

# CLimate Information Processing system

- What is the CLIPs
- How do we use CLIPs

- Climate Prediction Team, APCC
- Han, Jeongmin ([goal@apcc21.org](mailto:goal@apcc21.org))

2018. 10. 15



# Climate information processing system

## - Why do you need to make CLIPs



- Needs

- Climate and climate application data not only have numerous file formats, but also require enormous time and resources to process.

- Purpose

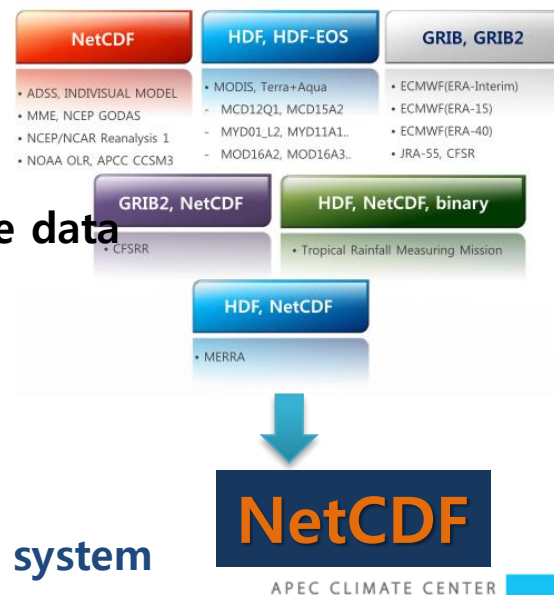
- Development of collect-standard-distribute integrated system based on Climate Data

- Developing plans

- Correlation analysis between application and climate data
  - Support to Indexes for correlation among the data
  - Applying the bigdata processing method

- Final destination

- Construction of integrated knowledge management system
- Data extraction + Visualization + Statistics + Correlation-ship



# Climate information processing system

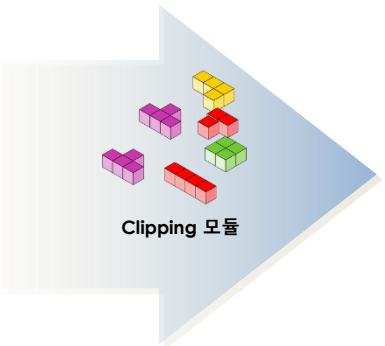
## - System configuration of CLIPs



### CLIST

(data standadization to common format)

- about 1.50million (1949~2020)  
- capacity: 57TB

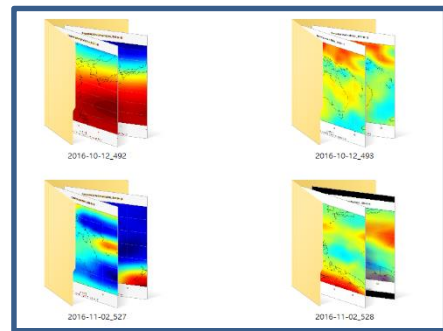
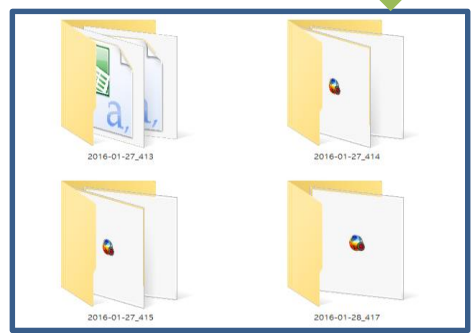


### CLIPs

(data extraction who wants to get area)



Climate Information processing system. (CLIPs)



# Climate information processing system

## - Data distributed parallel processing



Spark Master at spark://bigdata.a



Jobs Stages Storage Environment

URL: spark://bigdata.apcc21.org:7077  
 REST URL: spark://bigdata.apcc21.org:6066 (cluster mode)  
 Alive Workers: 2  
 Cores in use: 16 Total, 16 Used  
 Memory in use: 123.5 GB Total, 8.0 GB Used  
 Applications: 1 Running, 1 Completed  
 Drivers: 0 Running, 0 Completed  
 Status: ALIVE

