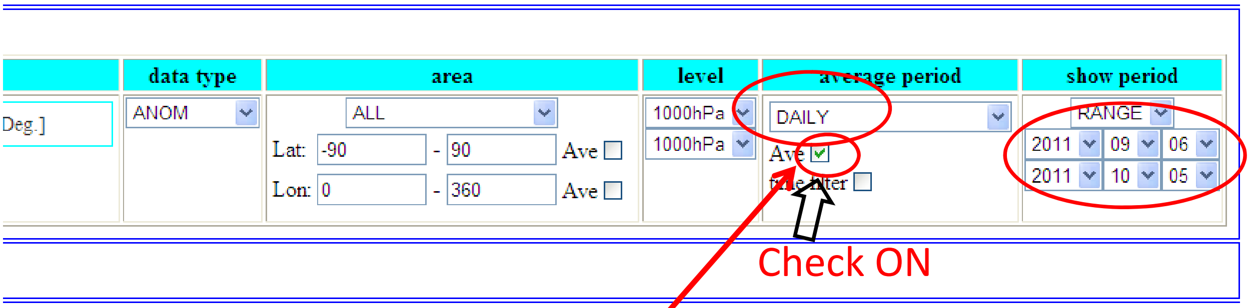
Let's know how to figure out the average of daily data.

Average of SST anomaly



First, select “dataset” - “SST” and its “element” - “Temperature”.
And, select “data type” - “ANOM”(anomaly data) and “area” – “ALL”.

Average of SST anomaly

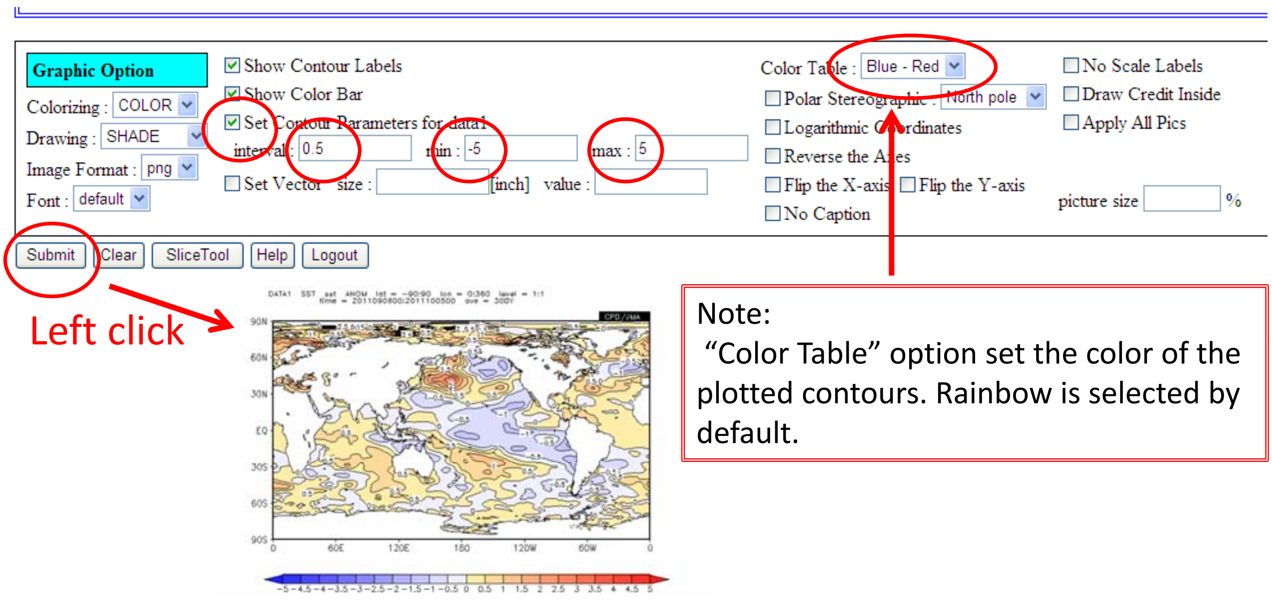


"Ave" gives average of data.

Note:
ITACS figures out the monthly data if you select "MONTHLY" in the "average period".

Next, please select "average period" – "DAILY" and check "Ave" – "ON(checked)".
And, select "show period"(2011.09.06 – 2011.10.05).

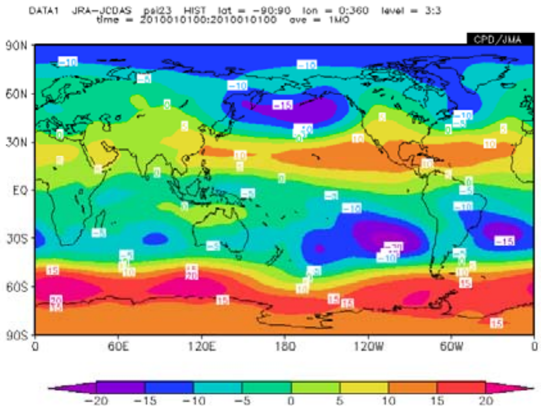
Average of SST anomaly



Note:
"Color Table" option set the color of the plotted contours. Rainbow is selected by default.

Finally, please select "Set Contour Parameters" of Graphic Option. Let's change "Color Table" if you want to set the color of the plotted contours. And, click "Submit".

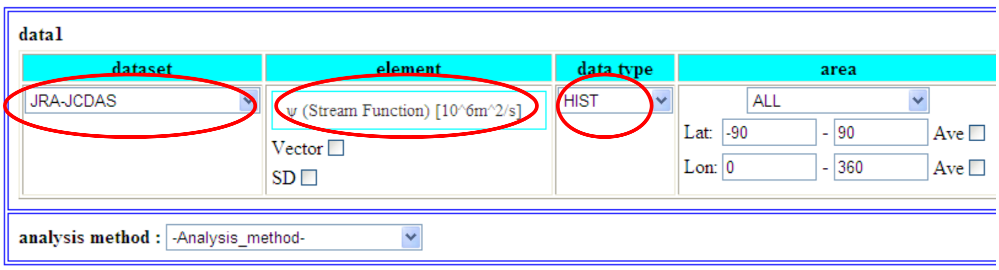
Stream function of historical data on 850hPa



Ψ(Stream function) of historical data on 850hPa

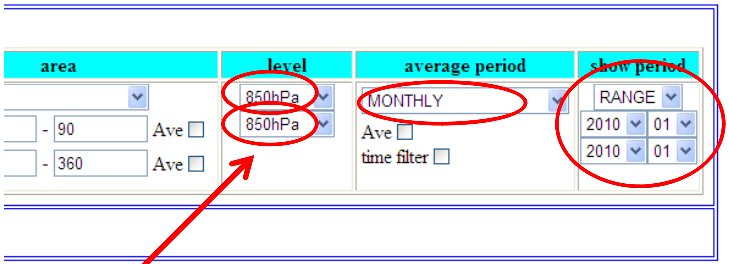
Let's know how to change vertical level of the data.

Stream function of historical data on 850hPa



First, please select "dataset" – "JRA-JCDAS", "element" – "Ψ(Stream function)" and "data type" – "HIST".

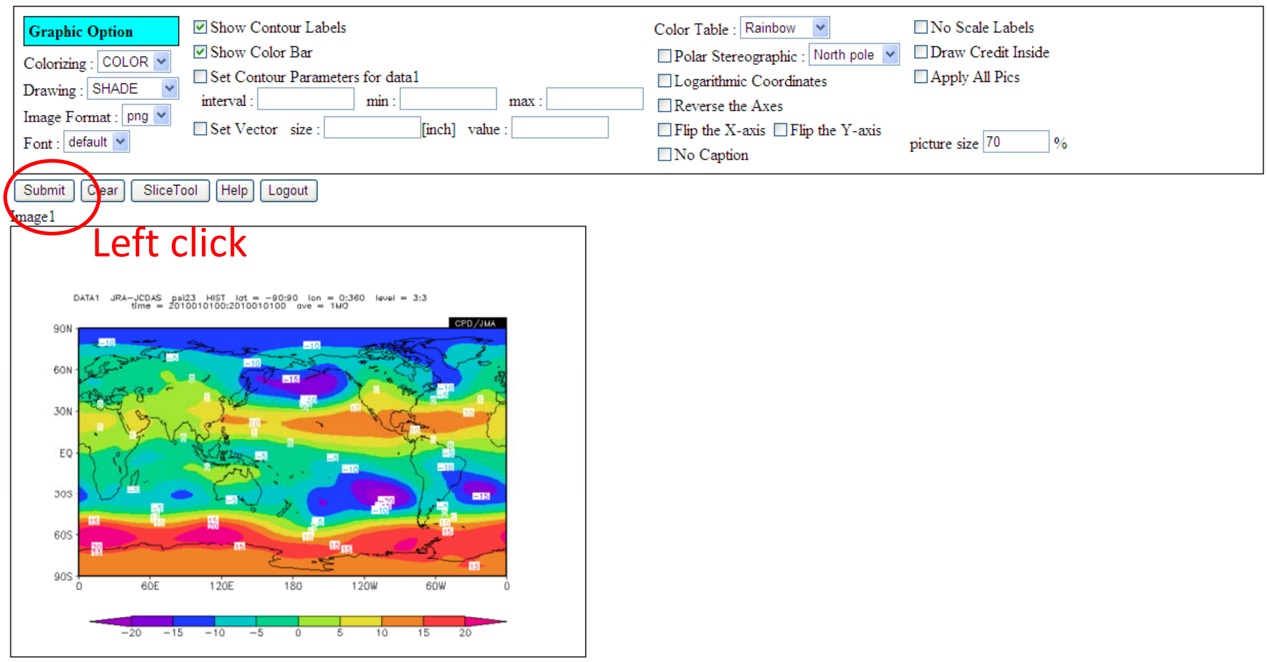
Stream function of historical data on 850hPa



Two pull-down menus are prepared in this field and available vertical levels are listed on them. If different levels are chosen from each menu by users, the drawing will be a vertical cross section chart.

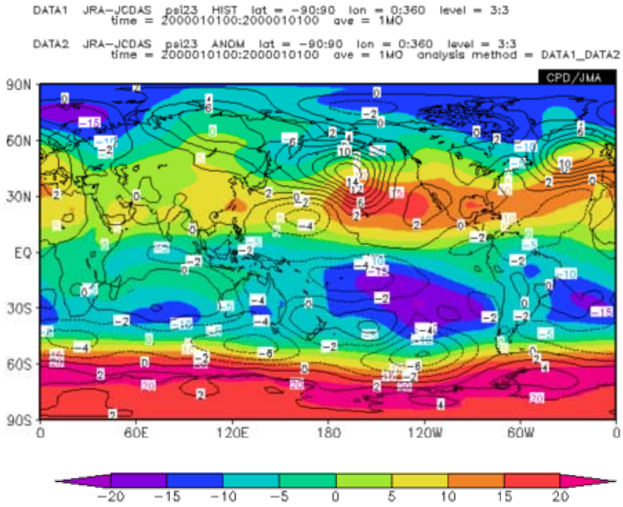
Secondly, please select "level" – "850hPa", "average period" - "MONTHLY" and "show period".

Stream function of historical data on 850hPa



Finally, click "Submit" button. A map on 850hPa will be made.

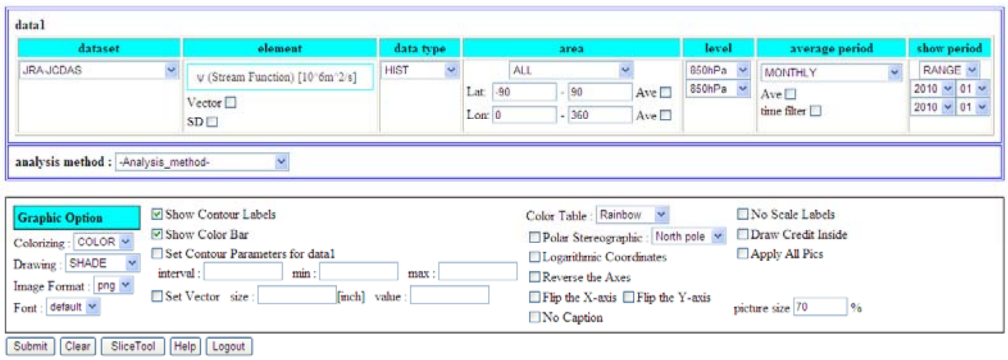
Stream function of historical data and anomaly data on 850hPa

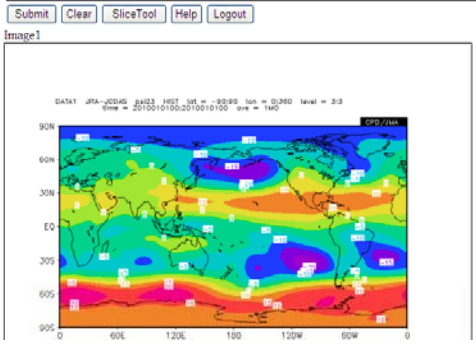


Stream function of historical data and anomaly data on 850hPa

Let's know how to superimpose a map on other map.

Stream function of historical data and anomaly data on 850hPa

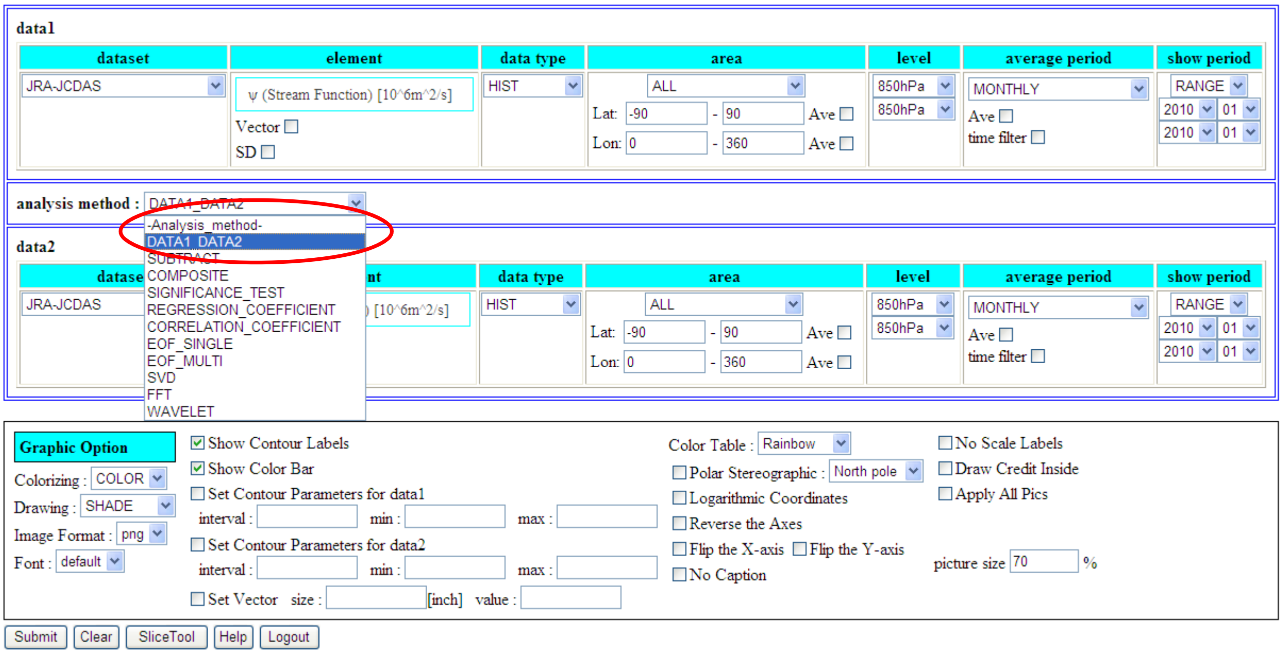




Dataset: JRA-JCDAS
Element: Ψ(Stream function)
Data type: HIST(historical data)
Area: ALL
Level: 850hPa
Average period: MONTHLY
Show period: 2010.01

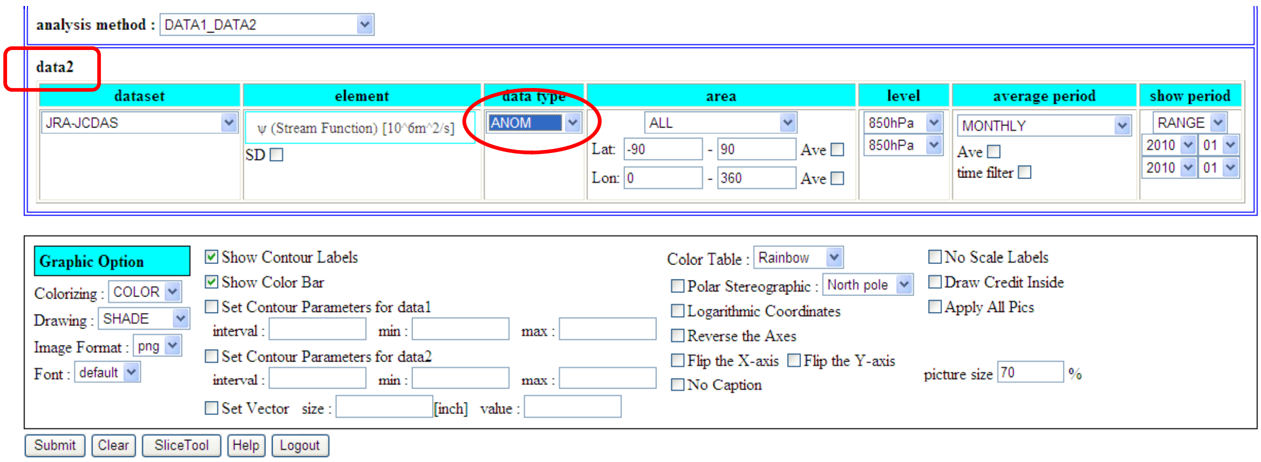
First, please draw Stream function of historical data on 850hPa.

