A User Based Approach to Comparing SST Fields

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

[1] University of Rhode Island
[2] University of California/San Diego

GHRSST Meeting
9 June 2008

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' $\Rightarrow$ GMPE; HR-DDS;
    - Advantages: Closer; fast; large suite of statistical analyses available
  - A user-contolled, on-the-fly tool
    - Advantages: Can compare any data sets available to the tool; easy to add new statistical analyses

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' ⇒ GMPE; HR-DDS;
    - Advantages: Global, fast, large suite of statistical analyses available
  - A user-contolled, on-the-fly tool
    - Advantages: User controls any data sets available to the tool, Easy to add new statistical analyses

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' ⇒ GMPE; HR-DDS;
    - Advantages: Closer, fast, large suite of established analysis routines
  - A user-contolled, on-the-fly tool
    - Advantages: User controls any data sets available to the user, Easy to add new statistical analyses

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' $\Rightarrow$ GMPE; HR-DDS;
    - Advantages: Global, fast, large suite of statistical analyses available
  - A user-contolled, on-the-fly tool
    - Advantages: Can compare any data sets available to the tool; easy to add new statistical analyses

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' $\Rightarrow$ GMPE; HR-DDS;
    - Advantages: Global, fast, large suite of statistical analyses available
  - A user-contolled, on-the-fly tool
    - Advantages: Can compare any data sets available to the tool; easy to add new statistical analyses

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' $\Rightarrow$ GMPE; HR-DDS;
    - Advantages: Global, fast, large suite of statistical analyses available
  - A user-contolled, on-the-fly tool
    - Advantages: Can compare any data sets available to the tool; easy to add new statistical analyses

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' $\Rightarrow$ GMPE; HR-DDS;
    - Advantages: Global, fast, large suite of statistical analyses available
  - A user-contolled, on-the-fly tool
    - Advantages: Can compare any data sets available to the tool; easy to add new statistical analyses

The Problem

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

Consider a user . . .

- Who would like to evaluate the output of a general ocean circulation model by comparing predicted SST with that observed in one or more satellite derived fields, or
- Who has been using SST fields available only for the North Atlantic to study mesoscale processes in the vicinity of the Gulf Stream, but would now like to study the Kuroshio using a similar approach, or
- Who like to know how available satellite-derived SST fields compare for a small region, say the western Med, in an effort to choose the one best for her project.
- The two extremes in approaches to dealing with this are:
  - A 'centralize approach' $\Rightarrow$ GMPE; HR-DDS;
    - Advantages: Global, fast, large suite of statistical analyses available
  - A user-contolled, on-the-fly tool
    - Advantages: Can compare any data sets available to the tool; easy to add new statistical analyses

REAP

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- NSF funded a project called REAP, Realtime Environment for Analytical Processing .
- REAP has an ocean portion and a terrestrial portion.
- The ocean portion is to develop an on-the-fly user driven analysis tool for the comparison of SST fields.
- Let's look at the case of a user who wants to compare HYCOM SST fields to satellite-derived fields,

REAP

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- NSF funded a project called REAP, Realtime Environment for Analytical Processing .
- REAP has an ocean portion and a terrestrial portion.
- The ocean portion is to develop an on-the-fly user driven analysis tool for the comparison of SST fields.
- Let's look at the case of a user who wants to compare HYCOM SST fields to satellite-derived fields,

REAP

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- NSF funded a project called REAP, Realtime Environment for Analytical Processing .
- REAP has an ocean portion and a terrestrial portion.
- The ocean portion is to develop an on-the-fly user driven analysis tool for the comparison of SST fields.
- Let's look at the case of a user who wants to compare HYCOM SST fields to satellite-derived fields,

REAP

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- NSF funded a project called REAP, Realtime Environment for Analytical Processing .
- REAP has an ocean portion and a terrestrial portion.
- The ocean portion is to develop an on-the-fly user driven analysis tool for the comparison of SST fields.
- Let's look at the case of a user who wants to compare HYCOM SST fields to satellite-derived fields,