### Spark Jobs (?)

User: clips  
 Total Uptime: 40 s  
 Scheduling Mode: FIFO  
 Completed Jobs: 4

#### Event Timeline

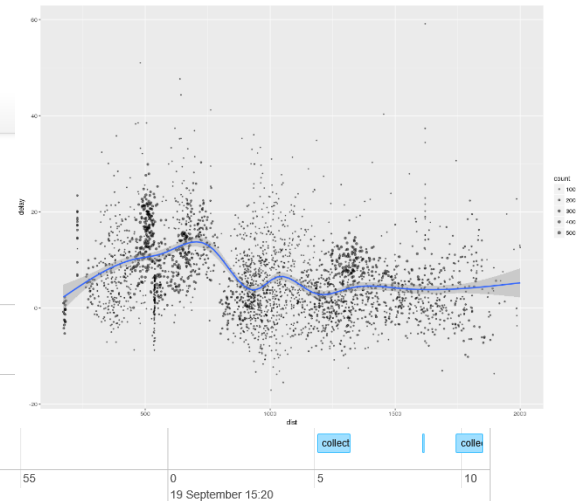
Enable zooming

#### Executors

Added  
 Removed

#### Jobs

Succeeded  
 Failed  
 Running



### Workers

Worker Id	Address
<a href="#">worker-20170911140750-210.98.49.163-59615</a>	210.98.4
<a href="#">worker-20170911140750-210.98.49.164-37510</a>	210.98.4

### Completed Jobs (4)

Job Id	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
3	collect at utils.scala:58	2017/09/19 15:20:09	0.9 s	1/1	1/1
2	collect at utils.scala:196	2017/09/19 15:20:08	91 ms	2/2	2/2
1	sqj at NativeMethodAccessorImpl.java:0	2017/09/19 15:20:06	2 s	2/2	2/2
0	collect at utils.scala:58	2017/09/19 15:20:05	1 s	1/1	1/1

### Running Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
<a href="#">app-20170919151944-0001</a>	(kill) sparklyr	16	4.0 GB	2017/09/19 15:19:44	clips	RUNNING	1.7 min

### Completed Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
<a href="#">app-20170911141307-0000</a>	sparklyr	16	2.0 GB	2017/09/11 14:13:07	clips	FINISHED	3.7 h

# Climate information processing system

## - Main functions



- **Job Management**
  - History management of user's working ( result, taking time, file amount, etc)
- **Search**
  - Searched list which including area , Time period
- **Data Extract**
  - Extracting the type of data and model, element of climate
  - Creating new file extracted from original Files
- **File Transfer**
  - Extracted data Transfer to personal computer
  - File format transformation to NetCDF, ASCII(txt), CSV
- **Reprocessing**
  - Re-extracting from made files in personal computer



# Climate information processing system

## - User interface



The screenshot shows the APCC CLIPs User Interface. At the top, there is a table with columns for Name, Added, Status, Done, Completed On, Extracting, and Downloading. Below the table is a 'New Data Processing' window with a map of East Asia and a configuration panel for coordinates and data parameters.

Name	Added	Status	Done	Completed On	Extracting		Downloading		
					Elapsed	Extracted	Size	Elapsed	Downloaded
sheet-4	2016-01-11 13:11:41	Completed	100.0%	2016-01-11 15:25:57	1 hours 29 min...	5028/5028	416.89322 MB	44 min. 52 sec.	5028/5028
sheet-5	2016-01-13 09:46:17	Completed	100.0%	2016-01-13 09:48:06	28 sec.	388/388	3.356084 MB	1 min. 21 sec.	388/388
sheet-6	2016-01-14 09:35:18	Completed	100.0%	2016-01-14 09:40:24	1 min. 15 sec.	37/37	34.838802 MB	3 min. 49 sec.	30/37