Stream function of historical data and anomaly data on 850hPa



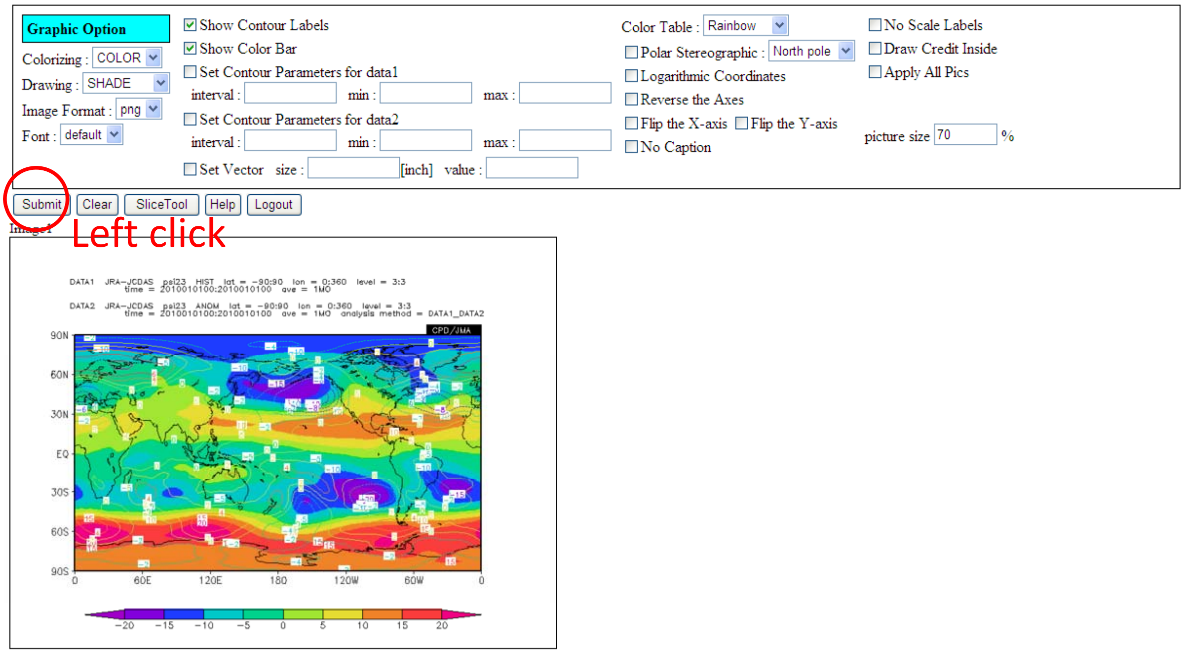
Secondly, please select “analysis method” – “DATA1_DATA2”.
Box “data2” will appear.

Stream function of historical data and anomaly data on 850hPa



And, please change “data type” – “ANOM” of box “data2”.
Don’t change other options.

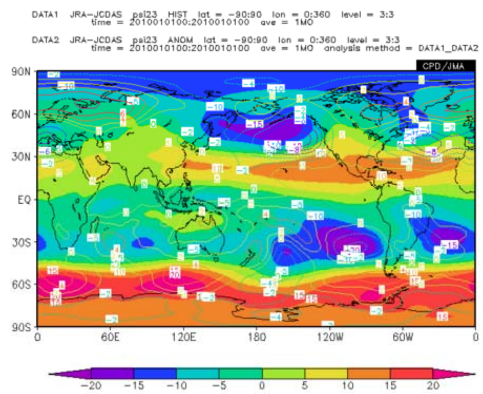
Stream function of historical data and anomaly data on 850hPa



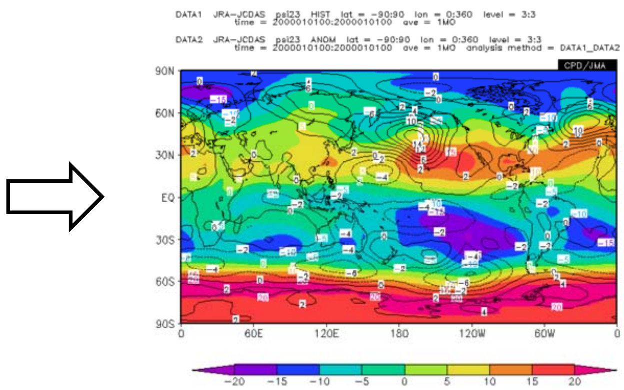
Finally, click “Submit” button and draw a map.
In addition, the color of contour can be changed… => go next page

Stream function of historical data and anomaly data on 850hPa

Before:

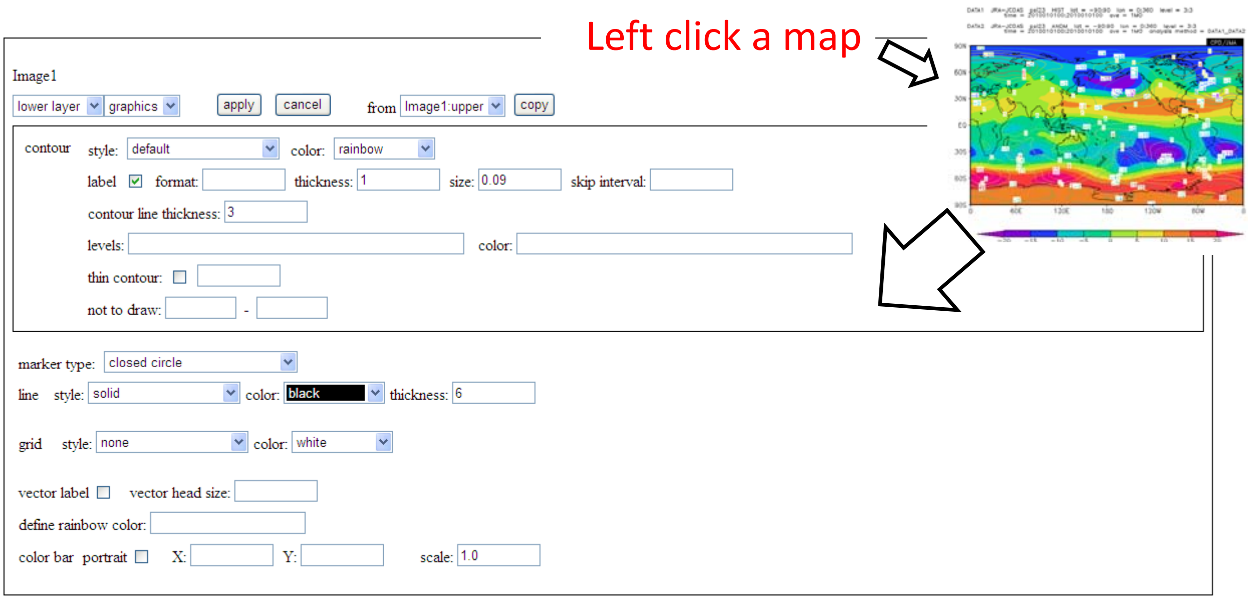


After:



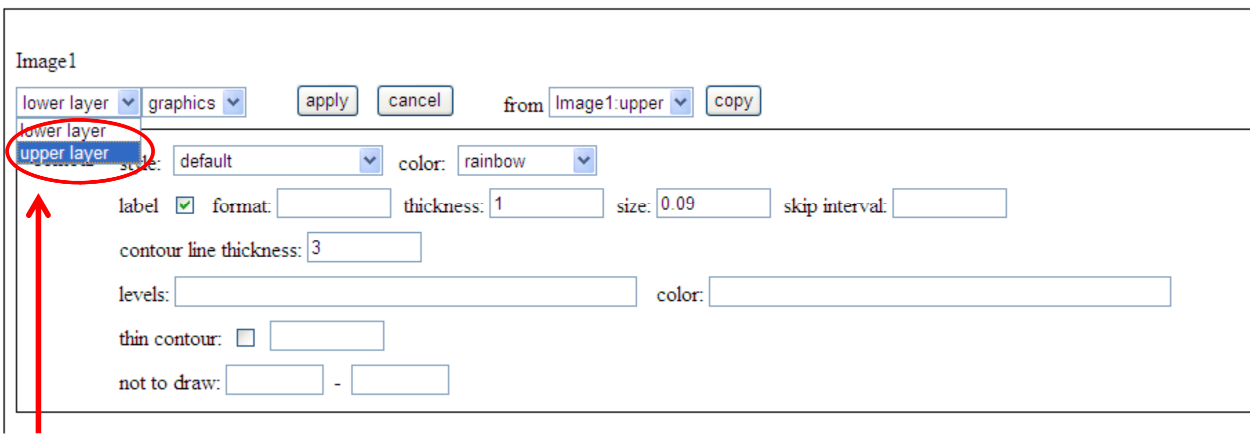
Let’s change the color of contour of upper layer.

Stream function of historical data and anomaly data on 850hPa



Please click a map. New option box "Image" will appear.

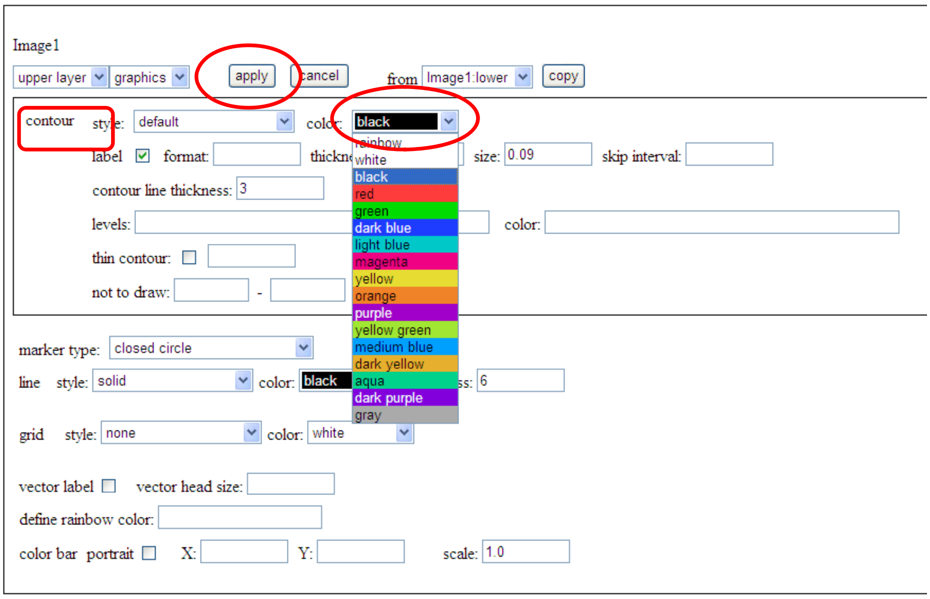
Stream function of historical data and anomaly data on 850hPa



Select layer you want to edit:
Lower layer: Data1
Upper layer: Data2

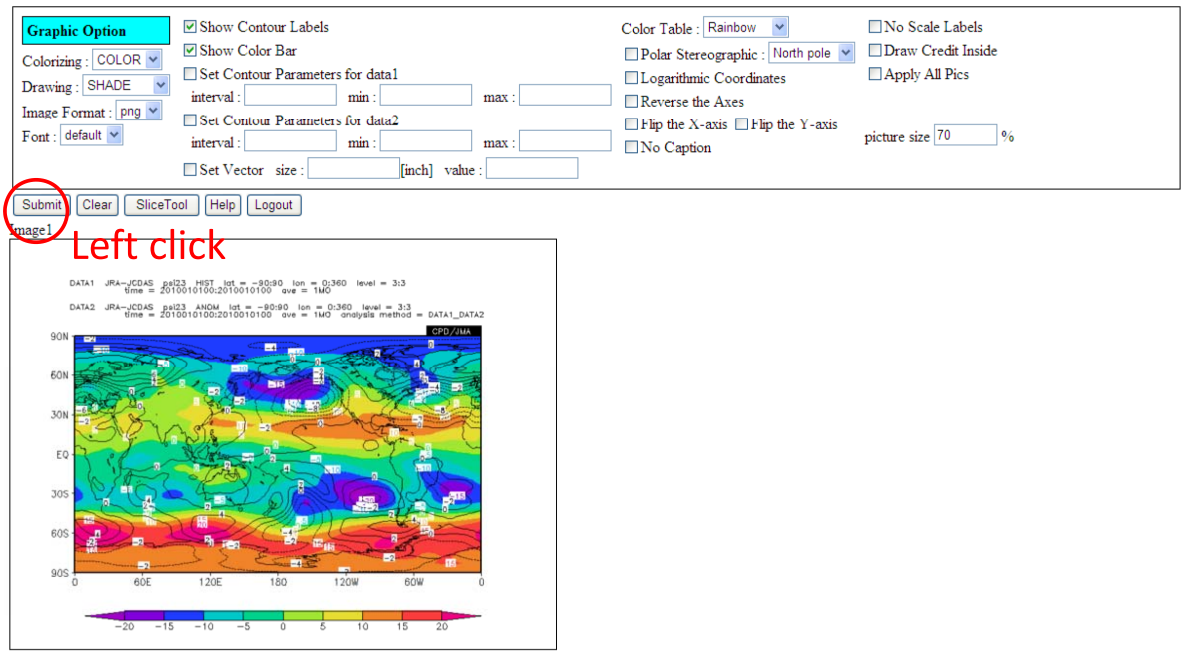
Secondly, please select layer you want to edit. In this example, select "upper layer".

Stream function of historical data and anomaly data on 850hPa



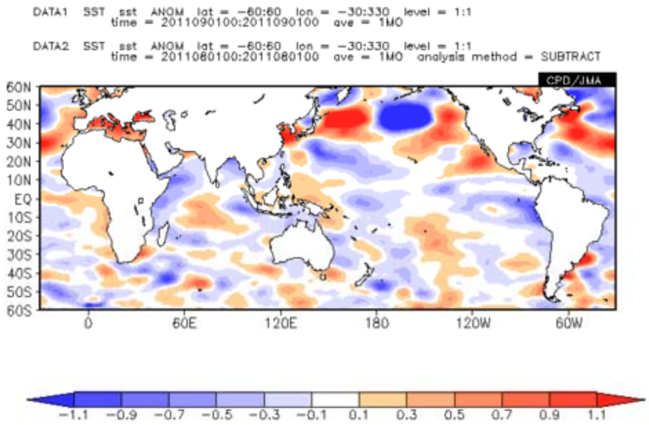
Next, please select “color” – “black” and click “apply” button.
Don’t forget to click “apply” button.

Stream function of historical data and anomaly data on 850hPa



Finally, click “Submit” button. The contour will be black.

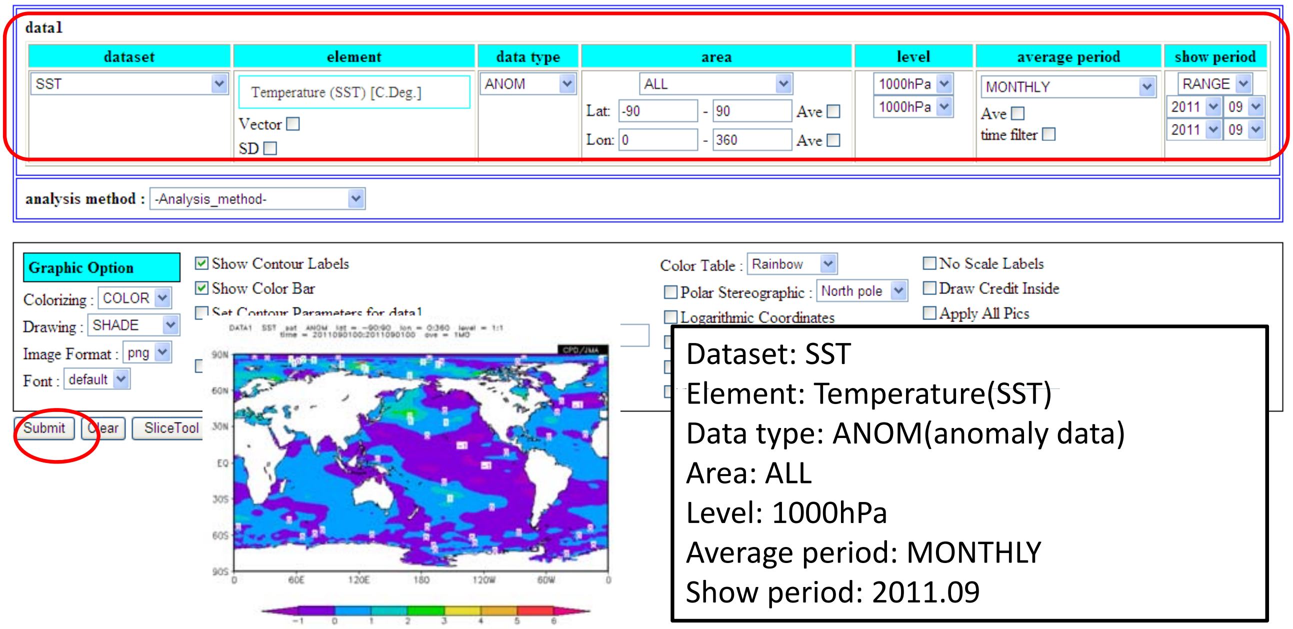
Subtraction of monthly SST



Subtraction of monthly SST between September and August

Let's know how to subtract data from data of other period.

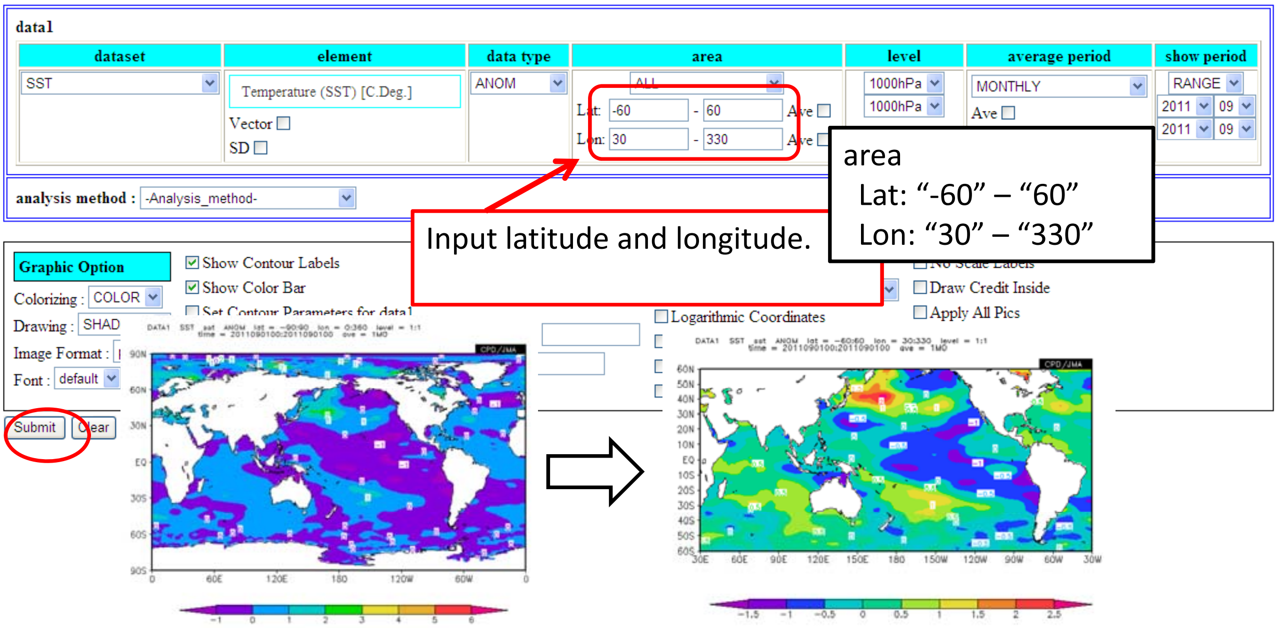
Subtraction of monthly SST



Dataset: SST
Element: Temperature(SST)
Data type: ANOM(anomaly data)
Area: ALL
Level: 1000hPa
Average period: MONTHLY
Show period: 2011.09

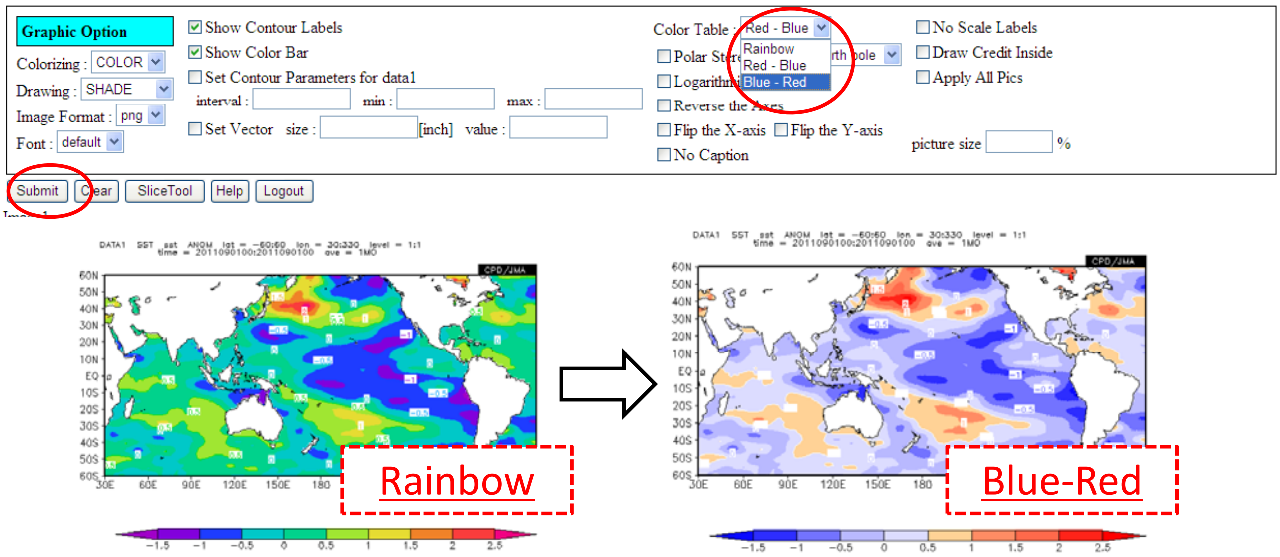
First, please make a map of monthly SST in September 2011 as mentioned above.