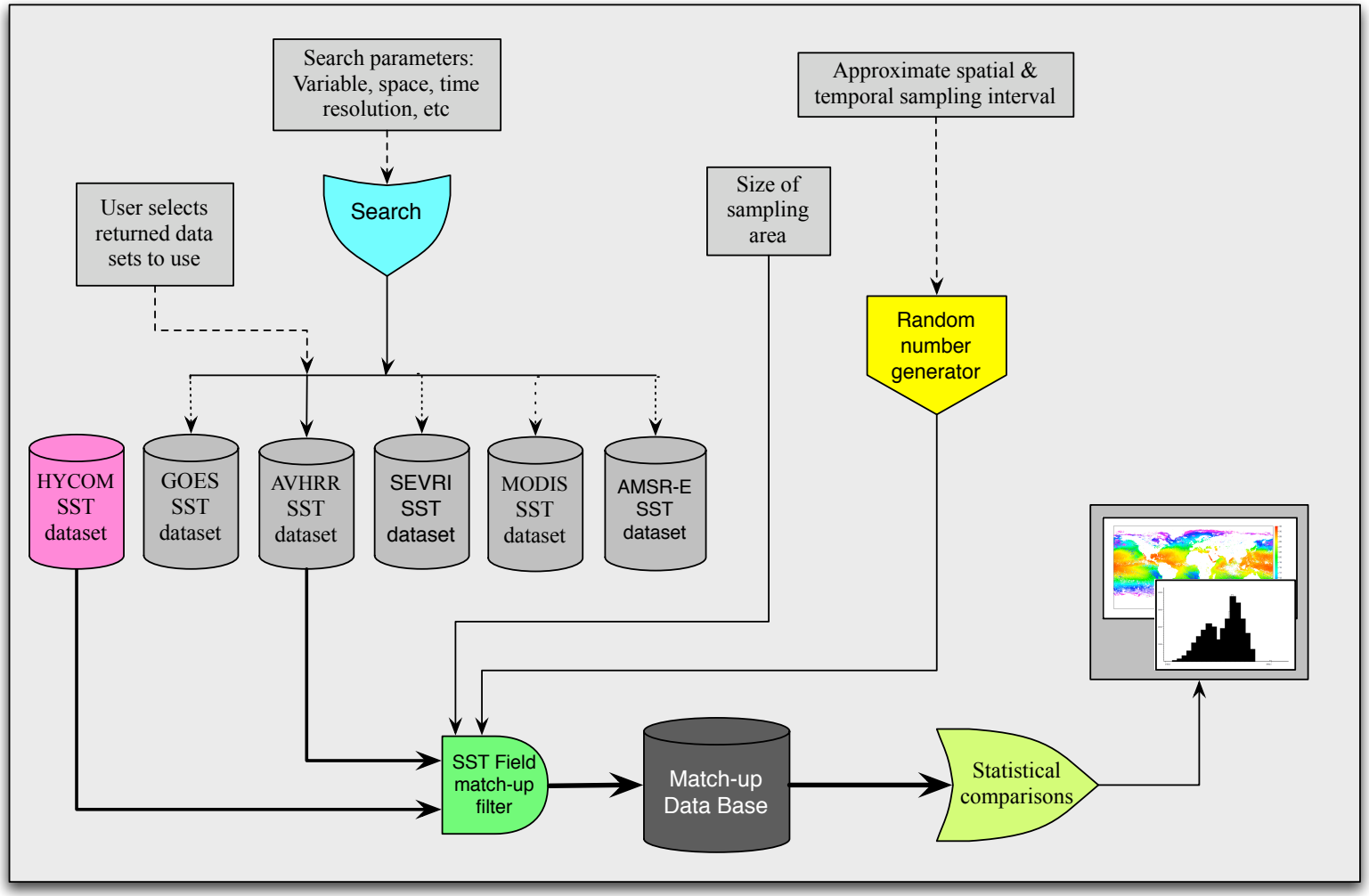
REAP

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- NSF funded a project called REAP, Realtime Environment for Analytical Processing .
- REAP has an ocean portion and a terrestrial portion.
- The ocean portion is to develop an on-the-fly user driven analysis tool for the comparison of SST fields.
- Let's look at the case of a user who wants to compare HYCOM SST fields to satellite-derived fields,
  But the user does not know what satellite-derived fields are available.

The General Idea – Basic Workflow for Grid Comparison