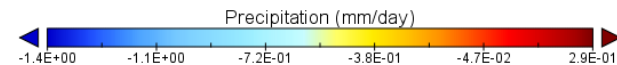
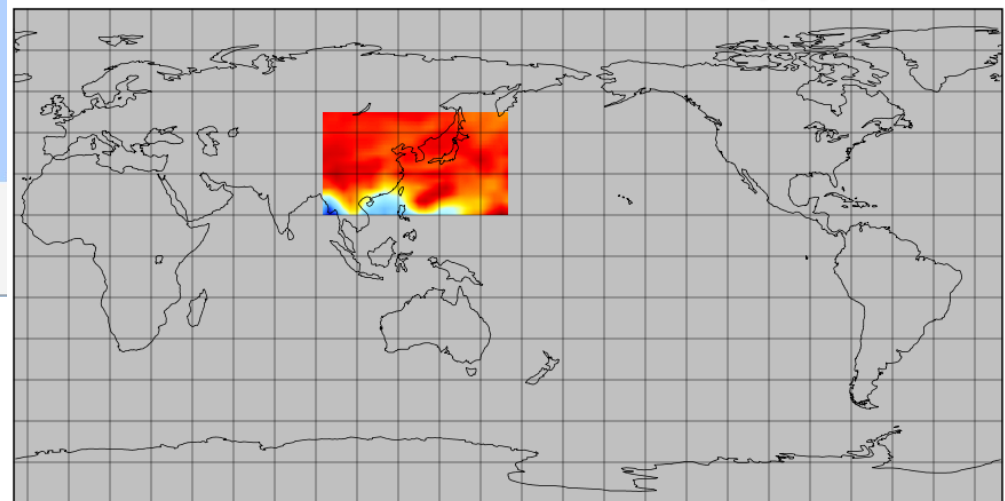
The 'New Data Processing' window includes a map of East Asia with a black rectangle highlighting a region. The configuration panel on the right shows:
 

- Coordinate: 16.9790, 92.5781, 156.3672, 50.0642
- Dataset: APCC-MME-3MON-FORECAST
- Model: CPM
- Time Period: Monthly
- Start Date: 2007-11-15 09:00:00