Subtraction of monthly SST



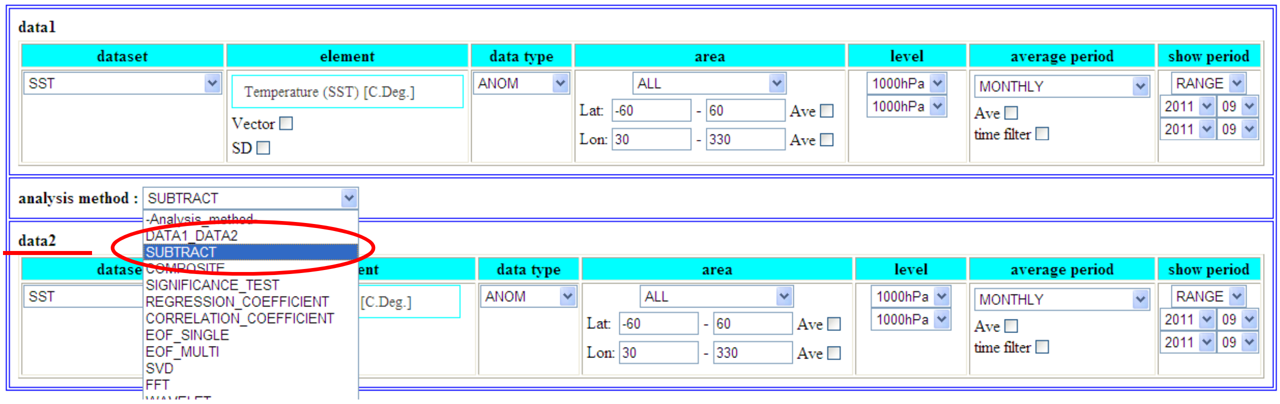
Now, let's try to change area. Please input latitude and longitude and click submit.

Subtraction of monthly SST



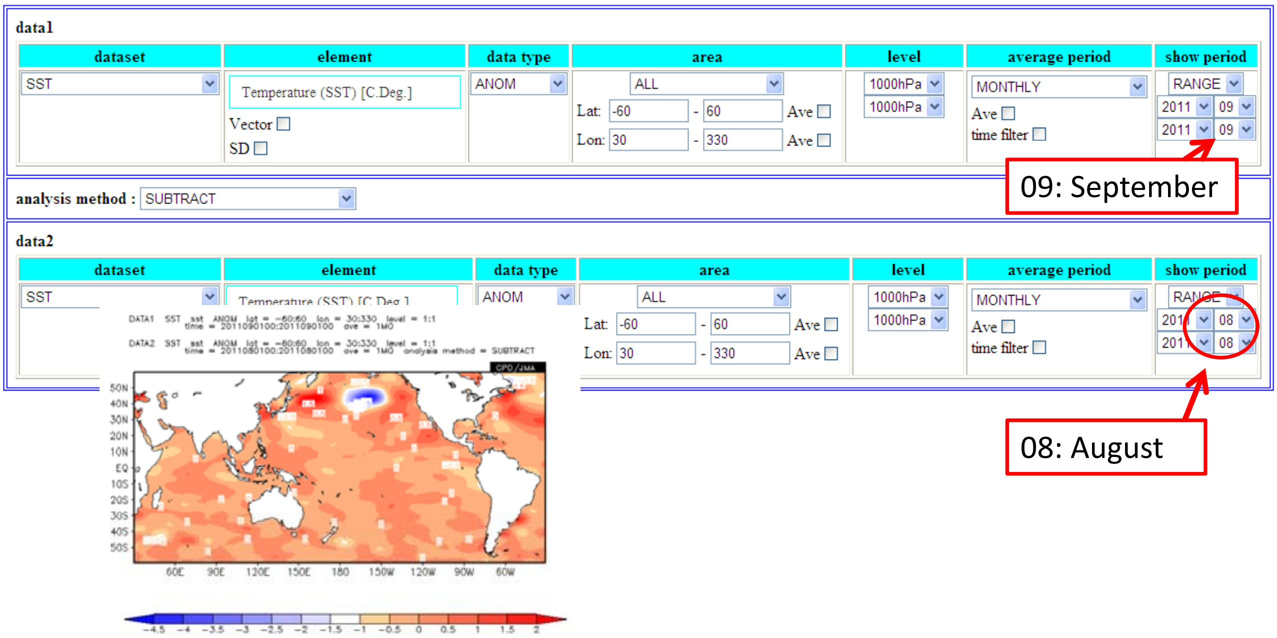
Next, let's change color table.
Please change "Color Table" – "Blue - Red" and click "Submit" button.

Subtraction of monthly SST



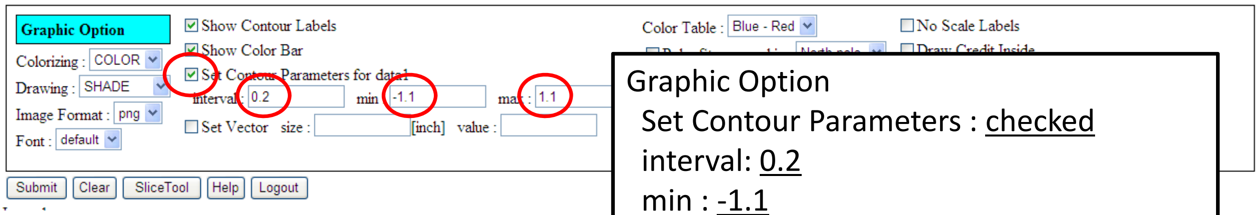
Please select “analysis method” – “SUBTRACT”.
Box “data2” will appear.

Subtraction of monthly SST

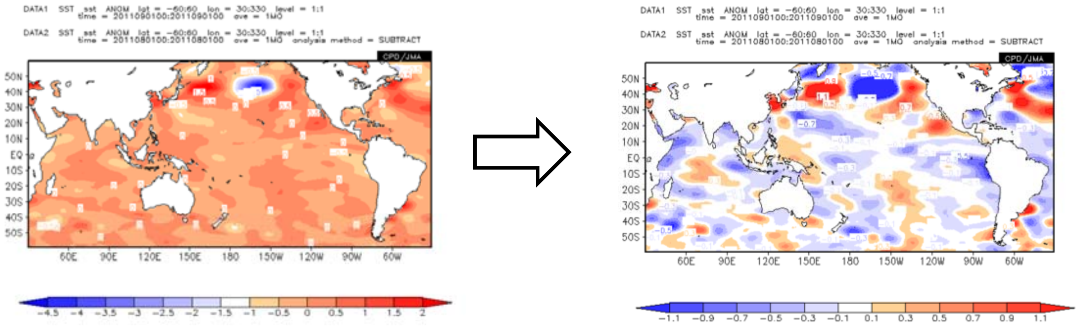


Next, please change month, in “data2” box and click “Submit” button.
Almost area of sea will be painted red. Let’s change contour parameter in next step…

Subtraction of monthly SST

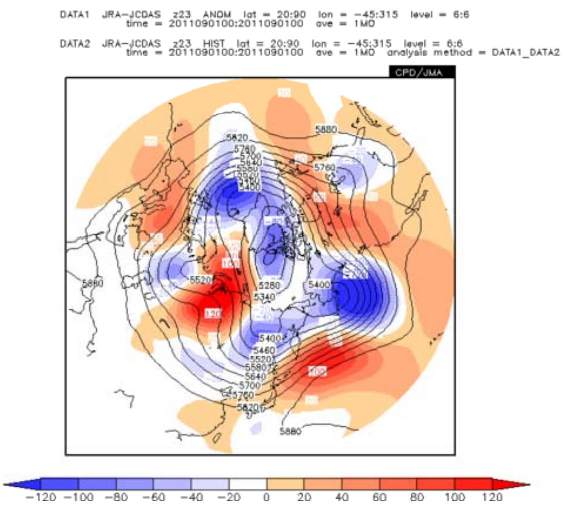


Graphic Option
Set Contour Parameters : checked
interval: 0.2
min : -1.1
max : 1.1



Finally, let's change the range of color bar to see change of SST in detail.

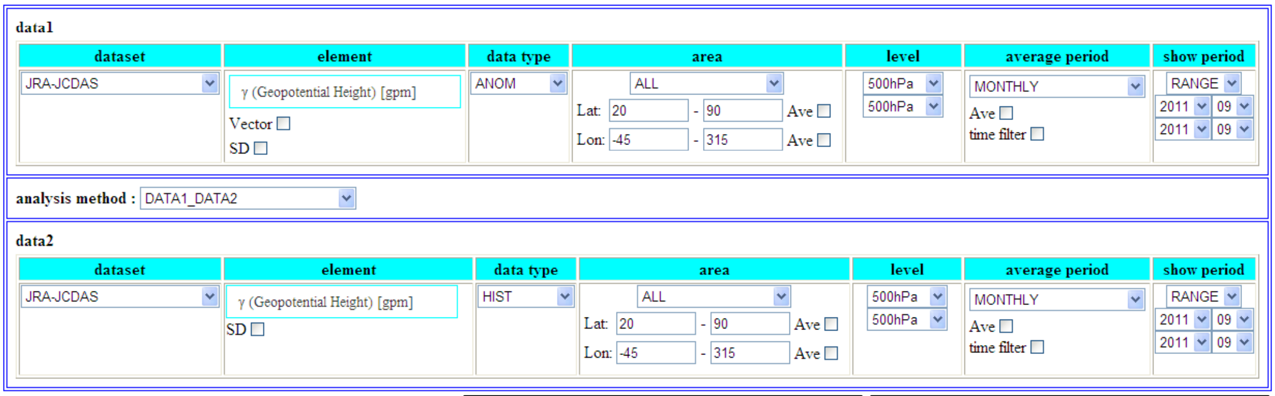
500-hPa height and anomaly

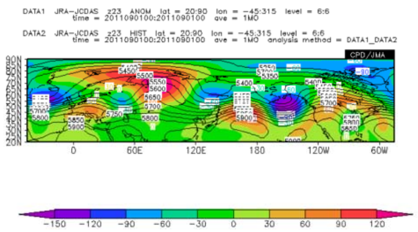


500-hPa height and anomaly

Let's know how to make a map as polar stereographic plot.

500-hPa height and anomaly



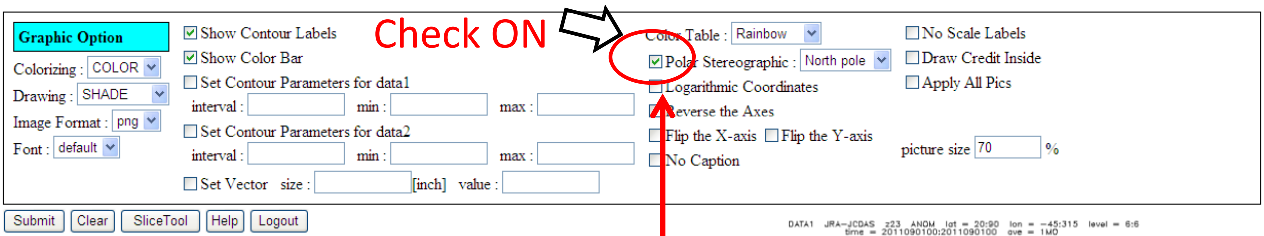


(data1)
Dataset: JRA-JCDAS
Element: γ (Geopotential Height)
Data type: ANOM
Area:
Lat: "20" - "90"
Lon: "-45" - "315"
Level: 500hPa
Average period: MONTHLY
Show period: 2011.09

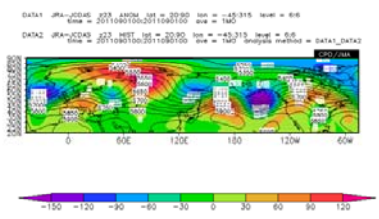
(data2)
Dataset: JRA-JCDAS
Element: γ (Geopotential Height)
Data type: HIST
Area:
Lat: "20" - "90"
Lon: "-45" - "315"
Level: 500hPa
Average period: MONTHLY
Show period: 2011.09

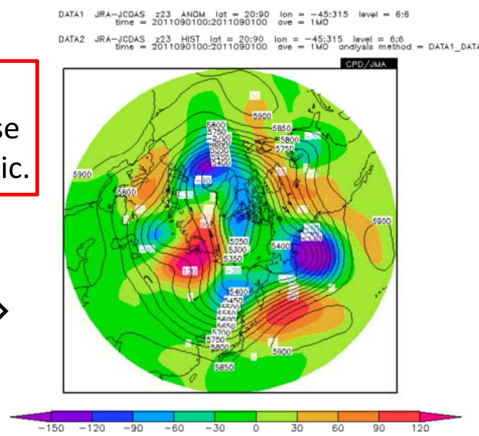
First, please make a map of 500-hPa height and anomaly in September 2011 as mentioned above.

500-hPa height and anomaly



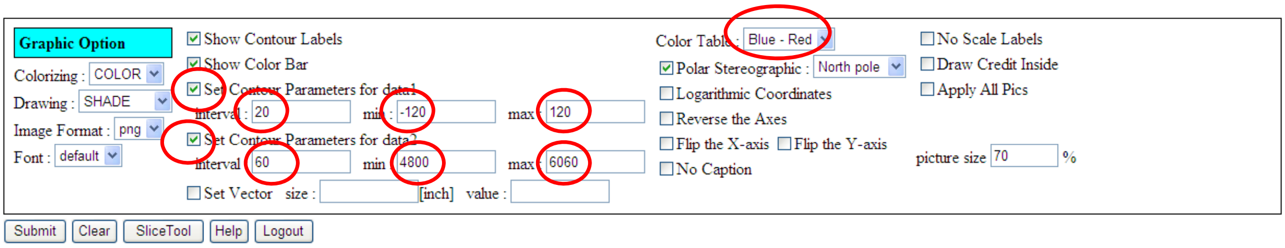
Polar Stereographic
Show polar stereographic plot, Check here and choose north polar stereographic or south polar stereographic.





Secondly, let's make a polar stereographic plot. Check "Polar stereographic" option as mentioned above and click "Submit" button.

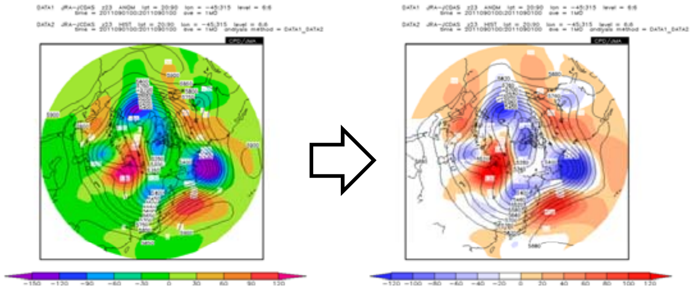
500-hPa height and anomaly



(Graphic Options)
Set contour Parameters for data1
Intervel: “20” Min: “-120” Max: “120”

Set contour Parameters for data2
Intervel: “60” Min: “4800” Max: “6060”

Color Table: “Blue-Red”



Finally, let’s change graphic options as mentioned above.

JMA's Ensemble Prediction System (EPS) for Seasonal Forecasting

JMA's Ensemble Prediction System (EPS) for Seasonal Forecasting

Climate Prediction Division, Japan Meteorological Agency (JMA/CPD)[1]

1. Introduction

(1) Forecast classification

In line with WMO's Manual on the Global Data-Processing and Forecasting System[2], forecasts are classified by their ranges as follows (Appendix I-4):

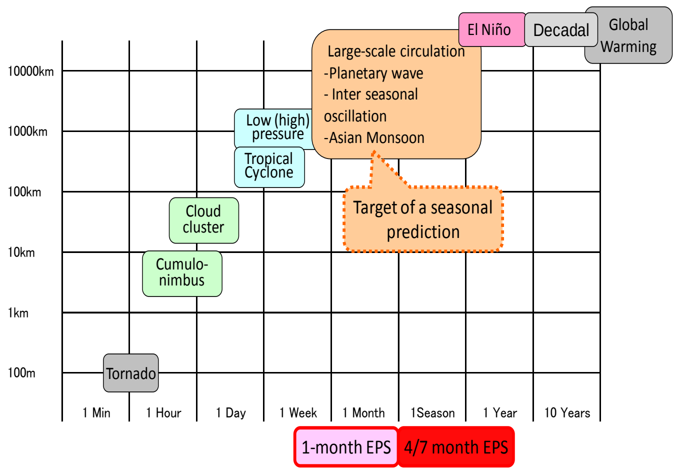
	Forecasting target period
Nowcasting	Up to 2 hours
Very short-range weather forecasting	Up to 12 hours
Short-range forecasting	Beyond 12 hours and up to 72 hours
Medium-range weather forecasting	Beyond 72 hours and up to 240 hours
Extended-range weather forecasting	Beyond 10 days and up to 30 days
Long-range forecasting	Beyond 30 days up to two years
Climate forecasting	Beyond two years

Seasonal prediction, which is the main topic of the TCC seminar, corresponds to extended-range and long-range forecasting (shaded in the above table).