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



Three Modules

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

The workflow can be viewed as consisting of 3 modues:

1. User input and data set selection.
2. Construction of the match-up data set.
3. Analysis.

Three Modules

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

The workflow can be viewed as consisting of 3 modues:

1. User input and data set selection.
2. Construction of the match-up data set.
3. Analysis.

Three Modules

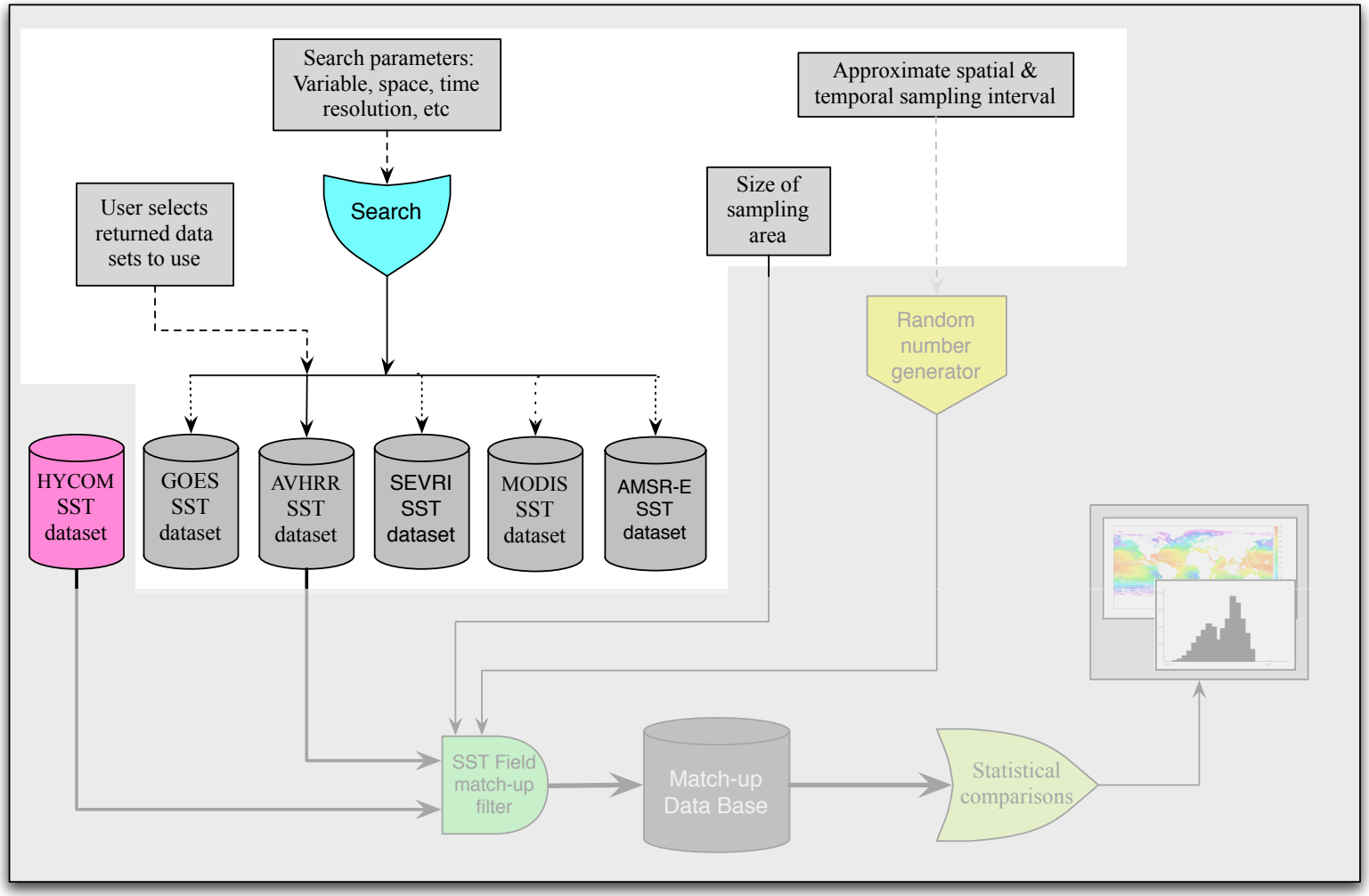
Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