NASA's Panoply Plot Tool



Precipitation



Data Min = -1.4E+00, Max = 2.9E-01, Mean = -1.3E-01

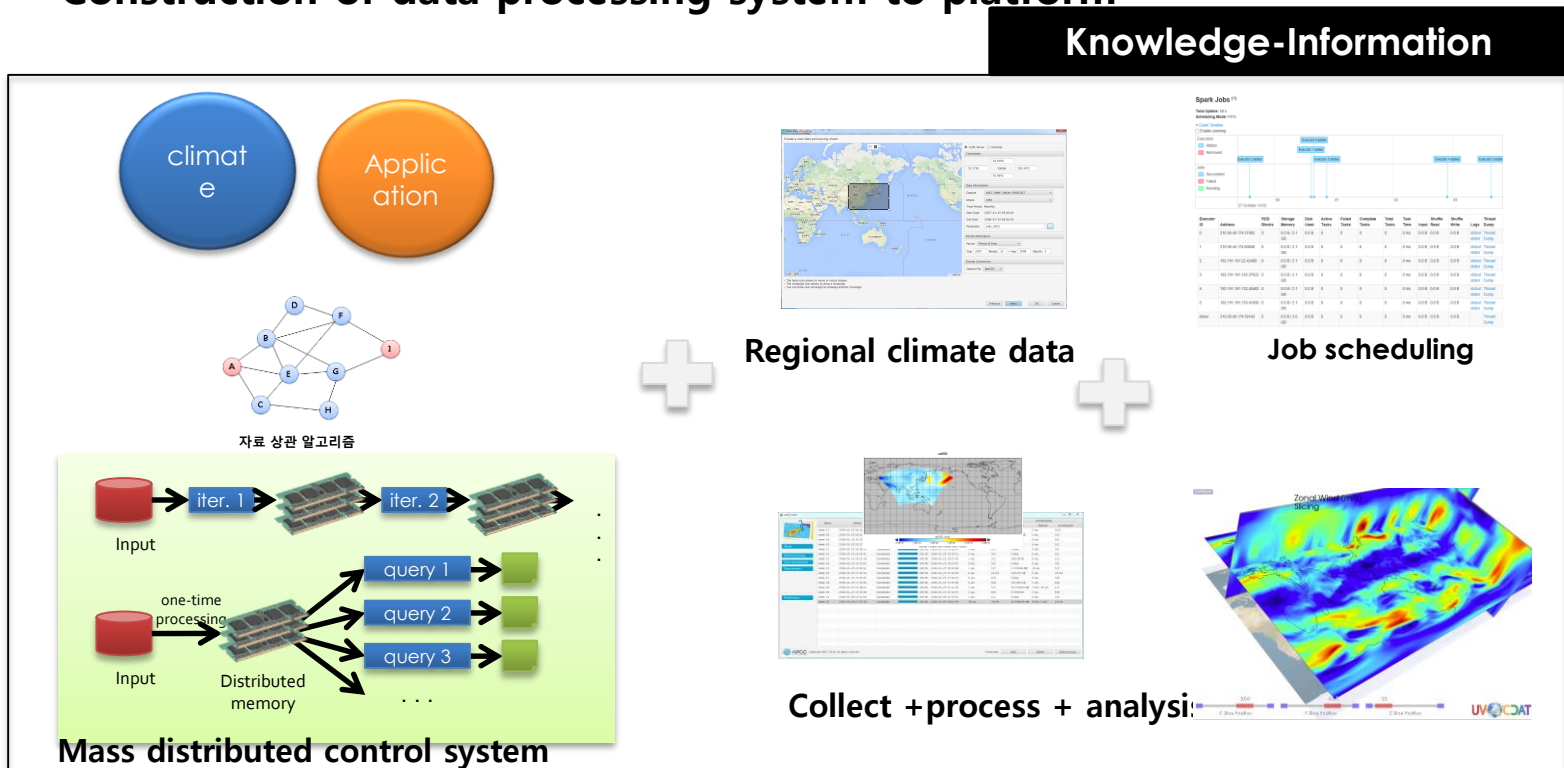
CLIPs User Interface



# Climate information processing system

## - CLIPs Data Analysis platform

- Development of correlation analysis algorithm
  - Climate element + application element
  - Construction of data processing system to platform



\* ETL: Extraction, Transformation, Loading

# Climate information processing system

## - Utilization method



- CLIPs Web Site: <http://clips.apcc21.org>

**Objectives**

Climate and climate application data not only have numerous file formats, but also require enormous time and resources to process. Researchers and policy makers need user-friendly file formats and data extraction services to conveniently utilize climate data. The average user tends to have difficulty dealing with climate data because a single climate data set includes time information, geographical information, and various factor values. Due to these various factors, expertise in data extraction is required. Therefore, the APEC Climate Center (APCC) Climate Information & Application Team has developed the Climate Information Processing System (CLIPs), which enables the average user to conveniently extract the specific data necessary for their research.

**Material**

1.49 million climate change scenarios, climate forecast MME data, climate change scenarios, NOAA Station observation data, and NASA satellite data all from 1960 to 2020, are collected and available on CLIPs. The collected data were formatted with NetCDF, CSV, H5, TXT, and various different formats relative to the characteristics of each dataset and institution.

To provide standardized data, all the data are reformatted with NetCDF (widely used in climate research) and the numerous attributes and units of the data have been converted into standardized units. Additionally, the team has built the database with the extracted temperature, precipitation, wind data and global climate factors.

In order to process large amounts of climate data, the system has two distinct functions, CLIST (Climate data Standardization Toolkit) and CLIPs (Climate Information Processing system) that collect and standardized data, then makes this data available to users. CLIST transforms various file formats into standardized file formats while CLIPs enables users to search for specific locations or values with the installed application on their PCs. Then the combination of these two systems conveys the search results to the users. Finally, for those users residing in those places that lack high speed Internet infrastructure, APCC has applied compression technology to CLIPs so that they can easily access this service.