(2) Targets of seasonal forecasting

An important difference between short-range and seasonal forecasting is time scale involved. The figure on the right shows the spatial and temporal scales of various phenomena. The main target of seasonal prediction is large-scale circulation of the atmosphere, which involves the following considerations:

- Blocking
  - Stationary Rossby waves (propagation of Rossby wave packets along the jet stream)
- Inter-seasonal oscillation phenomena (e.g., the MJO) and its influence in the mid-latitudes
- Large-scale convective activity related to oceanic variations in the tropics (e.g., ENSO)



Spatial and temporal scales of forecast targets

---

[1] Author: Masayuki Hirai, Oct. 2011
[2] http://www.wmo.int/pages/prog/www/DPS/Publications/WMO_485_Vol_I.pdf

(3) Predictability

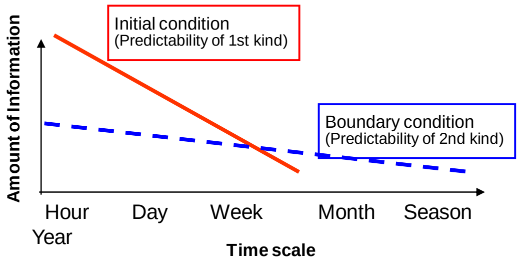
There are three basic types of uncertainty in forecasting:

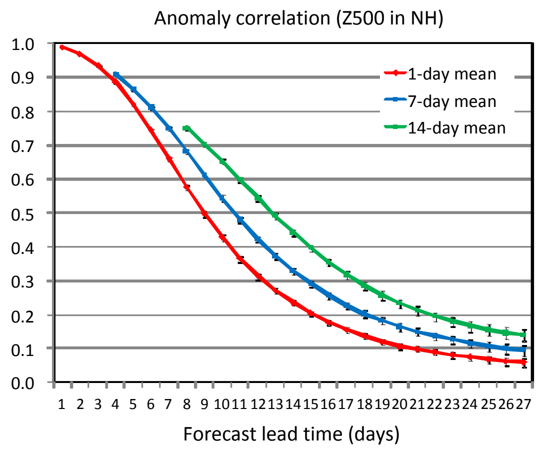
- Uncertainty relating to initial conditions: predictability of the first kind
- Uncertainty relating to boundary conditions: predictability of the second kind
- Uncertainty relating to models (not discussed here)

Due to the chaotic nature of the atmosphere, which stems from its characteristic of strong non-linearity, the limit for deterministic forecasting is about two weeks. Accordingly, the ensemble prediction system (EPS) is essential in seasonal forecasting.

However, the influence of boundary conditions is important for longer-range forecasting models. In particular, forcing on the sea surface is an important signal in seasonal forecasting. However, in extended-range forecasting, initial conditions (the first kind) and boundary conditions (the second kind) are both important.

The figure on the right shows prediction skill as an anomaly correlation of Z500 in the Northern Hemisphere over time, and indicates decreasing values (7- and 14-day means). This indicates that the influence of signals is greater for longer time periods.





Prediction skill dependency on time scale

2. Outline of JMA's EPS for seasonal forecasting

(1) Seasonal forecasts issued by JMA

Output	Date of issue	Content
Early Warning Information on Extreme Weather	Every Tuesday and Friday	Probability of extreme events based on the 7-day mean air temperature for the next 5 – 14 days
One-month forecast	Every Friday	Tercile probabilities of temperature, precipitation, sunshine duration and snowfall anomaly for week-1, week-2 (one week forecast), week-3,4 (2-week forecast) and week-1,2,3,4 (1-month mean forecast)
Three-month forecast	25th of every month	Tercile probabilities of temperature, precipitation and snowfall anomaly for 1, 2 and 3 months ahead and the 3-month mean
Cold/warm season forecast	25th of Sep., Oct., Feb., Mar. and Apr.	Tercile probabilities of temperature, precipitation and snowfall for the DJF mean (cold season) and the JJA mean (warm season)

(2) Ensemble prediction system for seasonal forecasts

JMA/CPD operates two ensemble prediction systems (EPSs) for seasonal prediction. These are the 1-month EPS and the 4/7-month EPS. Additionally, JMA/NPD (JMA's Numerical Prediction Division) also operates a 1-week EPS and a Typhoon EPS. The specifications of the two EPSs run by JMA/CPD are shown in the table below.[3]

	1-month EPS	4/7-month EPS
Model	AGCM (Atmospheric General Circulation Model)	CGCM (Coupled Ocean-atmosphere General Circulation Model)
Resolution	Horizontal: approx. 110 km (TL159) Vertical: 60 levels up to 0.1 hPa	* Atmospheric component Horizontal: approx. 180 km (TL95) Vertical: 40 levels up to 0.4 hPa * Oceanic component Horizontal: 1.0º longitude, 0.3–1.0º latitude (75˚S – 75˚N) Vertical: 50 levels
Forecast range	Up to 34 days	7-month (initial month of Sep., Oct., Feb., Mar., Apr) 4 months (other initial month)
Ensemble method	Combination of Breeding of Growing Modes (BGM) and Lagged Average Forecast (LAF)	
Ensemble size	50 (25 BGMs & 2 days with 1-day LAF)	51 (9 BGMs & 6 days with 5-day LAF)
Frequency of operation	Every Wednesday and Thursday	Every 5 days
Frequency of model product creation	Once a week Every Friday	Once a month Around the 20th (no later than the 22nd) of every month

[3] Details are provided on the TCC website at http://ds.data.jma.go.jp/tcc/tcc/products/model/outline/index.html.

(3) Hindcasts

A hindcast is a set of systematic forecast experiments for past cases. Hindcast experiments are performed using a forecasting model identical to the current operational model. Hindcast datasets are used to determine the systematic bias and skill of models and to establish/evaluate statistical models.

As hindcasting involves model calculations for a large number of past events, huge computing resources are required. Accordingly, hindcast specifications (e.g., ensemble size, calculation frequency) are more limited than those of operational system forecasts, having a smaller ensemble size and a longer initial-date interval. Differences between hindcasts and operational system output are shown in the tables below.

* 1-month EPS

	Hindcast	Operational system
Ensemble size	5 (5 BGMs, not using LAF)	50 (25 BGMs & 2 days with 1-day LAF)
Forecast range	Initial date + 33 days	2, 3, 4,...31, 32 days from the later initial date (Thursday)
Initial date	10th, 20th, end of month	Every Wednesday and Thursday
Target period of hindcast	1979 – 2009	---

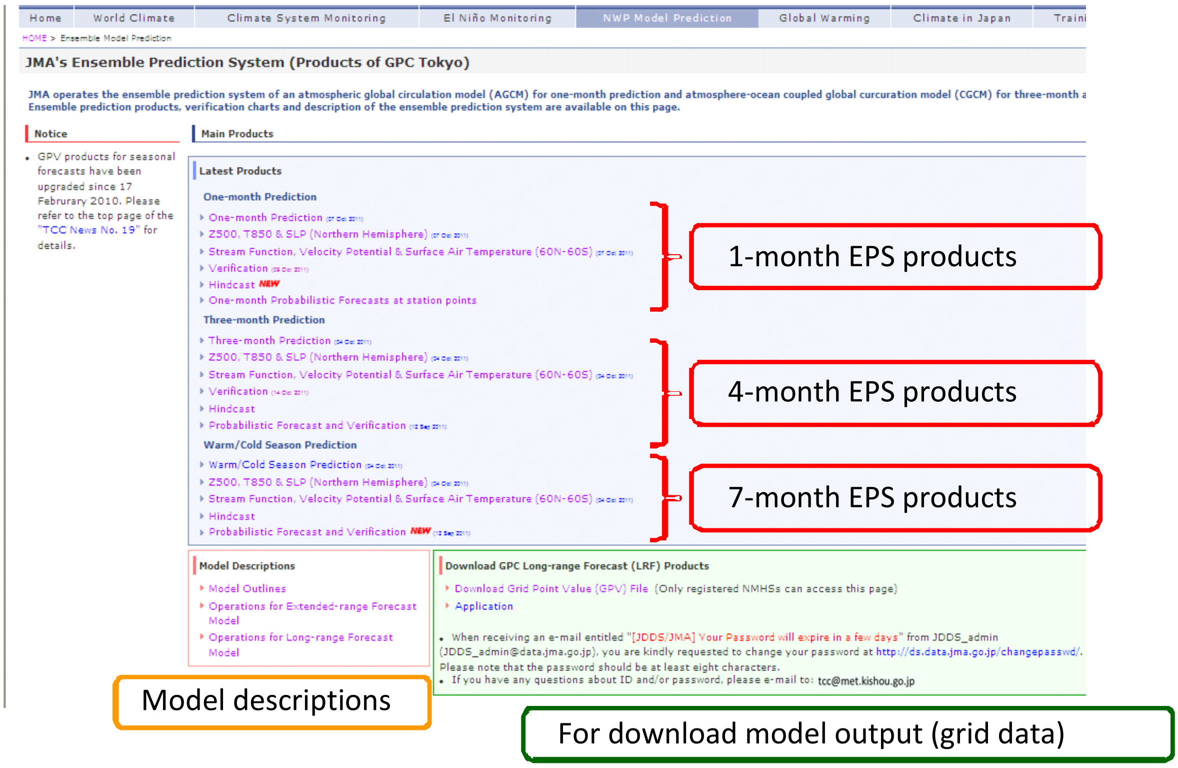
* 4/7-month EPS

	Hindcast	Operational system
Ensemble size	5 (5 BGM)	51 (9 BGMs & 6 days with 5-day LAF)
Forecast range	Lead time from 0 to 6 months as shown in the correspondence table below	(4-month EPS) Lead time from 1 to 3 as shown in the correspondence table below (7-month EPS) DJF (initial month of Oct.) JJA (initial months of Feb., Mar. and Apr.)
Initial date	24 initial dates a year (1st Jan., 16th Jan., 15th Feb., 2nd Mar., 17th Mar.,... 2nd Dec. and 17th Dec.	Once a month
Target period of hindcast	1979 – 2008	–

Correspondence between lead times (months) and initial dates

Target month Initial date	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1-Jan, 17-Dec	0	1	2	3	4	5	6					
31-Jan, 16-Jan		0	1	2	3	4	5	6				
2-Mar, 15-Feb			0	1	2	3	4	5	6			
1-Apr, 17-Mar				0	1	2	3	4	5	6		
1-May, 16-Apr					0	1	2	3	4	5	6	
31-May, 16-May						0	1	2	3	4	5	6
30-Jun, 15-Jun	6						0	1	2	3	4	5
30-Jul, 15-Jul	5	6						0	1	2	3	4
29-Aug, 14-Aug	4	5	6						0	1	2	3
28-Sep, 13-Sep	3	4	5	6						0	1	2
28-Oct, 13-Oct	2	3	4	5	6						0	1
2-Dec, 17-Nov	1	2	3	4	5	6						0

3. Seasonal forecasting model products on the TCC website



Model product page on the TCC website
(http://ds.data.jma.go.jp/gmd/tcc/tcc/products/model/index.html

3.1 1-month EPS products

* Forecasting maps — Weekly forecasting maps
  Animated forecasting maps (experimental)
* Real-time verification of routine model forecasts
* Hindcast verification charts
* Calibrated probabilistic forecasts at station points

3.2 4/7-month EPS products

* Forecasting maps — Monthly forecasting maps
* Hindcast verification charts
* Calibrated probabilistic forecasts

3.3. Details of grid data (GPVs)

3.3.1. 1-month EPS operational system grid data

* Daily data on statistics (ensemble mean and anomaly)

Forecasting target period for each record:

2, 3, 4,...31, 32 days from the later initial date (Thursday)

Elements:

Ensemble mean forecast and anomaly

Level	Pressure or height (Z)	Rain	T	U	V	RH	Stream function	Velocity potential
Surface	O (SLP)	O	O (2 m)					
850 hPa			O	O	O*	O*	O	
700 hPa			O					
500 Pa	O (Z)							
200 Pa				O	O*		O	O
100 hPa	O (Z)							

*: Ensemble mean only (anomaly not included)

* Daily data on individual ensemble members

Forecasting target period for each record:

0, 1, 2,...33 days from later initial date (Thursday)

Elements:

Individual ensemble member forecast

(Anomalies not included)

Level	Pressure or height (Z)	Rain	T	U	V	RH	Stream function	Velocity potential
Surface	O (SLP)	O	O (2 m)					
1,000 hPa	O (Z)		O	O	O	O	O	
850 hPa	O (Z)		O	O	O	O		
700 hPa	O (Z)		O	O	O	O		
500 Pa	O (Z)		O	O	O	O		
300 Pa	O (Z)		O	O	O	O	O	O
200 hPa	O (Z)		O	O	O			
100 hPa	O (Z)		O	O	O			

3.3.2 1-month EPS hindcast grid data

* Daily data on individual ensemble members

Forecasting target period for each record:

0, 1, 2,...33 days from the initial date

Elements

Individual ensemble member forecast

(Anomalies not included)

Level	Pressure or height (Z)	Rain	T	U	V	RH	Stream function	Velocity potential
Surface	O (SLP)	O	O (2 m)					
1,000 hPa	O (Z)		O	O	O	O	O	
850 hPa	O (Z)		O	O	O	O		
700 hPa	O (Z)		O	O	O	O		
500 Pa	O (Z)		O	O	O	O		
300 Pa	O (Z)		O	O	O	O	O	O
200 hPa	O (Z)		O	O	O			
100 hPa	O (Z)		O	O	O			

3.3.3. 4/7-month EPS operational system grid data

* Monthly data on ensemble statistics (ensemble mean, anomaly and spread)

Forecasting target period for each record:

- 1-month mean (records 1 to 3)
  - Lead 1, 2 and 3 months from the initial month for 4-month EPS
  - DJF or JJA for 7-month EPS
- 3-month mean (record 4)
  - Lead 1 to 3 months for 4-month EPS
  - DJF or JJA for 7-month EPS

Elements

Ensemble mean forecast, anomaly and spread

Level	Pressure or height (Z)	Rain	T	U	V	RH
Surface	O (SLP)	O	O (2 m, SST*)			
850 hPa			O	O	O	
500 Pa	O (Z)					
200 Pa				O	O	

*: Ensemble mean and anomaly only (ensemble spread not included)

* Monthly data on individual ensemble members (forecast and anomaly)

Forecasting target period for each record:

- 1-month mean (records 1 to 3)
  - Lead 1, 2 and 3 months from the initial month for 4-month EPS
  - DJF or JJA for 7-month EPS
- 3-month mean (record 4)
  - Lead 1 to 3 months for 4-month EPS
  - DJF or JJA for 7-month EPS

Elements

Individual ensemble member forecast and anomaly

Level	Pressure or height (Z)	Rain	T	U	V	RH	Specific humidity
Surface	O (SLP)	O	O (2 m, SST*)				
850 hPa	O (Z)		O	O	O	O	O
500 Pa	O (Z)		O	O	O		
300 hPa	O (Z)						
200 Pa	O (Z)		O	O	O		
100 hPa	O (Z)						

3.3.4 4/7-month EPS hindcast grid data

* Monthly data on individual ensemble members (forecast and anomaly)

Forecasting target period for each record:

Lead 0, 1, 2,...6 months from the initial month (7 records)

Elements

Individual ensemble member forecast

Level	Pressure or height (Z)	Rain	T	U	V	RH	Specific humidity
Surface	O (SLP)	O	O (2 m, SST*)				
850 hPa	O (Z)		O	O	O	O	O
500 Pa	O (Z)		O	O	O		
300 hPa	O (Z)						
200 Pa	O (Z)		O	O	O		
100 hPa	O (Z)						

Using Grid Point Value Data Provided on the TCC Website

Using Grid Point Value Data Provided on the TCC Website - Visualization using GrADS -

Climate Prediction Division, Japan Meteorological Agency[1]

1. Purpose and agenda

TCC (the Tokyo Climate Center) has provided global grid datasets (grid point values, or GPVs) of one-month forecasts since June 2002, three-month forecasts since September 2003, and seven-month forecasts since February 2004.

Visualization of data provided by TCC is the main subject of this document, which also explains how to obtain data from the TCC website and the use of tools to display them.

A free program called GrADS (Grid Analysis and Display System) developed by COLA/IGES in the USA enables visualization of GPV data from the TCC website.

The GPV data provided by TCC are coded in GRIB2 format (GRIB Edition 2) for all one-month, three-month and seven-month forecasts. Conversion from GRIB2 format to GrADS format is a significant issue that needs to be addressed.

The GRIB2 format is explained on the WMO website at http://www.wmo.int/pages/prog/www/WMOCodes.html.

* Forecast data from JMA's latest models start from March 19, 2008, for one-month forecasts, and from February 2010 for three-month and seven-month forecasts. One-month forecast data are currently available only from April 2009 due to the capacity limitations of the TCC website.

2. Computing environment requirements

A computer that can run GrADS (Grid Analysis and Display System) is required. MS Windows, Linux and most UNIX computers are appropriate.

2.1. Installation of GrADS (Linux and UNIX)

- Go to **http://grads.iges.org/grads/index.html**.
- Go to the **Downloading the Software** section and click on the "**Downloading the Software"** link.
- Chose the appropriate version of the runtime program and download it.
- Extract the downloaded file.
  - E.g., **tar zxvf grads-1.9b4-linuxRH9.tar.gz**
- Run the program.
- Implement path setting.

  export GRADSHOME=/user/local/grads

  (The path depends on the installation location. Check where GrADS is installed and replace the text here with the relevant path.)

[1] Author: Ryoji Kumabe, Oct. 2010
Last revised by: Masayuki Hirai, Oct. 2011

export PATH=$PATH:${GRADSHOME}/bin
export GADDIR=${GRADSHOME}/data
Add above lines to the file "bash_profile" or "bashrc" for convenience.

2.2. Installation of wgrib2 (Linux and UNIX)

All grid point data on the TCC website are provided in GRIB2 format. Therefore, conversion of file format from GRIB2 to GRADS is required. "wgrib2" is a computer program used to convert GRIB2 formatted data into binary data that GrADS can handle.

* COLA/IGES has made GrADS ver. 2 compatible with GRIB2, although this compatibility has not been confirmed by the author. Users may establish their own way to use GrADS without the wgrib2 program.
* All binary files are compatible among computers with the same byte order, while conversion is needed for those with different byte orders (Windows and most Linux systems vs. most UNIX systems)

- Go to **ftp.cpc.ncep.noaa.gov/wd51we/wgrib2**.
- Download **wgrib2.tgz.1.7.8b**.
- Run the following commands in the directory to which the file has been downloaded:
  - **tar zxvf wgrib2.tgz.v1.7.8b**
  - **cd grib2**
  - **gmake -f makefile_all_libs_new**
  - **cd wgrib2**
- Copy the wgrib2 program to the directory in which GrADS is installed.