The workflow can be viewed as consisting of 3 modues:

1. User input and data set selection.
2. Construction of the match-up data set.
3. Analysis.

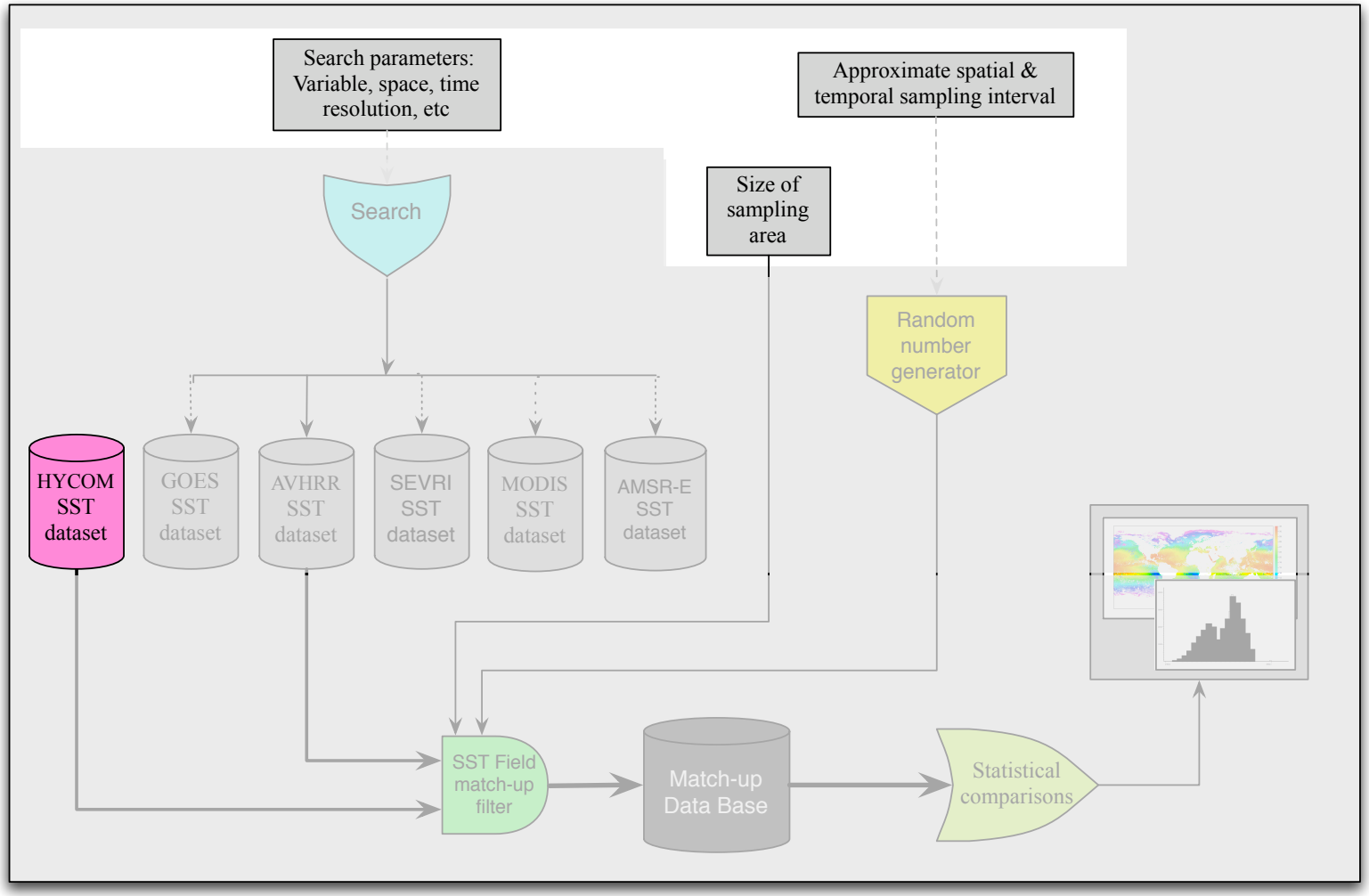
The General Idea – User Entry Module

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



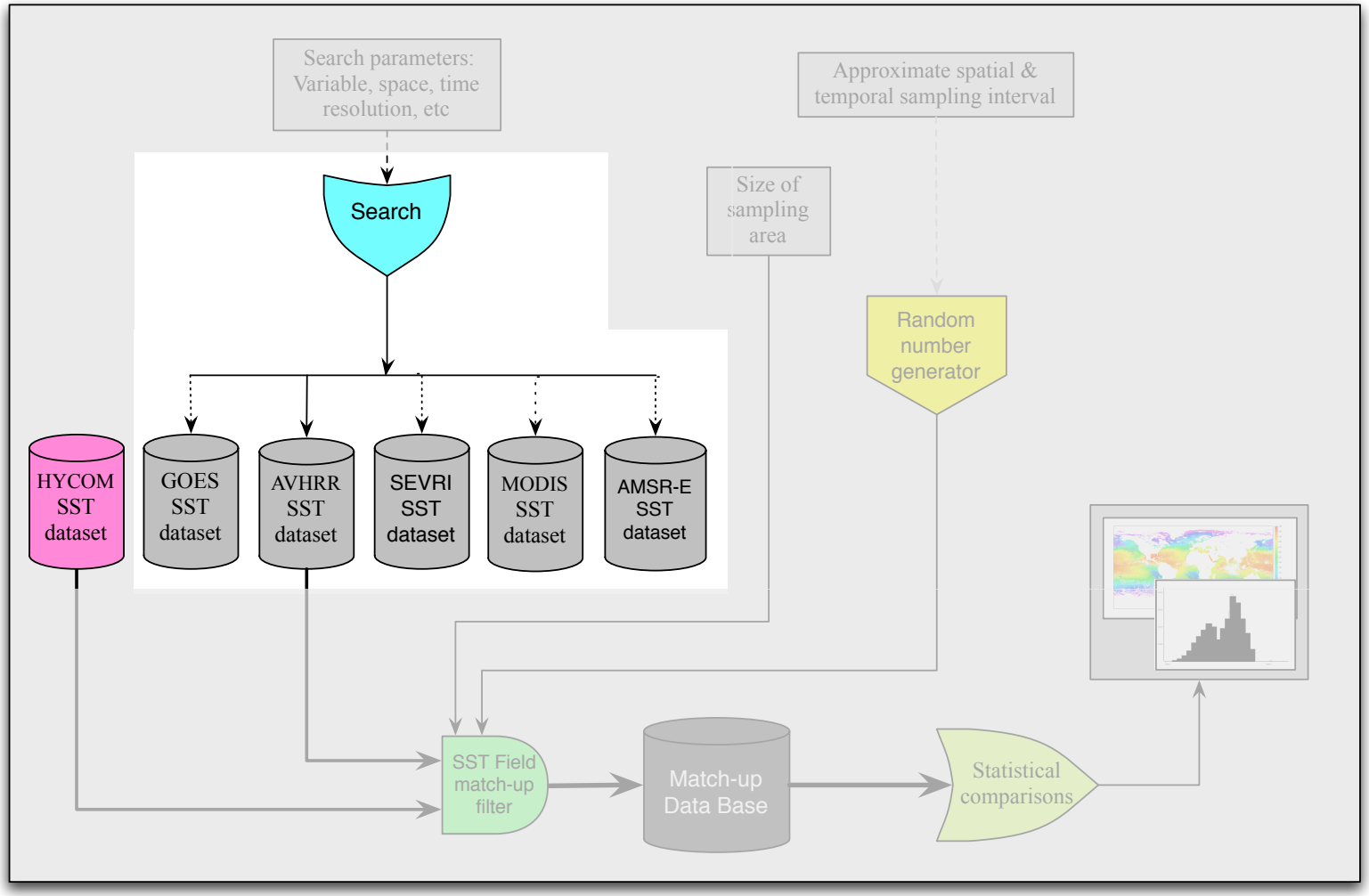
The General Idea – User Entry Module – Enter Basic Parameters