[Download CLIPs](#)

APCC Climate Center | 11, Centum 7-ro, Yuseong-gu Daejeon 30538 Korea | Copyright(c) APCC All rights reserved

**Clips software**

**Download**

In order to use CLIPs, users have to download the CLIPs client program. This program can work on any version of Windows, but does not work on Macs.

**Software**

Users can download the CLIPs Client Software below and use the application from their desktops. In order to sign-up, the user is required to input basic information (email and file directory).

**Manual**

The CLIPs manual is provided for new users. Both English and Korean versions are available. If a newer version of CLIPs is released, the manual will be updated accordingly.

**Tools**

For user convenience, CLIPs provides information on various application programs. CLIPs does not provide "Plot Service," but provides information on NASA's Panoply Plot Tools.

[CLIPs User Manual \(English\)](#)

[CLIPs User Manual \(Korean\)](#)

[Panoply Plot Tools](#)

CLIPs Client v2.0\_ Released in 2017.

**Clips software**

**Manual(Kor, Eng)**

**Plot tool**

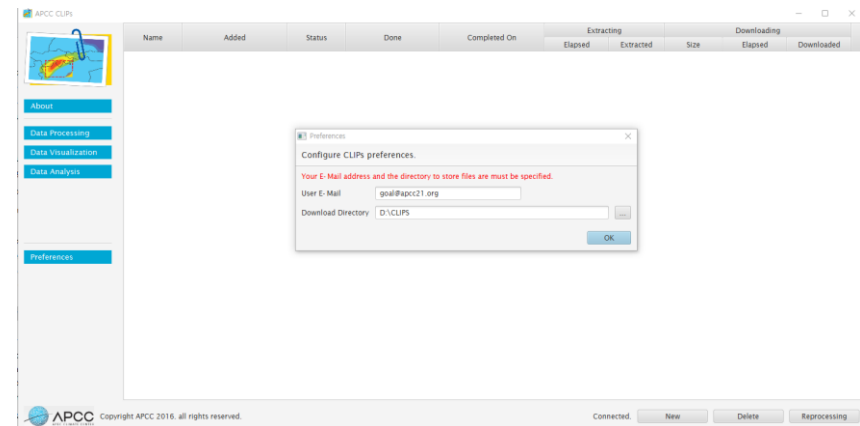
**Using it on personal computer**

# Climate information processing system

## - Installation of CLIPs

- Setting to run
  - Download CLIPs SW from website(<http://clips.apcc21.org>)
  - Unzip the “CLIPs\_Client” to everywhere
  - Double click the “start\_bat” icon
  - input your email and local directory to save

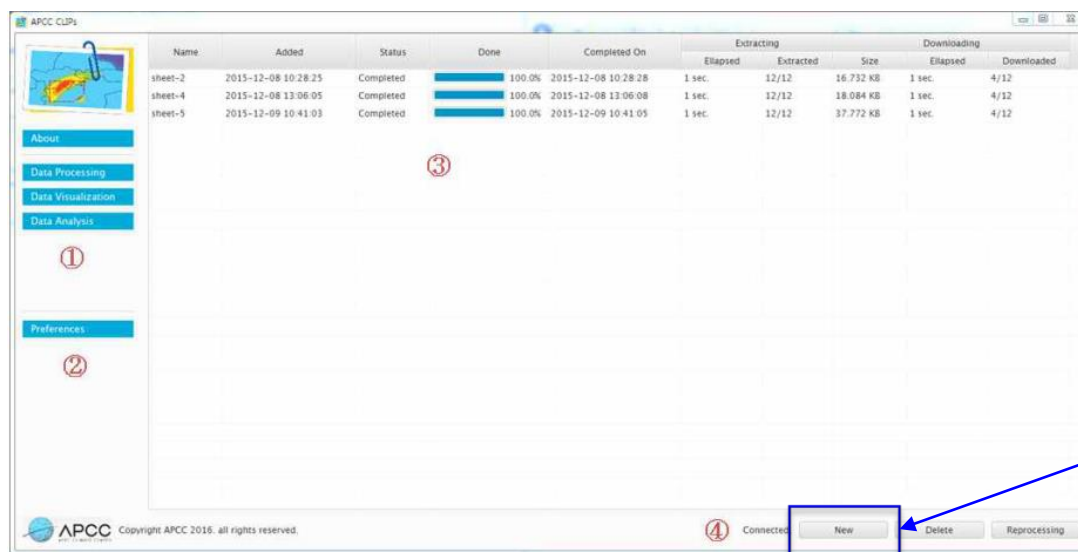
이름	수정된 날짜	유형	크기
yre	2017-02-03 오전...	파일 폴더	
lib	2017-02-03 오전...	파일 폴더	
logs	2018-06-12 오전...	파일 폴더	
org	2017-02-03 오전...	파일 폴더	
win32-x86-64	2017-02-03 오전...	파일 폴더	
derby.log	2018-06-12 오후...	텍스트 문서	2KB
log4j.xml	2016-01-25 오후...	XML 문서	2KB
logo.gif	2015-10-20 오전...	GIF 파일	7KB
preferences.production.properties	2017-02-14 오전...	PROPERTIES 파일	1KB
preferences.production.properties.bak	2017-02-03 오전...	BAK 파일	1KB
start.bat	2016-01-06 오전...	Windows 배치 파일	1KB