2.1. Installing GrADS (Windows)

While COLA/IGES in the US provides a Windows version of GrADS, US-based company SourceForge, Inc. has developed and provides an extension version of GrADS based on the original GrADS program from COLA/IGES called "OpenGrADS", which offers useful utilities such as wgrib2. The procedure for installing "OpenGrADS" from SourceForge, Inc. is described here. (Note that the procedure may change with their web design policy.)

- Go to the Website of the OpenGrADS at **http://sourceforge.net/projects/opengrads/**.
- Click on the banner "**Download"** and save the file to the desired directory.
- Run the file.
  - Check "Add application directory to your system path" during installation. This automatically takes care of the path setting procedure required for the Linux version.

For Windows and OpenGrADS, as wgrib2 is also packaged with OpenGrADS, there is no need to install wgrib2 separately as there is for Linux.

For more information, see the GrADS website at http://grads.iges.org/grads/index.html.

3. Downloading GPV data from the TCC website

The procedures for downloading GPV data from the TCC website are outlined here. GPV data can be downloaded for one-month, three-month and seven-month forecasts.

Open your web browser (e.g., Internet Explorer) by double-clicking on its icon.
TCC website URL: http://ds.data.jma.go.jp/tcc/tcc/index.html
During the procedures outlined below, the user ID and password provided by TCC will need to be input.

Downloading one-month predictions

Visit the **TCC top page** → Click on **NWP Model Prediction** in the top menu → Click on **Download Grid Point Value (GPV) File** → Input the user name and password → Click on **Daily Statistics** (for ensemble mean forecasts) or on **All Members** (for individual ensemble member forecasts), in the **1-month** category of **NWP Model Prediction** → Choose **Download Grid Point Value (GPV) Data**.
Select a file from the listed GPV data.

* The characters 'yyyymmdd' indicate the year (four digits), the month (two digits) and the day (two digits) for the initial time of the prediction.

Data are uploaded every Friday evening (Japanese time) for predictions in the next one-month period.

Downloading three-month predictions

Visit the **TCC top page** → Click on **NWP Model Prediction** in the top menu → Click on **Download Grid Point Value (GPV) File** → Input the user name and password → Click on **Statistics** for mean forecasts in each month over the next three months, or on **All Members** for forecasts in each month of all ensemble members in the **3-month** category of **NWP Model Prediction** → Choose **Download Grid Point Value (GPV) Data (201002 – present)**.
Select a file from the listed GPV data.

Data are uploaded by around the 20th of each month for three-month predictions starting at the beginning of the next month.

Downloading seven-month predictions

Visit the **TCC top page** → Click on **NWP Model Prediction** in the top menu → Click on **Download Grid Point Value (GPV) File** → Input the user name and password → Click on **Statistics** for mean forecasts in each month during the last three months of the prediction period, or on **All Members** for mean forecast in each month of all ensemble members in the **7-month** category of **NWP Model Prediction** → Choose **Download Grid Point Value (GPV) Data (201002 – present).**
Select a file from the listed GPV data.

Data are uploaded by around the 20th of February, March and April (for the warm season) and September and October (for the cold season) for seven-month predictions starting at the beginning of the next month.

The total file sizes are approximately 1 MB for one-month statistics forecasts (26 files of about

83 – 144 KB each), 1.5 MB for one-month all-member forecasts (44 files of about 34 KB each), 2 MB for three-month mean forecasts (29 files of about 47 – 62 KB each), 124 MB for three-month all-member forecasts (40 files of 3.1 MB each), 2.4 MB for seven-month statistics forecasts (29 files of about 83 KB each), and 170 – 250 MB for seven-month all-member forecasts (40 files of about 4.3 – 6.4 MB depending on the initial month).

* The content and other information of each GPV data set are shown in Appendix 3: One-month forecast GPV products; Appendix 4: Three-month outlook GPV products; and Appendix 5: Seven-month forecasts.

4. Using one-month forecast GPV data

This section outlines the procedure for visualizing the one-month prediction GPV data provided on the TCC website using the GrADS program. The *italics* in examples of computer commands in this section and the following sections should be replaced with the appropriate file names and other information, e.g., assume p500_Phh_em.20101021 for *gribfile*.

4.1 Conversion of TCC-GRIB files into GrADS files

Program
wgrib2 converts GRIB2 data provided by TCC into GrADS format.

Input: GRIB2 data
Any GRIB2-format file on the TCC website

Output: GrADS files
A GrADS-formatted binary file (GrADS data) of all data elements in the input file is produced. Alternatively, the user can choose specific elements as shown in 1.2 below.

1) Running the program

Programs must be executed on a console (Linux, UNIX) or by command prompt (Windows). In Windows, the command prompt window can be opened by clicking on the command icon for the accessories section in the pull-down menu activated from the Start button. On the console or the command prompt, first navigate to the directory to which the data files have been downloaded.

1.1) To retrieve all data from **Statistics (GRIB2) files** and **All Members (GRIB2) files**
> **wgrib2 *grib2_file* -no_header -bin *output_file***

1.2) To retrieve a particular member from **All Members (GRIB2) files** (when retrieving the ensemble member +10)
>**wgrib2 *grib2_file* | grep ENS=+10 | wgrib2 -i *grib2_file* -no_header -bin *output_file***

* To find the ensemble number, wgrib2's grib2_file can be used.
ENS spans from +1 to +24 and from -1 to -24 for one-month prediction;
from +1 to +5 and from -1 to -5 for three-month prediction; and
from +1 to +4 and from -1 to -4 for seven-month prediction.

* Replacing "-bin" with "–text" produces a text file instead of a GrADS binary file. This can then be processed in a spreadsheet program such as Microsoft Excel.

2) Editing the control file

Two files are required when using the GrADS program: one is a binary data file created by the wgrib2 program, and the other is a control file. A control file is a text file containing meta-information of GrADS data such as data dimensions, time, elements, etc.

Text files are written in a text editor such as Notepad. The text should be as follows:

Ex. 1) A statistics (GRIB2) file for one-month prediction

filename = *psea.ctl* (any filename.ctl is OK)

```
-----------------------------------------------------------------------------------------
dset ^surf_Ppp_em.yyyymmdd.bin        : GrADS file name
undef -1.e+20                         : Infinite number
title example                         : Title
xdef 144 linear   0 2.5               : Number of grids, starting lon, distance in x-direction
ydef  73 linear -90 2.5               : As above, but in y-direction
zdef   1 linear   0 1                 : As above, but in z-direction (vertical)
tdef   4 linear 03oct2009 7dy         : Time steps, start time, time unit
                 (meaning of time units: dy means days, mo means months)
vars 1                                : Number of variables (elements)
PSEA   0 0   PSEA                     : Name of variable
endvars                               : Closing remark
-----------------------------------------------------------------------------------------
```

Note: the symbol '^' in the dset section indicates that the path to the file is relative to the directory where the control file is located. Paths in GrADS are expressed by "/" as also seen in UNIX and Linux, in contrast to the "¥" used in Windows.

Ex. 2) An all-member (GRIB2) file for one-month prediction

filename = psea.ctl (any filename.ctl is OK)

```
-----------------------------------------------------------------------------------------
dset ^surf_Ppp_mb.yyyymmdd.bin
undef -1.e+20
title example
xdef 144 linear   0 2.5
ydef  73 linear -90 2.5
zdef   1 linear   0 1
tdef  34 linear 03oct2009 1dy      (The spaces here are important.)
vars 1
PSEA   0 0   PSEA
endvars
-----------------------------------------------------------------------------------------
```

4.2. Running GrADS and opening the control file

> **grads**
------- GrADS command -------
> **open** ***ctl_file.ctl*** ex) **open** **psea.ctl**
> **d** ***psea***
> **quit**

(See Appendix 1 for GrADS basic commands.)

5. Examples of long-range prediction data visualization using GrADS

Exercise and tutorial
Visualization with GrADS → TCC website
(http://ds.data.jma.go.jp/tcc/tcc/products/model/tips/vis.html)

This section outlines the visualization of GPV data for one-month forecasts using GrADS. GPV data in GrADS format and the control file are first prepared as outlined in Section 3.1.

5.1 Example 1

This example outlines how to:

a) represent 500-hPa geopotential height using contours;
b) represent 500-hPa geopotential height anomaly using shading; and
c) save images for each prediction time as png files.

Data preparation:
Download p500_Phh_em.yyyymmdd and p500_Paa_em.yyyymmdd.
Convert both GRIB2 files into GrADS format using wgrib2.
Edit the control files and save them as p50h.ctl and p50ha.ctl.

Script preparation:
Edit the script file below and save it with a name such as sample1.gs in the same directory as the data.
Scripts called solutions (cbarn.gs and color_k.gs) can be downloaded from http://ds.data.jma.go.jp/tcc/tcc/products/model/tips/vis.html.
Scripts should all be stored in the same directory.

GrADS operation:
Execute the GrADS program in the same directory as that used in the above step.
In GrADS, type the command “run sample1.gs” and watch the processing.

5.2 Example 2

This example outlines how to:

a) plot divergence/convergence of low-level and high-level wind fields; and
b) represent wind with stream lines.

Data preparation:

The procedure is the same as that for Example 1, but with p850_Pwu_em.yyyymmdd and p850_Pwu_em.yyyymmdd for the data, and u85.ctl and v85.ctl for the control file names.

Script preparation:

As outlined for Example 1, but with sample2.gs as the script file name.

GrADS operation:

As outlined for Example 1, but with run sample2.gs as the GrADS command.

6. Pointers for data handing in GrADS

GrADS data have four dimensions: X, Y, Z and T. However, as display is two-dimensional, the degree of freedom needs to be restricted to a maximumof two dimensions (except for time sequences expressed as animations). To achieve this, a 'set' command such as 'set t 1' is used.

Example (typical GrADS data)

```
--------------------------------------------------------------------------------------
dset ^mg1ws
undef -19999.0
title 1mE_GPV.20070920
ydef    37 linear -90.0000 2.5
xdef 144 linear 0.000000 2.5
tdef      4 linear 20Sep2007 7dy
zdef      7 levels 1000 900 850 700500 300 150
vars 20
Z500    0 99 ** Geopotential height at 500 hPa [gpm]
T850    0 99 ** Temp. at 850 hPa [K]
………..
      (continues)
………..
PSEAA 0 99 ** Sea level pressure anomaly [Pa]
Z100A 0 99 ** Geopotential height anomaly at 100 hPa [gpm]
ENDVARS
--------------------------------------------------------------------------------------
```

Display types: contours, shading, barbs, etc.
Projection: latlon, polar stereo, etc.

GrADS script library: Useful scripts can be downloaded from the Internet at http://grads.iges.org/grads/gadoc/library.html.

GrADS data sets

Data file: a binary file with 4-byte floating point values

Ctl file: a file describing the corresponding data file

GrADS script file: a file containing a collection of commands

Command to run the GrADS program with a GrADS script file silently:

echo “quit” | grads –lbc “run *gsfile*”

7. Advanced techniques for obtaining and using TCC data

A combination of tools available on the Internet will facilitate the use of data provided by TCC in terms of automatic downloading and visualization.

1. Shell scripts: Such scripts can be used to execute a collection of commands at one time. They can also be used to control commands such as repeating and conditional reactions.

Linux: installed by default

Windows: use the cygwin system available at http://www.cygwin.com/.

2. wget commands: Such commands can be used for automatic ftp and http access when operated within a shell script.

Linux: installed by default

Windows: choose wget from the web components upon cygwin system installation. Alternatively, download the Windows version of wget from http://users.ugent.be/~bpuype/wget/.

Use .wgetrc (or wgetrc) files to set web IDs, passwords and proxies as necessary.

3. Running GrADS with GrADS scripts

Typing the command grads -bcl "run sample1.gs" on a Linux or UNIX machine runs GrADS without opening the GrADS interactive screen and allows access to png image files. This is convenient for automatic processing of download data and for visualization. However, this approach does not work on the Windows version of GrADS.

4. The crontab command can be used in Linux to set scheduled commands for a particular time/day either on a single occasion or repeatedly.

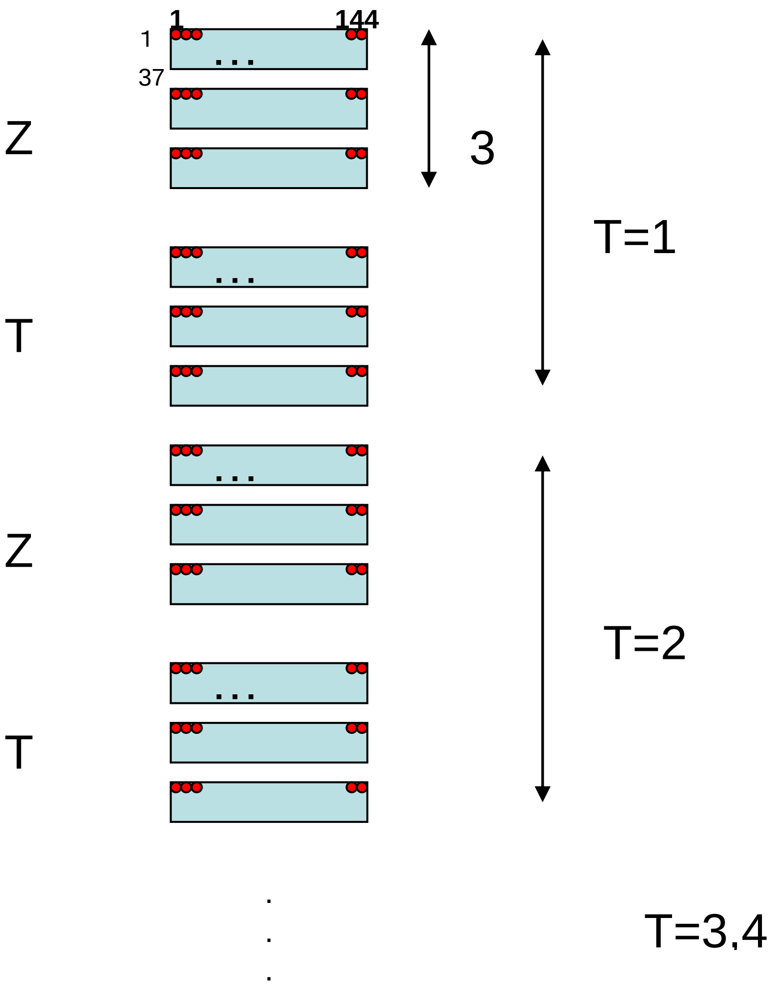
8. GrADS data format

GrADS data consist of a simple series of floating computer point values. The data alignment is shown in the corresponding control file. Below is an example of a control file (slightly modified for simplicity):

```
---------------------------------------------------------------------------------------
dset ^mg1ws
undef -19999.0
title 1mE_GPV.20070920
```

```
xdef 144 linear 0.000000 2.5
ydef    37 linear -90.0000 2.5
zdef      3 levels 1000 850 700
tdef      4 linear 20Sep2007 7dy
vars 2
Z    3 99 ** Geopotential height at 3 levels
T    3 99 ** Temp. at 3 levels
ENDVARS
--------------------------------------------------------------------------------------
```

The GrADS data file would be as follows:



Below is an example of programming to create GrADS data.

C programming

Creates a data set covering the whole globe (lat: -90 – 90; long: 0 – 360) with five time steps.

Below is an example of the corresponding control file.

DSET   ^test.dat

```
UNDEF 999.9
XDEF 361 LINEAR    0.0 1.0
YDEF 181 LINEAR    -90.0 1.0
ZDEF 1 LEVELS 1000
TDEF 1 LINEAR 1JAN2000 1DY

VARS 1
element        1     99     Plot time series
ENDVARS
```

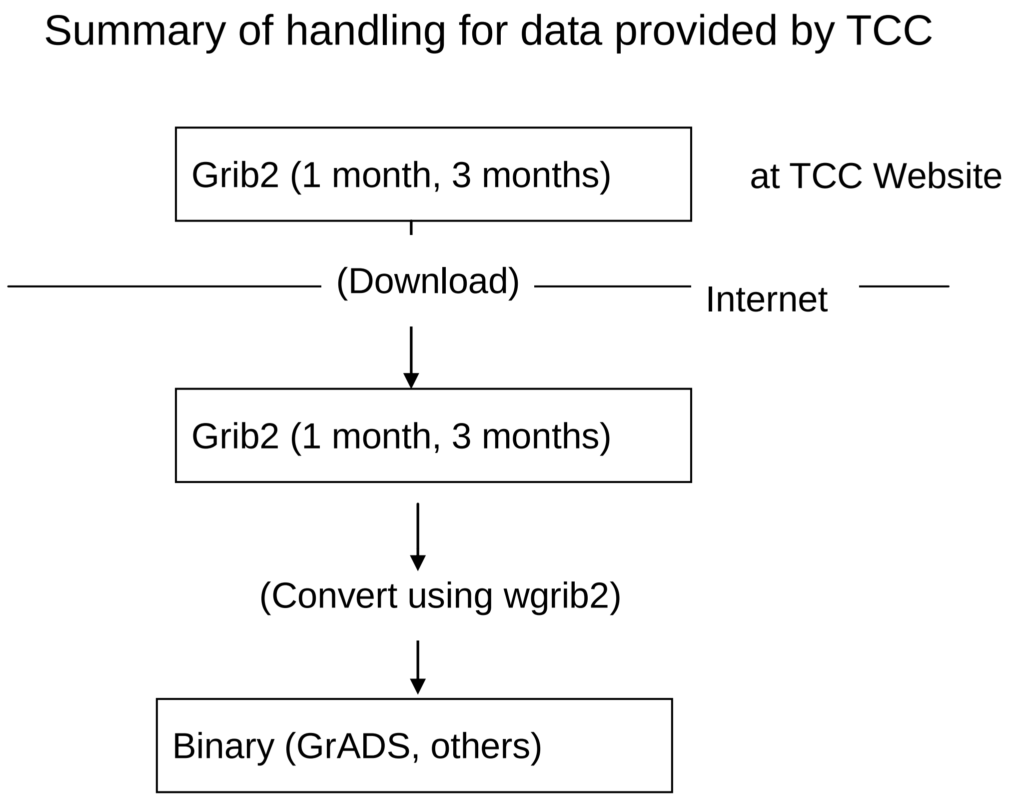
Ex. 8.1

Modification of the sample program and control file so that they contain time steps of five days

Ex. 8.2

Creation of a control file indicating that the data file created in Exercise 7.1 contains five levels of data instead of five days

Appendix 1 Summary of handling for data provided by TCC



Appendix 2 GrADS basic commands

a) General

> open ctl_file
> q file (Query info of the open file) or q ctlinfo
> quit (Leave GrADS)
> reset (Reset to the initial state without closing the file)

b) Displaying variables

> d variable_name
> d expression ex.) tempe-273.15, mag(ugrd,vgrd)
> clear

c) Changing area, level, time or map projection

> set lon lon1 lon2 ex.) set lon 0 180
> set lat lat1 lat2 ex.) set lat 0 90
> set z level1 level2 (If a 3D representation is specified as a single level, the chart will be a cross section.)
> set t n m (If nm is set, the chart shows time changes as for an animation.)
> set mproj nps (For northern polar stereo projection)

d) Changing graphic type

> set gxout [contours, shading, vectors, streams, grids, etc.] (Choose one of the parameters in parentheses.)

e) Changing graphic options

> set ccolor color_index ex.) set ccolor 1 (white) : set color
> set cint interval_value : set interval
> set cmin minimum_value : set minimum value
> set clab on/off : control label
> set cthick [1-10] : set line width

f) Changing map data set resolution

> set mpdset lowres | mres | hires

g) Output of GrADS metafiles (*.gmf)

> enable print filename.gmf
> d variable_name
> print
> disable print

h) Viewing GrADS metafiles (*.gmf)

> gv32 filename.gmf

With Gv32 (GrADS Metafile Viewer), it is possible to print images and save files as Windows metafiles, allowing charts to be included in Windows applications (e.g., Microsoft PowerPoint).

i) Outputting png/gif files (*.png/*.gif)
> printim *fname.(png / gif)*

j) Using scripts for GrADS
GrADS scripts allow construction of if/while/for clauses, etc. and keyboard input. Examples:
> set gxout shaded
> d expression
> run cbarn.gs

k) Changing graphic options (alternative way)
> run color_k.gs
> set ccols 47 46 45 44 43 23 24 25 26 27
> set clevs -120 -90 -60 -30 0 30 60 90 120

l) Creating titles
> draw title *String*

Outline of Guidance and Related Evaluation

Outline of Guidance and Related Evaluation

Climate Prediction Division, Japan Meteorological Agency (JMA)

1. Introduction

Guidance is a statistical downscaling technique based on grid point value (GPV) data predicted using a numerical model (Fig. 1). Although it is possible to directly use GPV data for forecasts such as those of surface temperature and precipitation, even good numerical models involve a certain degree of error due to the approximation of meteorological dynamics, the limitations of computational resources and other factors. Against this background, the accuracy of forecasts can be increased using guidance. In general, guidance for one-month forecasting involves elements such as geopotential height and temperature over target areas. The purpose of this training is to teach those on the course how to make regression model and to implement guidance for their own countries.