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



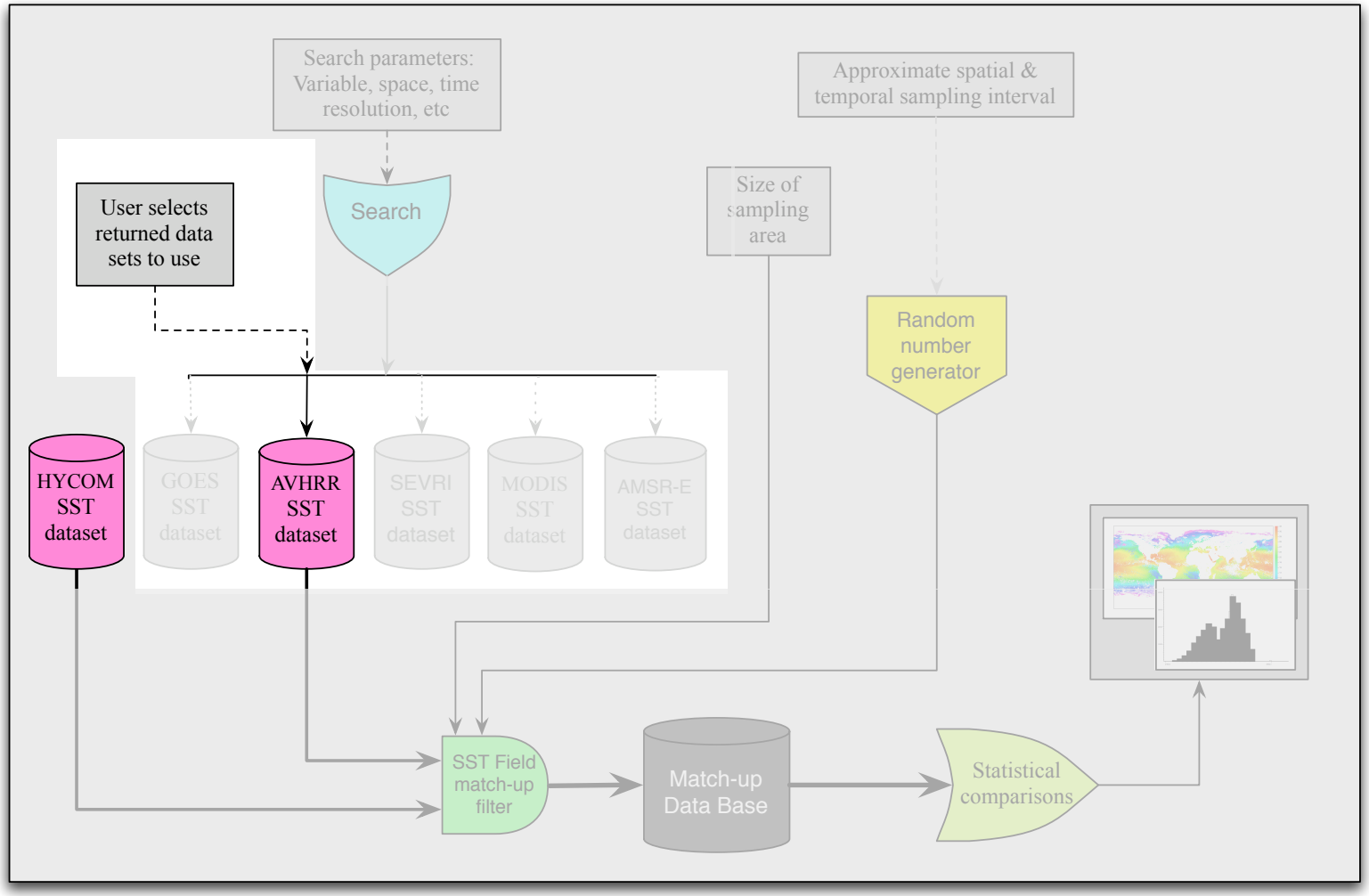
The General Idea – User Entry Module – Search for Data Sets

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



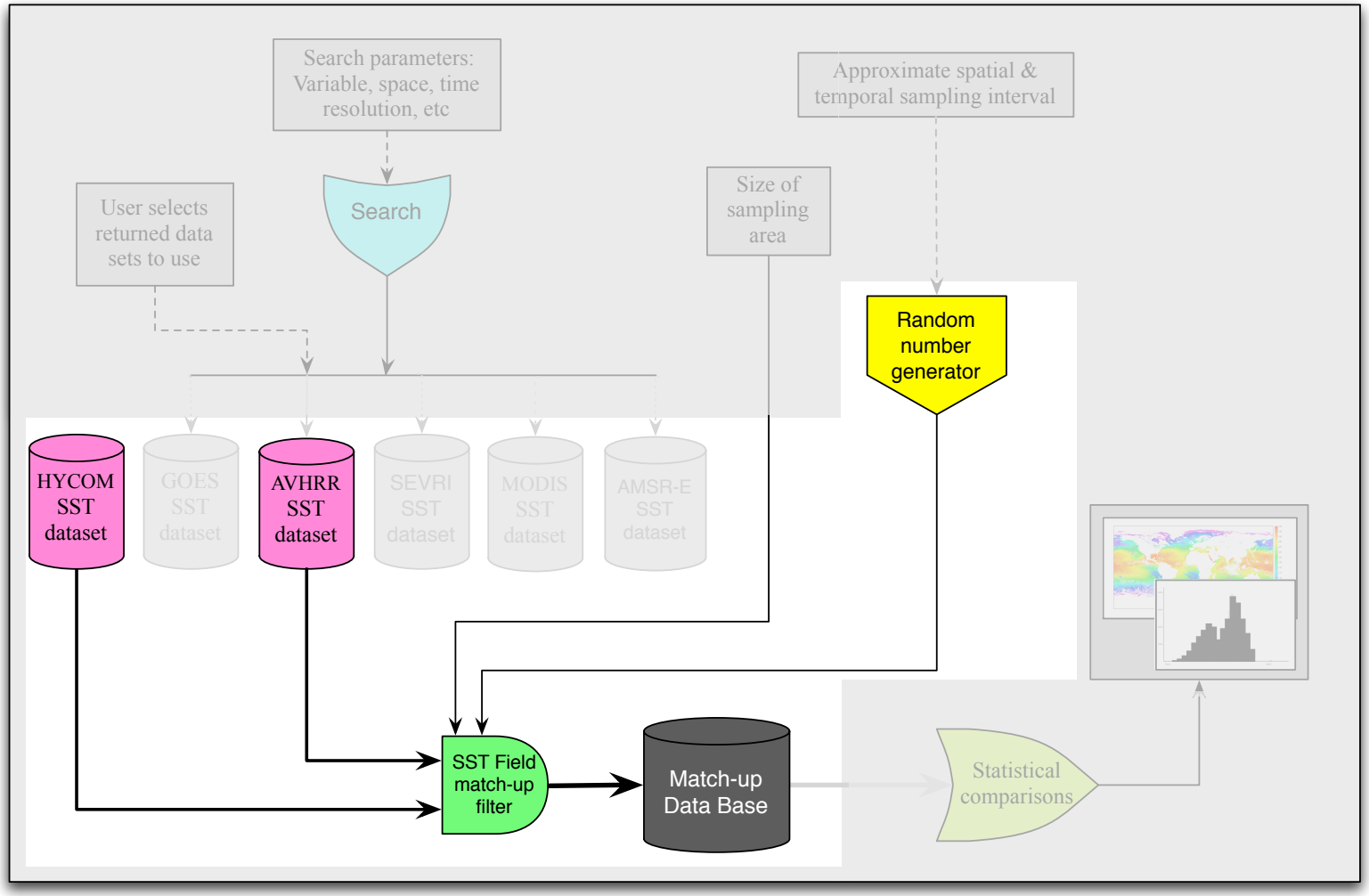
The General Idea – User Entry Module – Select Data Sets