# Climate information processing system

## - Runs of CLIPs

- Main user interface and how to Run



Click

- ① CLIPs about and submenu
- ② CLIPs system preference menu
- ③ Display of the current task progress and work history
- ④ Indication of the communication status between the CLIPs server and the user interface.

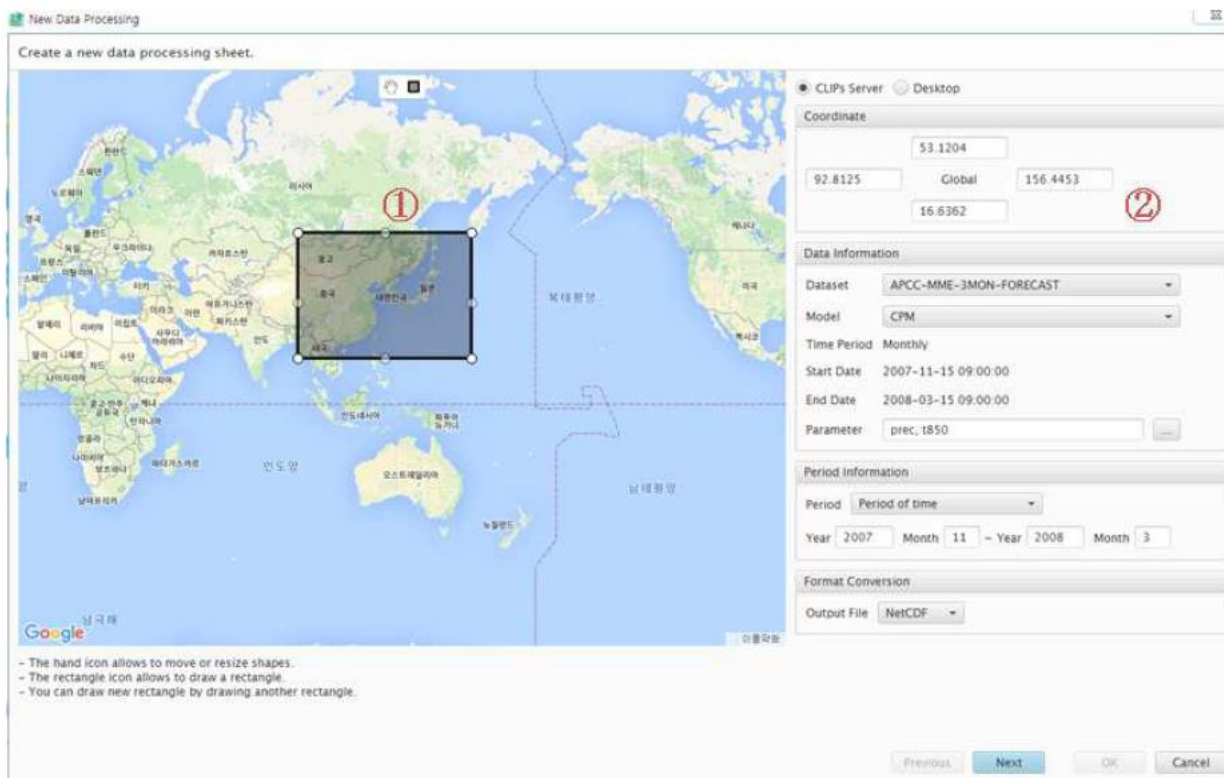
# Climate information processing system

## - Select of Data region on map

- Select of region to want to get data

① Drag and select your target region