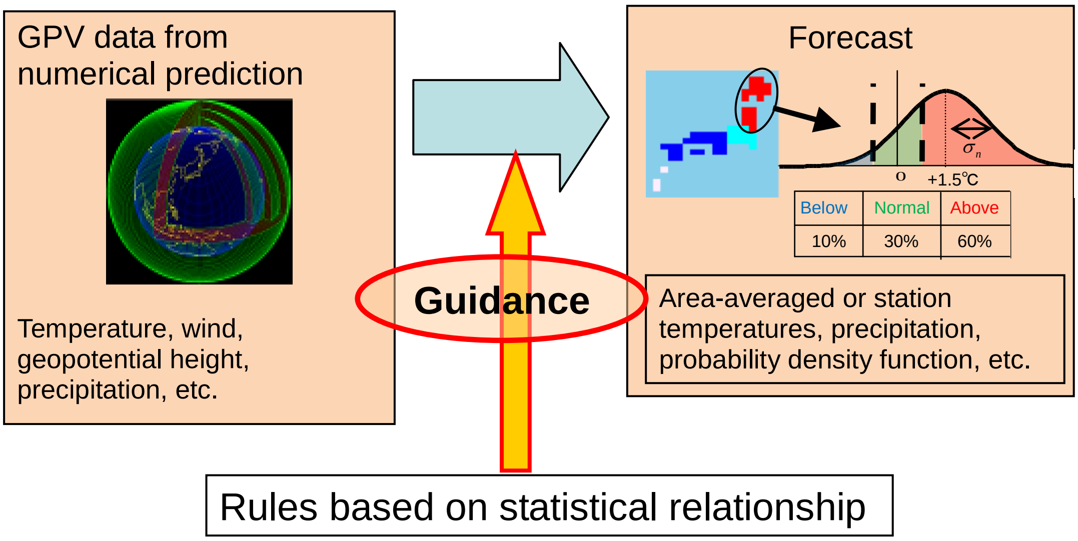


Figure 1 Schematic diagram outlining the concept of guidance

2. Regression method

Two types of time series are used in implementing guidance. One involves meteorological variables for issued forecasts, and the other involves variables obviously related to the former. The former and latter elements are referred to as predictands and predictors, respectively. Our purpose is to predict the future value of predictands using the statistical relationship between predictands and predictors.

To clarify the regression method, here we consider the simplest case, namely “a single regression model.” This is a predictive approach using a single predictor.

The single regression model is written as

$$\mathbf{Y = ax + b + \varepsilon}$$

Here, **Y** is the objective variable (i.e., the predictand), **x** is the predictor, **a** is a regression coefficient, **b** is a constant, and ε is an error term.

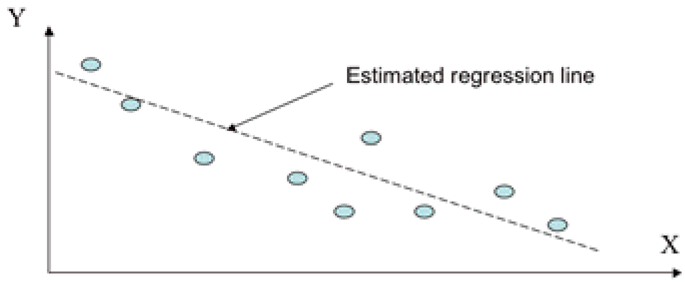


Figure 2 Single regression model

Y is the predictand.

X is another variable that can be used to predict demand.

2-1. Normalization of precipitation

Temperature histograms are generally based on normal (Gaussian) distribution, while precipitation histograms usually have gamma distribution and differ from those with Gaussian distribution. As the error distribution of regression models is assumed to involve normal distribution, precipitation data need to be normalized (Fig. 3). The simplest way to achieve this is to use the power technique. JMA's seasonal forecast guidance uses a power of 1/4 for precipitation and snowfall.

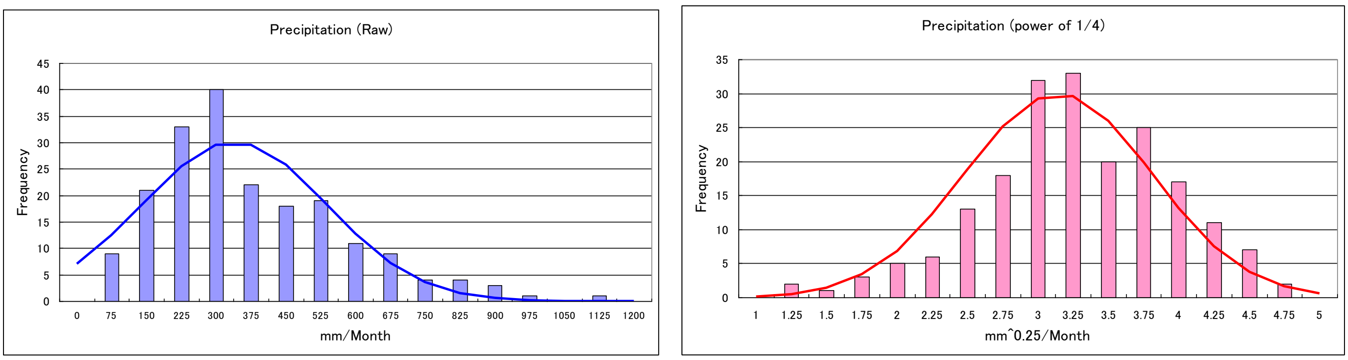


Figure 3 Precipitation histograms

Raw precipitation data differ from those with Gaussian distribution (left).

After applying a power of 1/4, the histogram fits Gaussian distribution (right).

Both bold lines indicate Gaussian distribution.

3. Probability forecasts

Even with guidance, long-range forecasting involves some level of uncertainty due to the chaotic nature of atmospheric flow. It is therefore necessary to take this uncertainty into account, and the optimal method for this is probabilistic forecasting.

The probability density function (PDF) is assumed to have normal (Gaussian) distribution with mean x_s and standard deviation σ_n (Fig. 4).

The mean x_s is predicted using the regression model, and the standard deviation σ_n is assumed to be the root mean square error of the regression model.

In this training, we will make a two-category probability forecast based on guidance (Fig. 5).

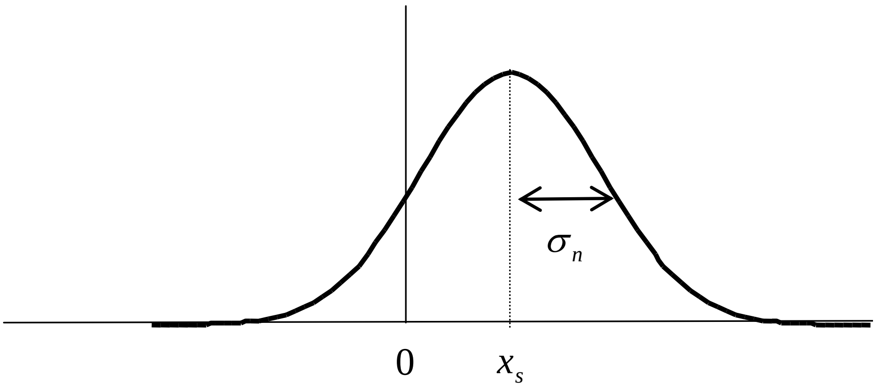


Figure 4 Schematic diagram showing the forecast probability density function (PDF)

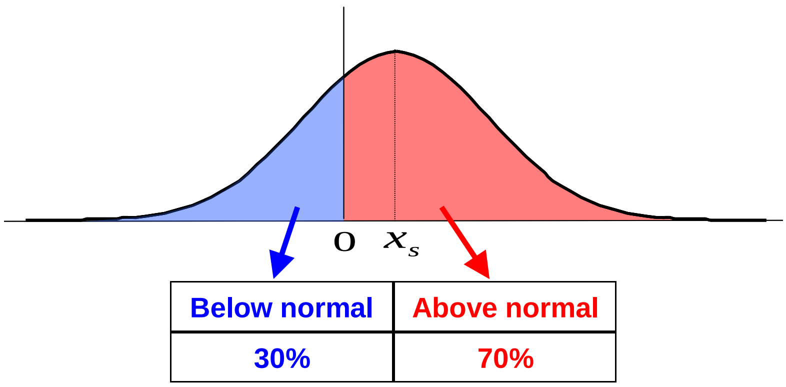


Figure 5 Schematic diagram showing a two-category forecast

4. Evaluation of probability forecasts

In this chapter, we will look at two techniques for evaluating probability forecasts. One is a reliability diagram that shows the occurrence frequencies of events based on forecast probability and the Brier score used to evaluate forecasts, and the other is the relative operating characteristic (ROC) curve, which is drawn based on the hit rates and false alarm rates often used to evaluate forecasting techniques.

a) Reliability diagrams

Reliability diagrams relate to the Brier score, and help us to understand the characteristics of probability forecasting. This section explains what the lines in such diagrams indicate.

A reliability curve – the most important element in a reliability diagram – plots the occurrence frequency of an event against the forecast probability. This shows how often the event actually occurs in relation to the forecast probability. If a forecast has absolute reliability, the probability coincides with the observed frequency, and the curve will be diagonal. Accordingly, the reliability of a probability forecast is indicated by the proximity of the curve to the diagonal. Any deviation from the diagonal indicates the characteristics of bias, with points below the diagonal indicating over-forecasting (excessive forecast probability) and those above it indicating under-forecasting (deficient forecast probability).

The horizontal line in a reliability diagram represents climatic frequency; generally, this is near 50% for two-category forecasts and near 33% for three-category forecasts. A forecast based on climatic frequency makes no discrimination between signals and no-signals for events, and has no resolution. Hence, proximity to the horizontal line indicates lower forecast resolution.

The frequencies of individual forecast probabilities may be plotted on a reliability diagram either in histogram form or in line graph form. Such graphs indicate the tendency of forecast probability. If the level of reliability is uniform, higher occurrence frequencies of extreme probabilities such as 0% or 100% indicates that the forecast has higher resolution and is closer to a deterministic forecast.

In light of these considerations, a number of conclusions can be drawn from the reliability diagrams shown in Fig. 6. The reliability curve on the left, which roughly follows the diagonal, shows that the forecast has a high level of reliability. Additionally, as the frequency of probability forecasting is higher at the extreme probabilities, it can be considered to have high resolution. Conversely, the reliability curve on the right shows lower reliability than that on the left, as well as a clear forecasting tendency for large values of forecast probability. The peak of forecast frequency is near the climatic frequency, and this indicates that the resolution is lower than that shown on the left.

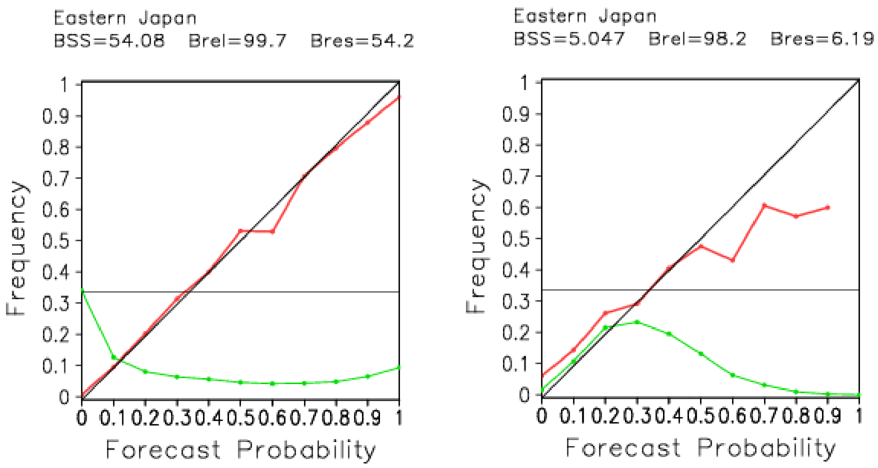


Figure 6 Reliability diagram of JMA guidance for temperature in eastern Japan
Left: 1st week; right: 3rd – 4th week
The red line shows occurrence frequencies by forecast probability.
The green line shows frequencies of forecast probabilities.
The values of BSS, Brel and Bres at the top of each panel denote the Brier skill score, the reliability skill score and the resolution skill score (%), respectively.

b) Brier score

The Brier score (hereinafter abbreviated to BS) is used to enable comprehensive evaluation of probability forecasting. The BS for two categories (written as b) is defined by formula (1):

$$b = \frac{1}{N}\sum_{i=1}^{N}(p_i - v_i)^2, \quad 0 \le p_i \le 1, \quad v_i \in \{0,1\}, \tag{1}$$

where N represents the total number of forecasts, p_i is the forecast probability value, and v_i is a variable assumed to be 1 when the predicted phenomenon occurs and 0 when it does not occur. Smaller values of b represent better results, and the BS is 0 for perfect prediction when forecasts are deterministic, with probability values assumed as either 0 or 100% (where all forecasts are accurate).

Murphy (1973) showed that b can be mathematically decomposed into three terms. If we rewrite (1) by separating the terms for the occurrence of a phenomenon from those for non-occurrence, we obtain

$$b = \frac{1}{N}\left\{\sum_t (p_t - 1)^2 M_t + \sum_t (p_t - 0)^2 (N_t - M_t)\right\}, \tag{2}$$

where t is a subscript assigned to each probability value (for example, $p_1 = 0.1$), M_t is the number of predicted phenomena that actually occur for the tth probability value, and N_t is the number of forecasts made for the same value. We can transform (2) further to obtain (3):

$$b = \sum_t \left(p_t - \frac{M_t}{N_t}\right)^2 \frac{N_t}{N} - \sum_t \left(\frac{M}{N} - \frac{M_t}{N_t}\right)^2 \frac{N_t}{N} + \frac{M}{N}\left(1 - \frac{M}{N}\right), \tag{3}$$

where $M = \sum_t M_t$, which represents the total number of predicted phenomena that actually occur.

The first term on the right of (3) represents reliability. If we allow it to be represented by $brel$, then

$$brel = \sum_t (p_t - \frac{M_t}{N_t})^2 \frac{N_t}{N}, \tag{4}$$

where $\frac{M_t}{N_t}$ is the proportion of predicted phenomena that occur to the total number of forecasts for the tth probability value (hereinafter referred to as the occurrence frequency). Smaller values of $brel$ represent larger numbers of forecasts made with forecast probabilities that match actual occurrence frequencies. If all points of the reliability curve are located on the 45-degree line where forecast probabilities and occurrence frequencies match, then $brel = 0$.

The second term in (3) represents resolution. If we allow it to be represented by $bres$, then

$$bres = \sum_t (\frac{M}{N} - \frac{M_t}{N_t})^2 \frac{N_t}{N}, \tag{5}$$

where $\frac{M}{N}$ is the climatic frequency of the event; this is generally $\frac{M}{N} = 0.5$ for a two-category forecast and $\frac{M}{N} = 0.33$ for a three-category forecast.

The negative value for $bres$ in (3) indicates that larger values of $bres$ represent better results. A larger $bres$ indicates that the difference between the occurrence frequency of each forecast probability and the climatic frequency of the event is greater. Accordingly, if reliability is uniform, higher occurrence frequencies of extreme probabilities such as 0% or 100% correlate to higher values of $bres$.

The third term in (3) represents uncertainty. If we allow it to be represented by $bunc$, then

$$bunc = \frac{M}{N}(1 - \frac{M}{N}). \tag{6}$$

This term is not related to the performance of forecasting. It depends on climatic frequency alone and shows the level of uncertainty for the phenomenon.

These scores allow us to make judgments on the relative performance of forecasts, but not on their significance. The improvement rate with respect to a climatic value forecast is often used as a measure of forecast significance. If we write the BS of a climatic value as bc, the improvement rate can be defined by

$$BSS = \frac{bc - b}{bc}, \tag{7}$$

where BSS is the Brier skill score. If $BSS \le 0$, the forecast is inferior to the climatic value forecast, and the maximum value of $BSS=1$ represents a perfect forecast.

Likewise, the reliability skill score and the resolution skill score can be defined by the following formulas:

$$Brel = \frac{bc - brel}{bc} \tag{8}$$

$$Bres = \frac{bres}{bunc} \tag{9}$$

$Brel = Bres = 1$ for a perfect forecast.