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



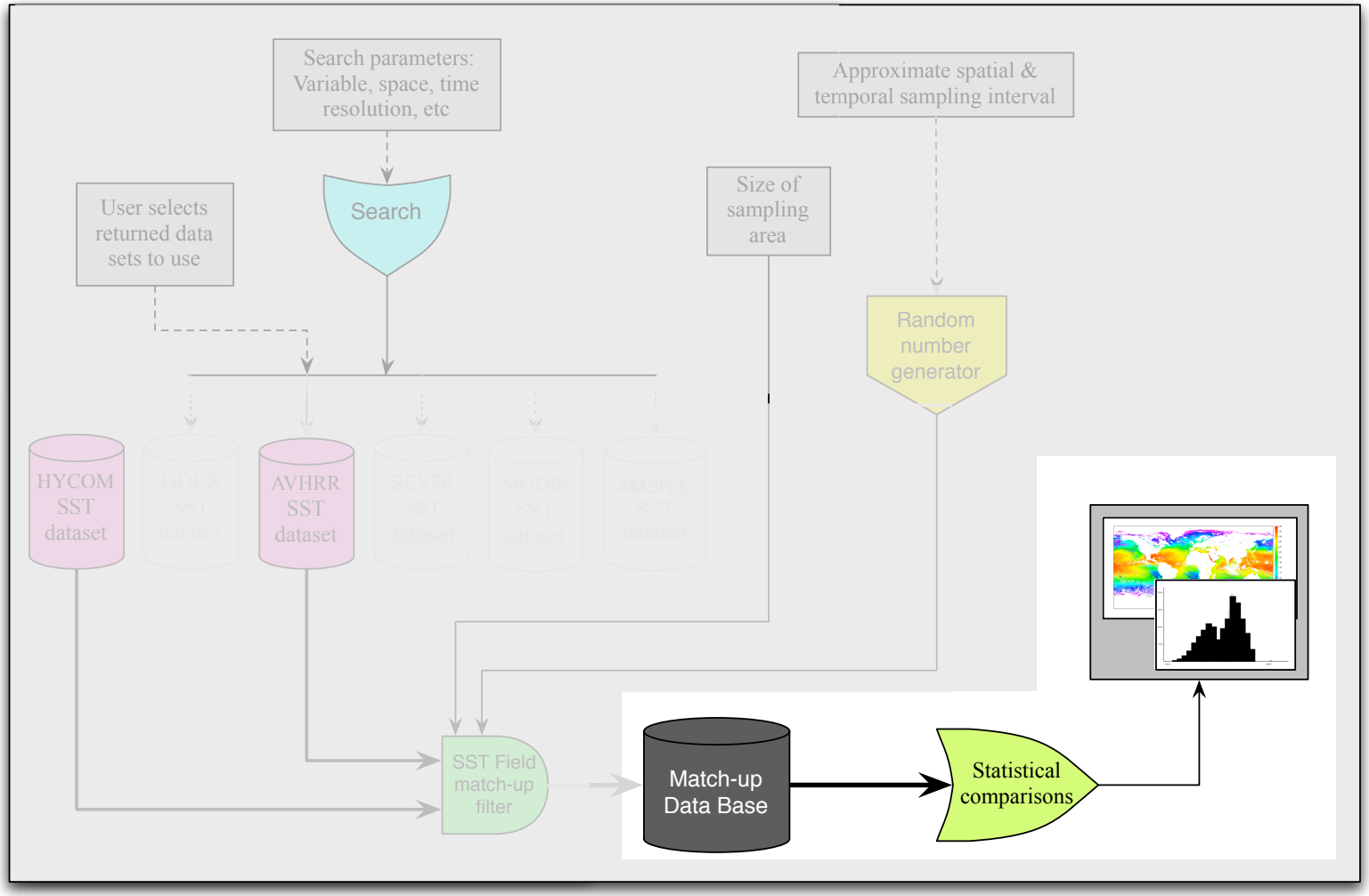
The General Idea – Construct Matchup Database