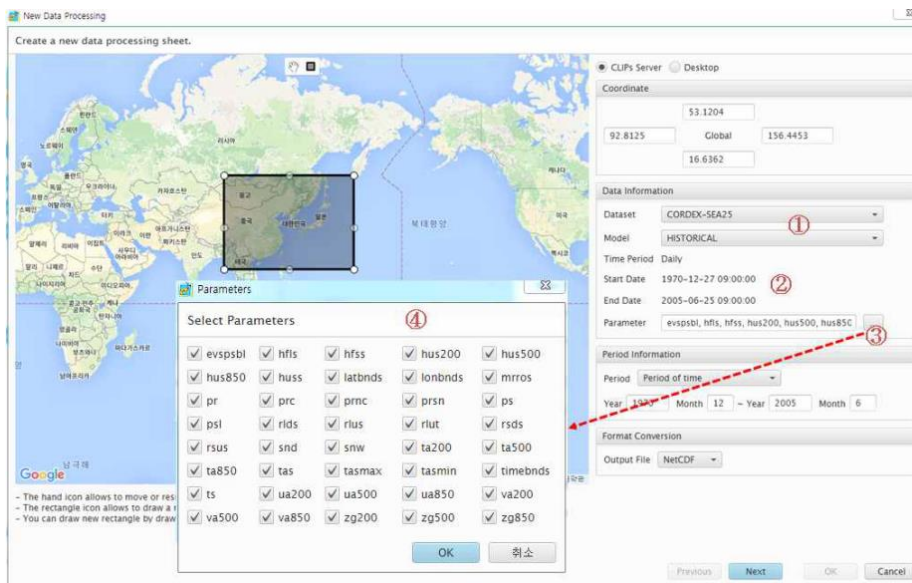
② Verify the coordinates of the selected area in the coordinates panel



# Climate information processing system

## - Select attributes on the tools

- Select the Dataset, Model, parameters

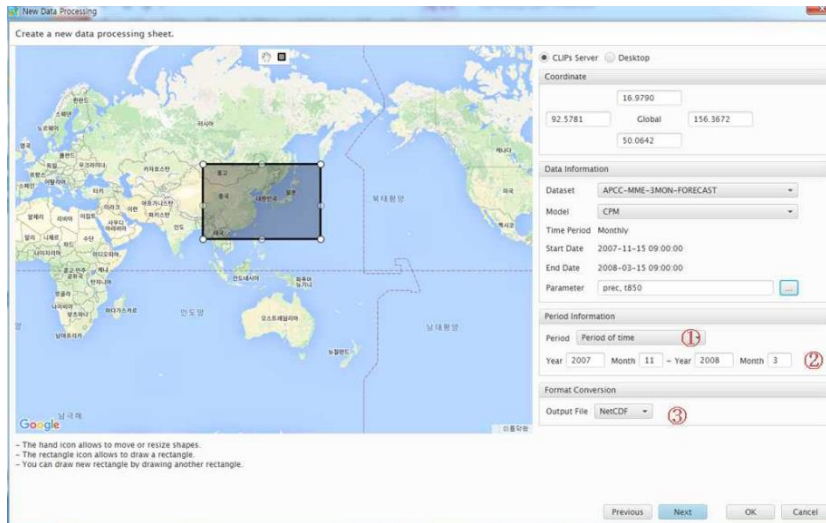


- ① Attributes of the dataset and the model.
- ② Find the time period for a specific model's data.
- ③ Click the button to find the parameter options for that specific model.
- ④ Select the parameters of your choice.