These skill scores are illustrated using a reliability diagram in many cases. In Fig. 6, although the *BSS* of the panel on the right is inferior to that on the left, the score shows a positive value and indicates that the forecast is made with a higher level of skill than the forecast of climatic values. It is found that the *BSS* difference between the right and left sides mainly originates in *Bres*.

c) Relative operating characteristics diagrams

Probability forecasts can be transformed into yes/no forecasts of category defined by any probability threshold. For all thresholds, the corresponding hit rates and false alarm rates can be computed, and relative operating characteristics (ROC) diagrams are used to display and interpret this information. Different users of forecasts will generally have different sensitivities based on economic values, and will choose different probability thresholds as triggers for the implementation of measures in response to certain events. In this way, ROC diagrams help users to decide probability thresholds.

Table 1 shows the relationship between forecasts and observations for a case in which an event is predicted with a forecast probability of p_i or higher. The hit rate hr_i and the false alarm rate fr_i are defined as follows:

$$hr_i = \frac{A_i}{A_i + C_i} = \frac{\sum_{t=i}^{10} M_t}{M} \tag{10}$$

$$fr_i = \frac{B_i}{B_i + D_i} = \frac{\sum_{t=i}^{10} (N_t - M_t)}{N - M}. \tag{11}$$

Here, hr_i represents the proportion of predicted phenomena out of the total number of phenomena that actually occur, and is called the hit rate. It should be noted that the proportion of predicted phenomena that occur to the total number of forecasts is also often referred to as the hit rate. The hit rate for an ROC curve refers to the level of coverage, which represents the degree to which phenomena are covered by forecasts. Meanwhile, fr_i represents the proportion of predicted phenomena out of the total number of phenomena that do not occur, and is called the false alarm rate. In an ROC diagram, the point corresponding to each threshold is defined by the hit rates on the vertical axis and the false alarm rates on the horizontal axis, and the ROC curve is drawn accordingly.

A plot in the upper left corner of an ROC diagram represents a perfect forecast (100% hit rate, 0% false alarm rate). Conversely, the diagonal line where the hit rate and the false alarm rate are the same represents a forecast with no skill. Accordingly, the closer the curve is to the upper left corner, the higher the skill of the forecast. The extent of the area under the ROC curve is used as a metric to evaluate forecast skill, and is called the AUC (area under curve) or the ROC area. If $AUC > 0.5$, the forecast has a certain level of skill, and AUC is at its maximum ($AUC=1$) for a perfect forecast.

		Observation	
		Yes	No
Forecast	Yes	Ai	Bi
	No	Ci	Di

Table 1 Correspondence between forecast and observation for a case in which a phenomenon was predicted with a forecast probability of p_i or higher

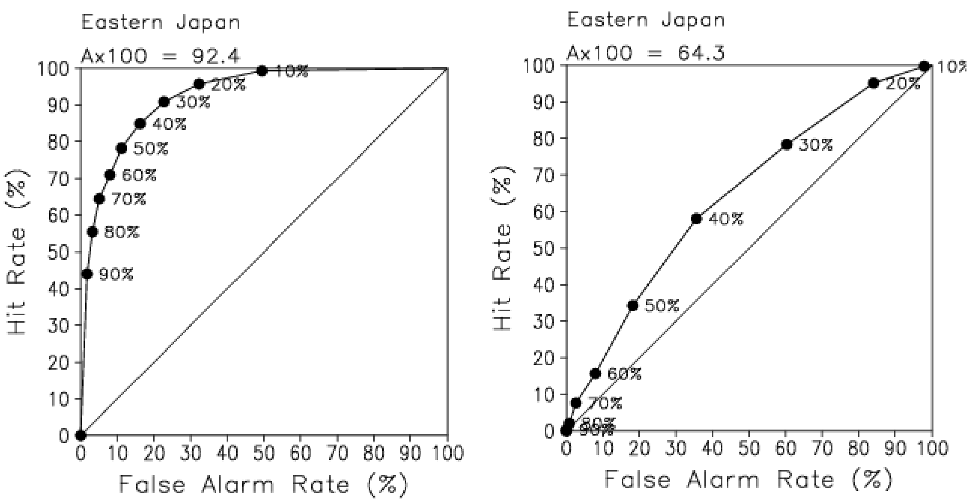


Figure 7 ROC curves of JMA guidance for temperature in eastern Japan

Left: 1st week; right: 3rd – 4th week

The curves show combinations of hit rates and false alarm rates based on forecast probability values.

The values of A × 100 at the top of each panel denote the ROC area (%).

Producing Site-specific Guidance Using Domestic Data

Producing Site-specific Guidance Using Domestic Data

Climate Prediction Division, Japan Meteorological Agency (JMA)

1. Objectives

The objectives of these exercises are as follows:

- To clarify how to produce guidance for surface temperature and precipitation.
- To identify effective predictors.
- To verify guidance based on deterministic and probabilistic methods.
- To produce one-month forecasts using guidance and the latest numerical prediction methods.

2. Procedures

A) Single regression model

B) Multiple regression model

C) Probabilistic forecast

D) Verification

E) Production of one-month forecasts for Nov. 2011

F) Creation of presentations

G) Presentations (10 minutes)

3. Preparation

- Observation data (by trainees)
- Predictors
  - GPV data over trainees' hindcast stations (init. 31st Oct. 1979 – 2010) *HindcastGPV.xls*
  - GPV data for latest prediction (init. 27th Oct. 2011) *LatestPredictionGPV.xls*
- Excel software for producing guidance *Exercise for Guidance.xls*
- Textbook *Exercise for Guidance.doc*

4. Exercises

4-1. Single regression model

Step 1

Open “Exercise for Guidance.xls”.

Paste the observation data into the Temperature/Precipitation worksheet.

Input “=AVERAGE(C4:C35)” in C36 to calculate the normal.

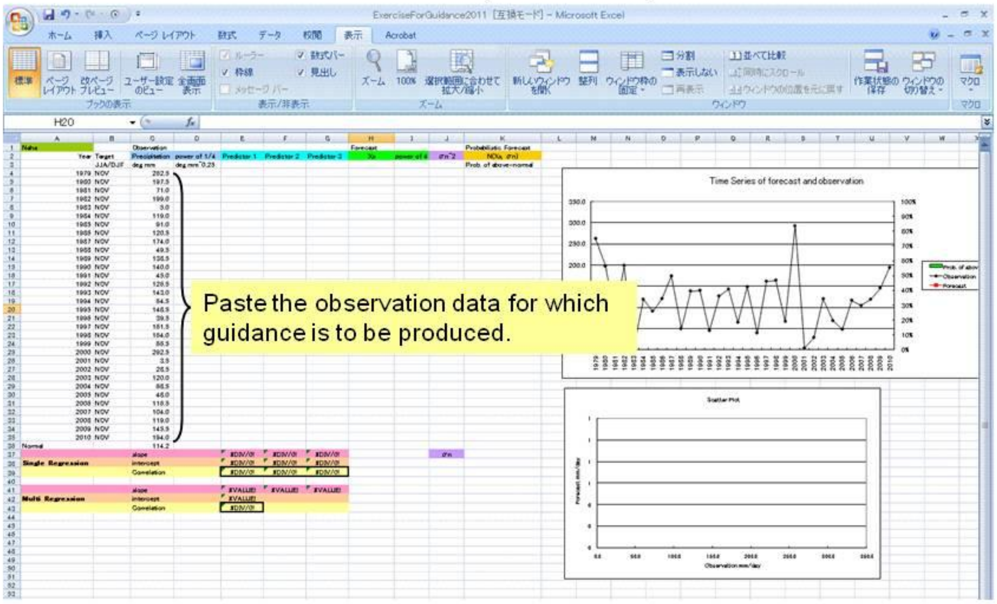
For precipitation, input “=C4^0.25” in D4 to calculate the power of 1/4, then copy D4 to D5:D35.

Then, input “=AVERAGE(D4:D35)” in D36 to calculate the normal.

Exercise for Guidance.xls

1. Single Regression Model

- Open “Exercise for Guidance.xls”.
- Paste the observation data into the Temperature/Precipitation worksheet.

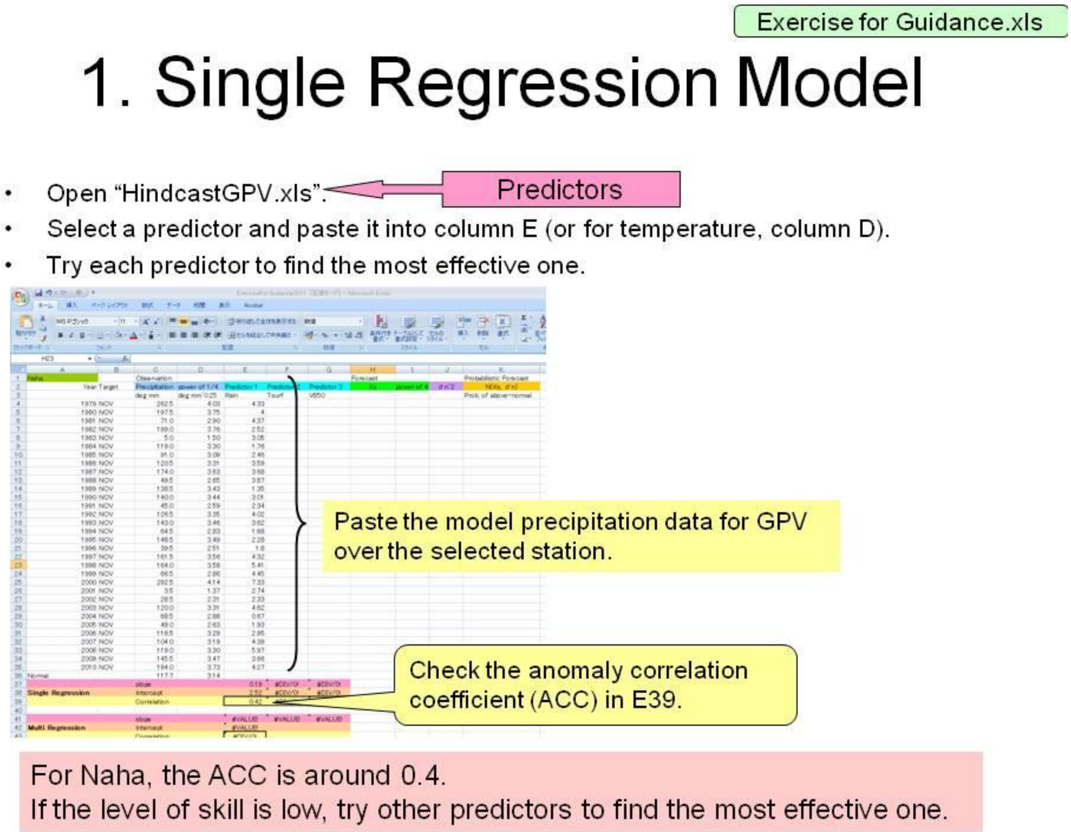


Step 2

Open "HindcastGPV.xls".

Select a predictor and paste it into column E (or for temperature, column D).

Try each predictor to find the most effective one.



Step 3

Check the regression coefficient in E37 and the constant term in E38.

Then, input the regression equation "=$E$37＊$E4+$E$38" in H4.

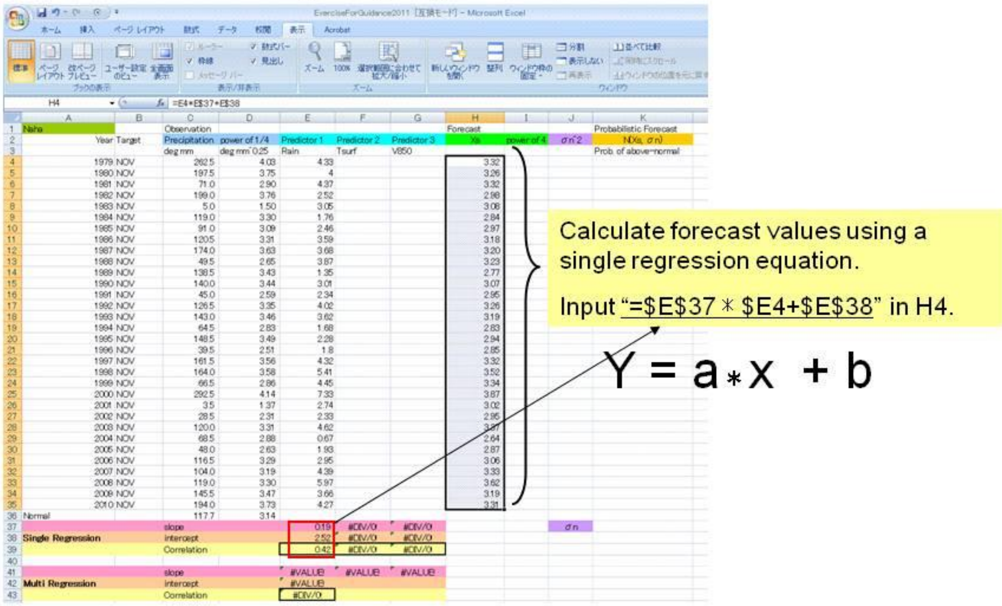
Then, copy H4 into H5:H35.

For temperature, E and H should be replaced with D and G.

Exercise for Guidance.xls

1. Single Regression Model

- Calculate forecasts using a single regression equation.

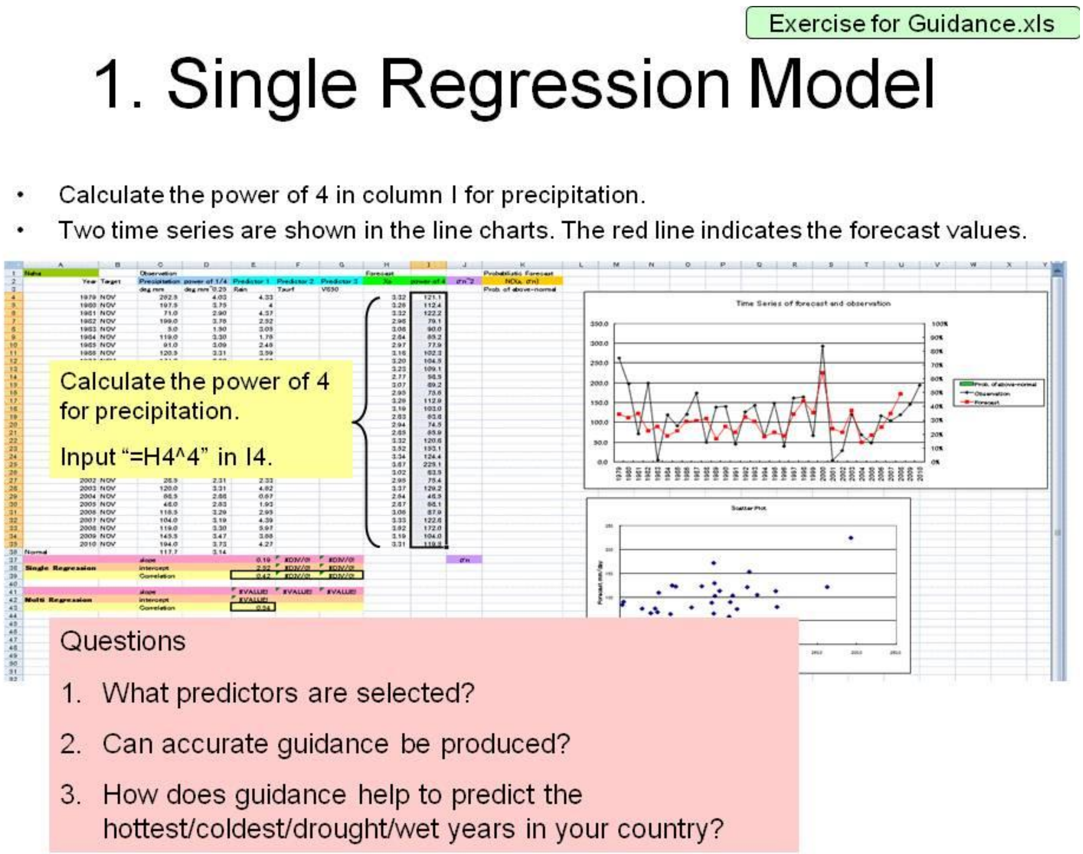


Step 4

This is the end of the exercise using the single regression model.

The time series line chart of the forecast and observation can be viewed.

The anomaly correlation coefficient appears in E39.



Questions

- What predictors are selected?
- Can accurate guidance be produced?
- How does guidance help to predict the hottest/coldest/drought/wet years in your country?

4-2. Multiple regression model

The multiple regression model assumes that the predictand is the sum of the linear combination of the predictors.

It can be written as

$$Y = a_t x_t + b + \varepsilon_t \qquad t = 1,2,3,\ldots\ldots,n,$$

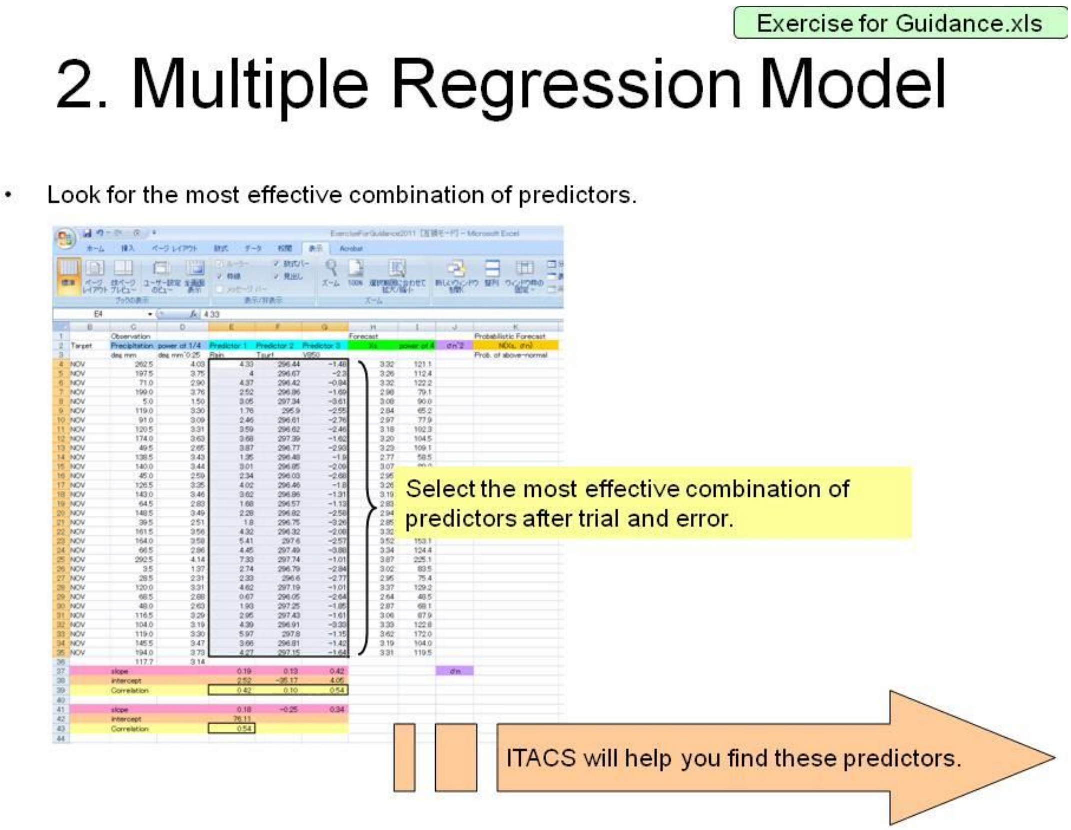
where **Y** is the objective variable (i.e., the predictand), **x** is the predictor, **a** is a regression coefficient, **b** is a constant, and ε is an error term.