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



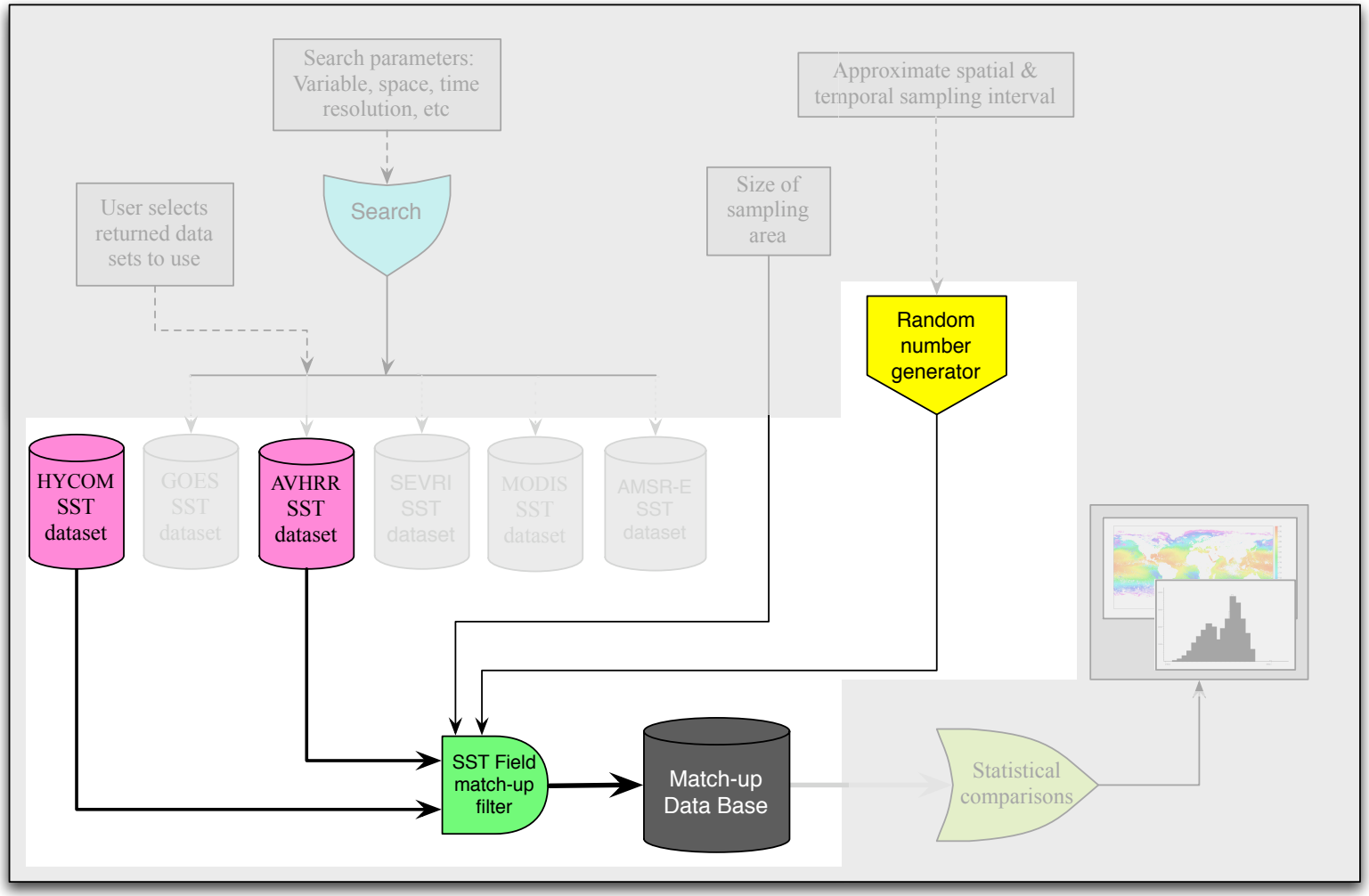
The General Idea – Analysis Module

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



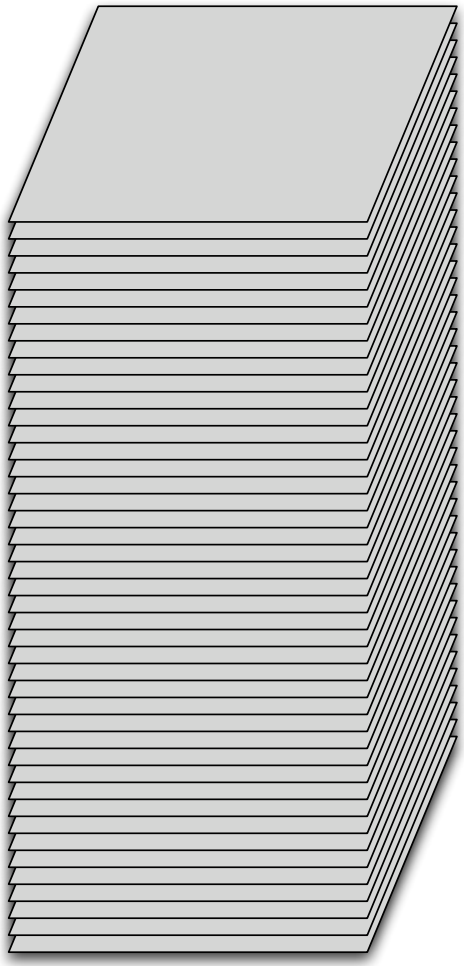
The General Idea – Construct Matchup Database