# Climate information processing system

## - Select period on the tools

- Select a period option



- ① Select a period option.
- ② Directly input the time values of the period.
- ③ Select a file format for the extracted data. NetCDF and ASCII (txt) formats are provided by default, and the CSV format is provided when the data is from NOAA.

# Climate information processing system

## - Processed data



- Data extraction and downloading progress indicator

Name	Added	Status	Done	Completed On	Extracting			Downloading	
					Elapsed	Extracted	Size	Elapsed	Downloaded
sheet-2	2015-12-08 10:28:23	Completed	100.0%	2015-12-08 10:28:28	1 sec.	12/12	16.732 KB	1 sec.	4/12
sheet-4	2015-12-08 13:06:03	Completed	100.0%	2015-12-08 13:06:08	1 sec.	12/12	18.084 KB	1 sec.	4/12
sheet-5	2015-12-09 10:41:03	Completed	100.0%	2015-12-09 10:41:05	1 sec.	12/12	37.772 KB	1 sec.	4/12
	2015-12-10 15:58:57	Downloading	50.0%		6 min. 59 sec.	148/148	560.8179 MB	1 sec.	

이름	수정된 날짜	유형	크기
_2004_prec.nc	2015-12-10 오후...	NC 파일	4KB
_2004_1850.nc	2015-12-10 오후...	NC 파일	4KB
_2004_2500.nc	2015-12-10 오후...	NC 파일	4KB
_2005_prec.nc	2015-12-10 오후...	NC 파일	4KB
_2005_1850.nc	2015-12-10 오후...	NC 파일	4KB
_2005_2500.nc	2015-12-10 오후...	NC 파일	4KB
_2006_prec.nc	2015-12-10 오후...	NC 파일	4KB
_2006_1850.nc	2015-12-10 오후...	NC 파일	4KB
_2006_2500.nc	2015-12-10 오후...	NC 파일	4KB
_2007_prec.nc	2015-12-10 오후...	NC 파일	4KB
_2007_1850.nc	2015-12-10 오후...	NC 파일	4KB
_2007_u850.nc	2015-12-10 오후...	NC 파일	4KB
_2007_v850.nc	2015-12-10 오후...	NC 파일	4KB
_2007_2500.nc	2015-12-10 오후...	NC 파일	4KB
_2008_prec.nc	2015-12-10 오후...	NC 파일	4KB
_2008_1850.nc	2015-12-10 오후...	NC 파일	4KB
_2008_u850.nc	2015-12-10 오후...	NC 파일	4KB
_2008_v850.nc	2015-12-10 오후...	NC 파일	4KB
_2008_2500.nc	2015-12-10 오후...	NC 파일	4KB
_2009_prec.nc	2015-12-10 오후...	NC 파일	4KB
_2009_1850.nc	2015-12-10 오후...	NC 파일	4KB
_2009_u850.nc	2015-12-10 오후...	NC 파일	4KB
_2009_v850.nc	2015-12-10 오후...	NC 파일	4KB
_2009_2500.nc	2015-12-10 오후...	NC 파일	4KB
_2010_prec.nc	2015-12-10 오후...	NC 파일	4KB
_2010_slp.nc	2015-12-10 오후...	NC 파일	4KB
_2010_1850.nc	2015-12-10 오후...	NC 파일	4KB

When you set the search conditions and run the program, it extracts the relevant data and transfers it to the directory you designated on your desktop.



# Hands on CLIPs

**Mission 1. Install the CLIPs in your computer**

**Mission 2. Extract the data using CLIPs**



# Question and Supports

- *If you need more help, feel free to email me*
  - Climate Prediction Team at APCC
  - Dr, Han, Jeongmin
  - [goal@apcc21.org](mailto:goal@apcc21.org)

# Thank You

- *APEC Climate Center, Dr. Jeongmin Han*
- *Email: [goal@apcc21.org](mailto:goal@apcc21.org)*