Step 1

Open GPVdata.xls.

Select the predictors you consider to be connected with the predictand and paste them into columns E, F and G. Try each predictor to find the most effective combination while checking the anomaly correlation coefficients in E39, F39 and G39.

For temperature, E, F and G should be replaced with D, E and F.



Step 2

If two predictors are selected, input "=$E$41＊$E4+$F$41＊$F4+$E$42" in H4.

If three are selected, input "=$E$41＊$E4+$F$41＊$F4+$G$41＊$G4+$E$42" in H4.

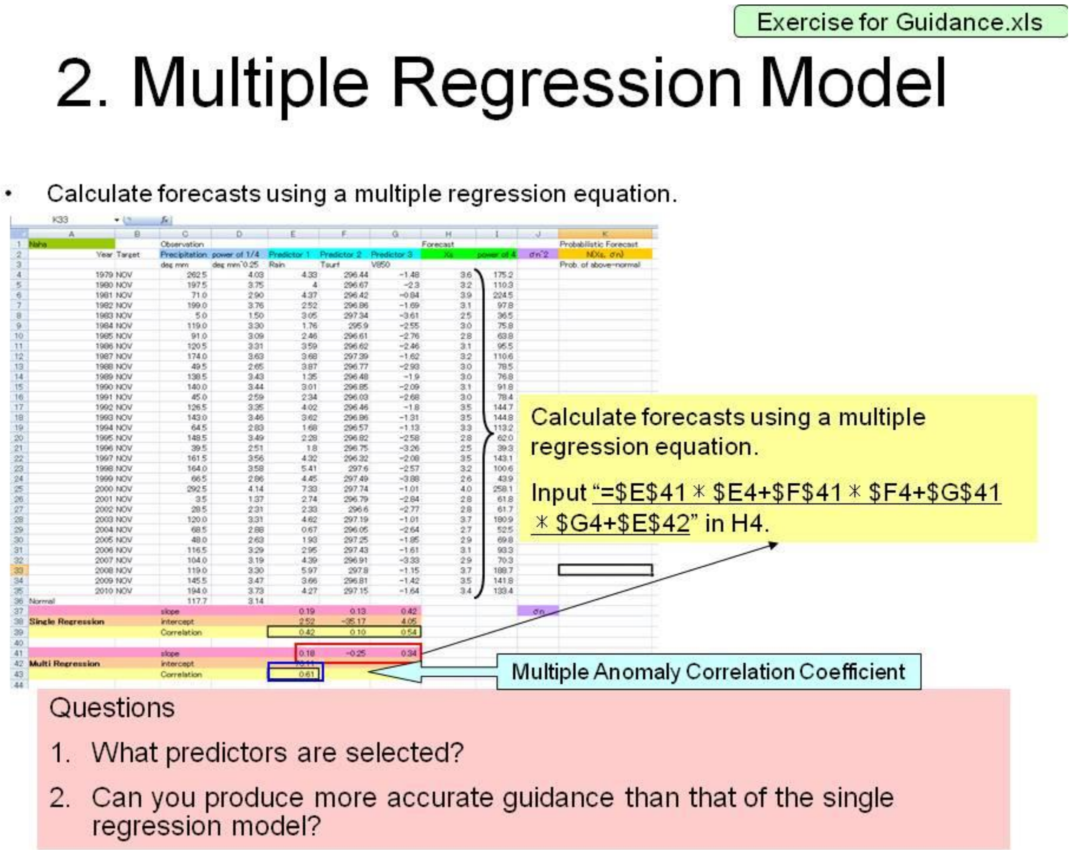
For precipitation, input "=H4^4" in I4 to calculate the power of 4.

Copy I4 into I5:I35.

Check the anomaly correlation coefficient in E43.

If the value of E43 is less than that of E39, F39 and G39, try other predictor combinations to find a higher-accuracy multiple regression equation.

For temperature, E, F, G and H should be replaced with D, E, F and G.



Questions

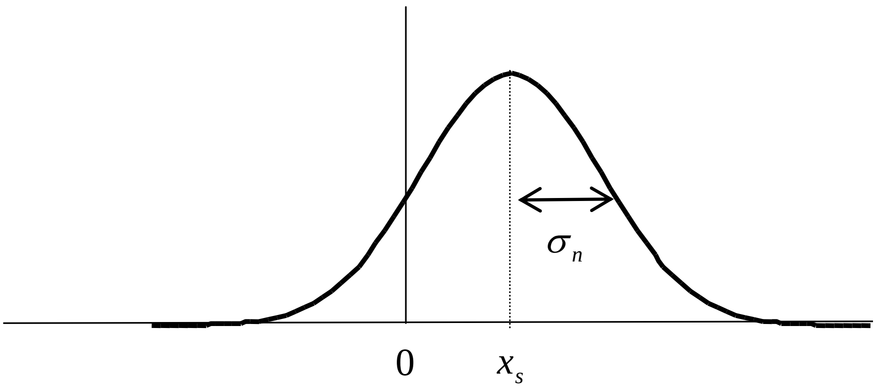
- What predictors are selected?
- Can you produce more accurate guidance than that of the single regression model?

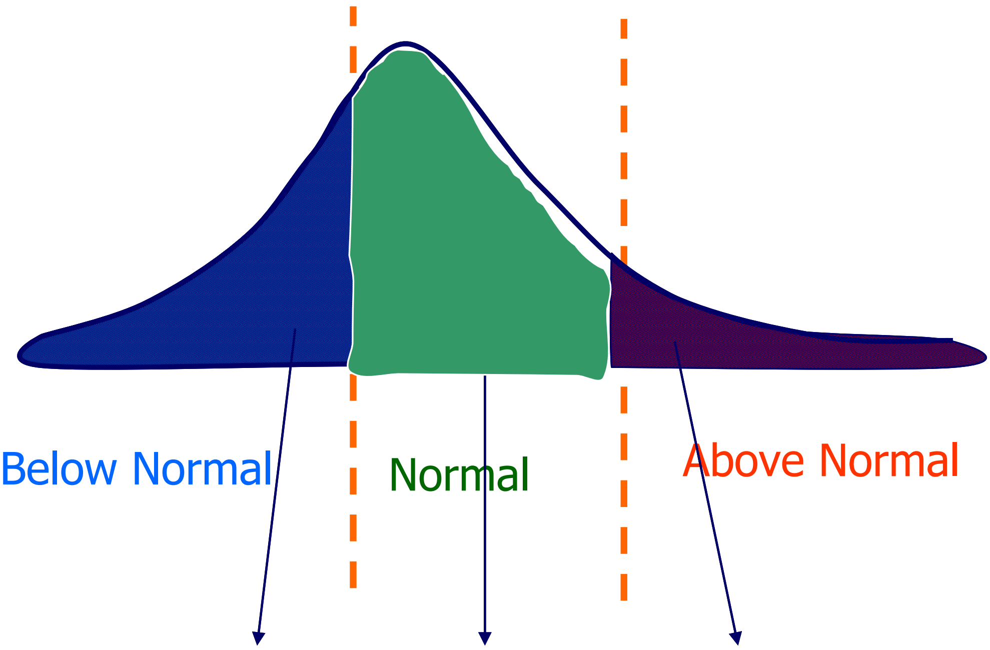
4-3. Probabilistic forecast

Seasonal forecasts involve some level of uncertainty due to the chaotic nature of atmospheric flow. It is therefore necessary to take this uncertainty into account, and the optimal method for this is probabilistic forecasting. Here, we outline probabilistic forecasts issued by the Japan Meteorological Agency (JMA) and related verification.

The probability density function (PDF) is assumed to have normal (Gaussian) distribution with mean x_s and standard deviation σ_n (Fig. 1).
The mean x_s is predicted using a single/multiple regression method, and the standard deviation σ_n is assumed to be the root mean square error of the regression model.

JMA issues a three-category probabilistic forecast based on this guidance around the 25th of every month (Fig. 2), but here we will look at a two-category probabilistic forecast for simplicity.





Below Nomal	Normal	Above Normal
40%	40%	20%

Figure 2 Schematic diagram showing the concept of a three-category forecast
The dashed lines indicate threshold values.

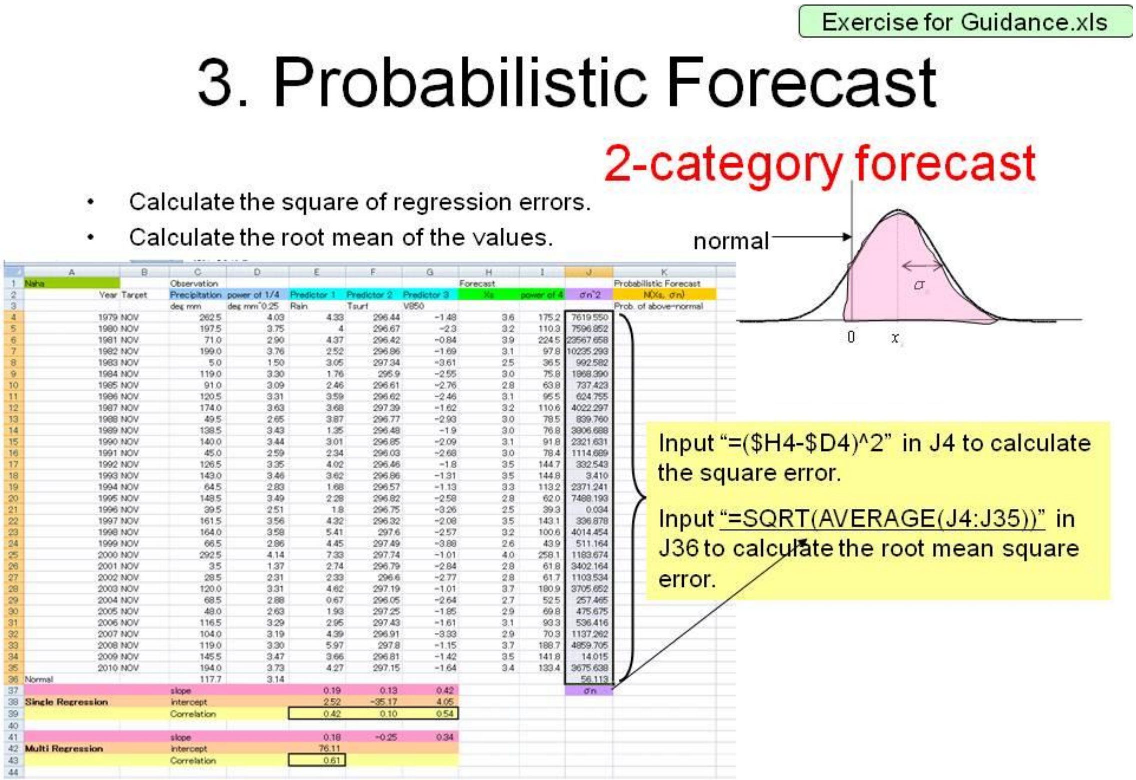
Step 1

To calculate the square of the regression error, input "=($H4-$D4)^2" in J4.

$\varepsilon^2 = \{y - (ax + b)\}^2$

Copy J4 into J5:J35.

Then, to calculate the root mean square error, input "=SQRT(AVERAGE(J4:J35))" in J36. This value will be used as the standard deviation of normal distribution.



Step 2

Normal distribution N(X_s, σ_n) is taken as the probability density function.

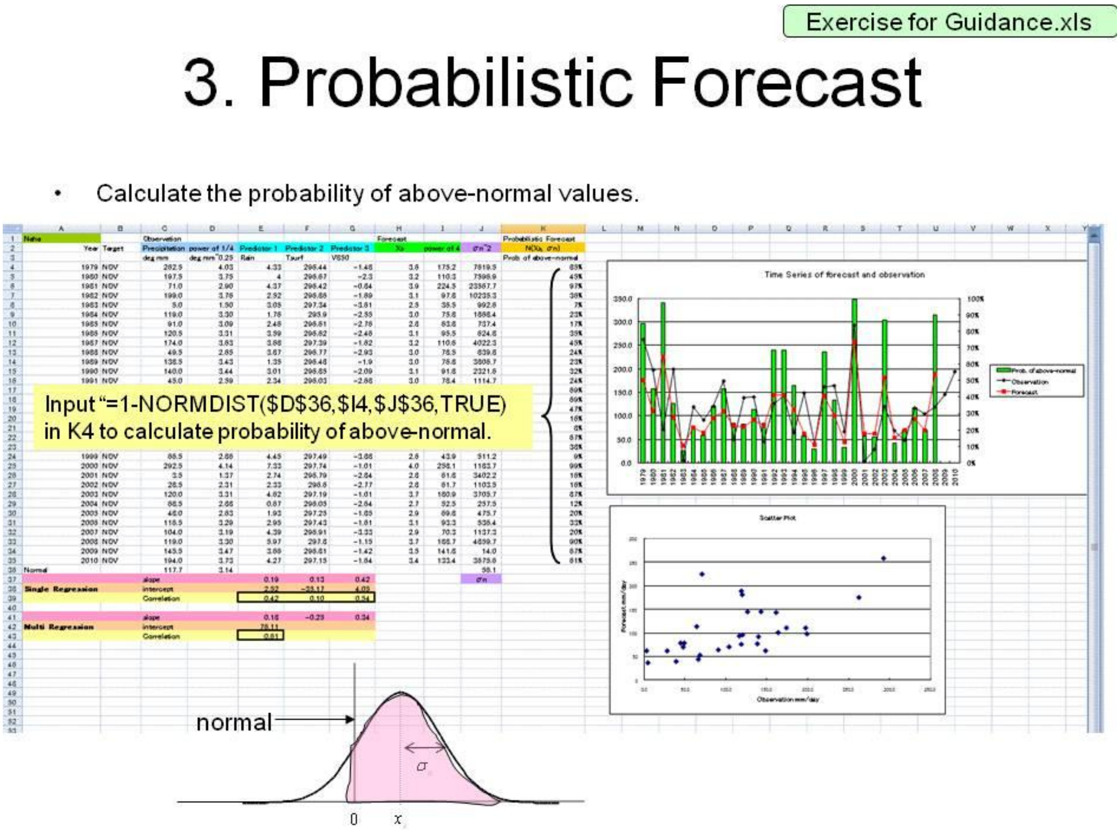
To calculate above-normal probabilities, input

"=1-NORMDIST($D$36,$H4,$J$36,TRUE)" in K4.

Copy K4 to K5:K35.

The values in column K indicate above-normal probabilities.

Finally, the time series bar chart of probabilistic forecasts can be viewed.



4-4. Verification

In many cases, probabilistic forecasts need to be verified, so here we will produce guidance for weather stations. Although it is possible to total all cases for each station and season, it is not necessary to total different variables such as temperature and precipitation because the level of skill is generally different for each variable.

If sufficient data can be gathered, we can make a reliability diagram and calculate the Brier skill score using Excel software.

Step 1

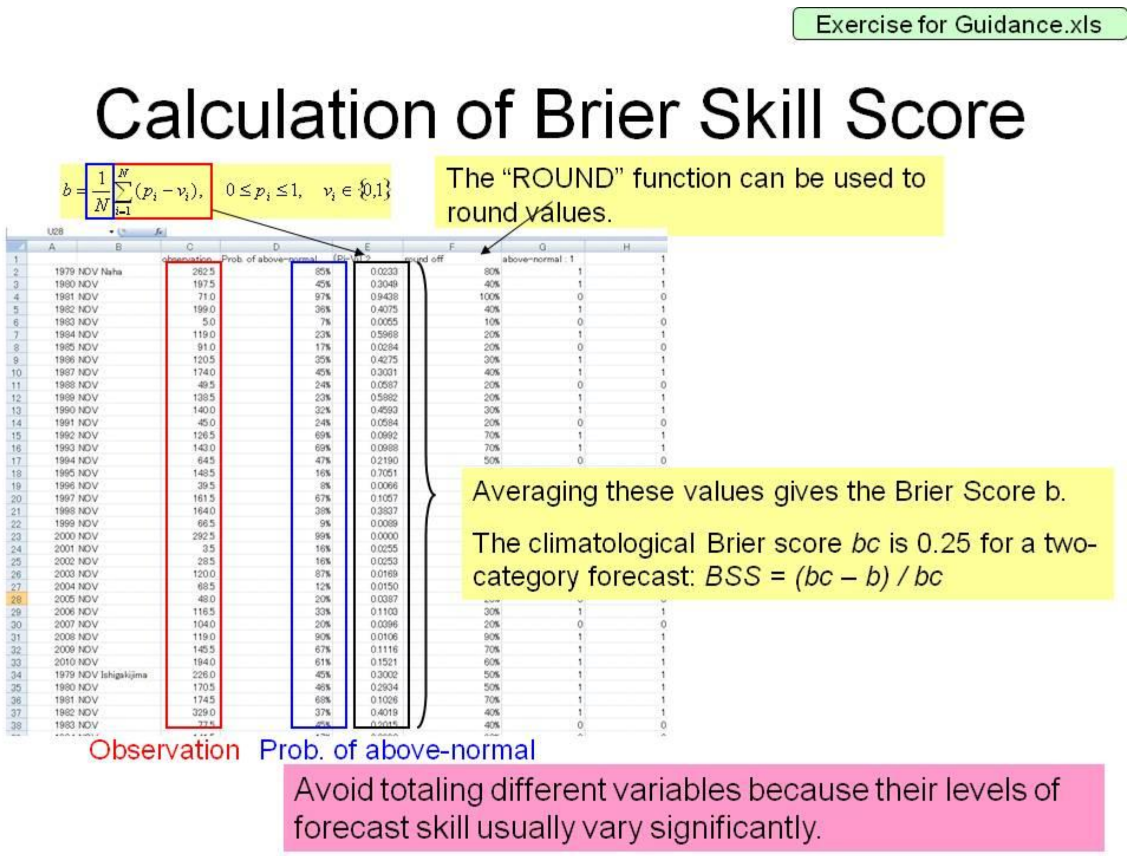
Open "Verification" worksheet.

First of all, probabilistic forecast errors need to be calculated and averaged to obtain the Brier score.

For a two-category forecast, the climatological Brier Score *bc* is 0.25.

Overwriting the samples in columns C and D produces the BSS for the guidance.

Note: Avoid totaling different variables because their levels of forecast skill usually vary significantly.



Step 2

We will now proceed with reliability diagram production.

This is the final step of the verification class.

First, count the number of cases with the same level of forecast probability using the “COUNTIF” function.

Then, count the observation frequency for each forecast probability using the “SUMPRODUCT” function.

The reliability diagram for your guidance can then be viewed.

This is the end of the verification class.