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

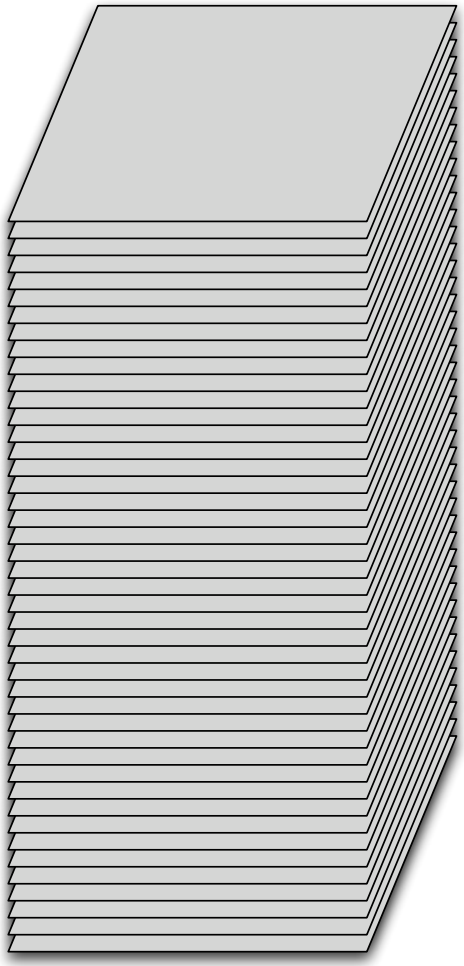


Time Series of SST Fields

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



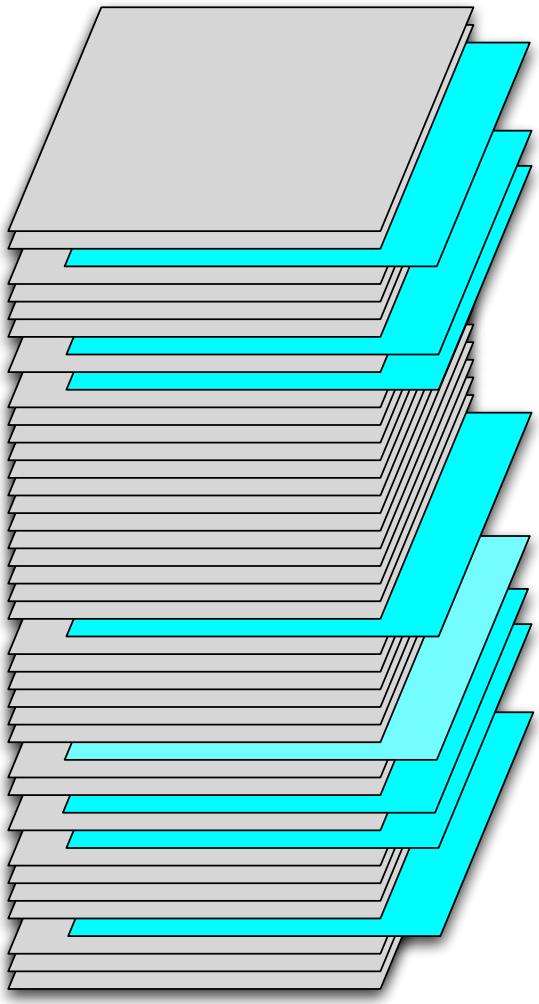
AVHRR



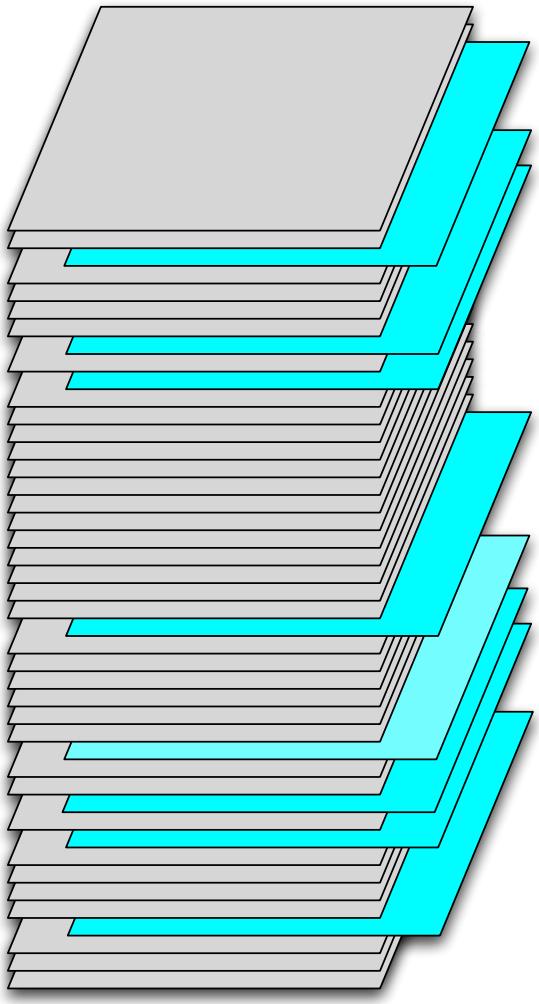
HYCOM

Time Series of SST Fields – 20% Randomly Selected

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



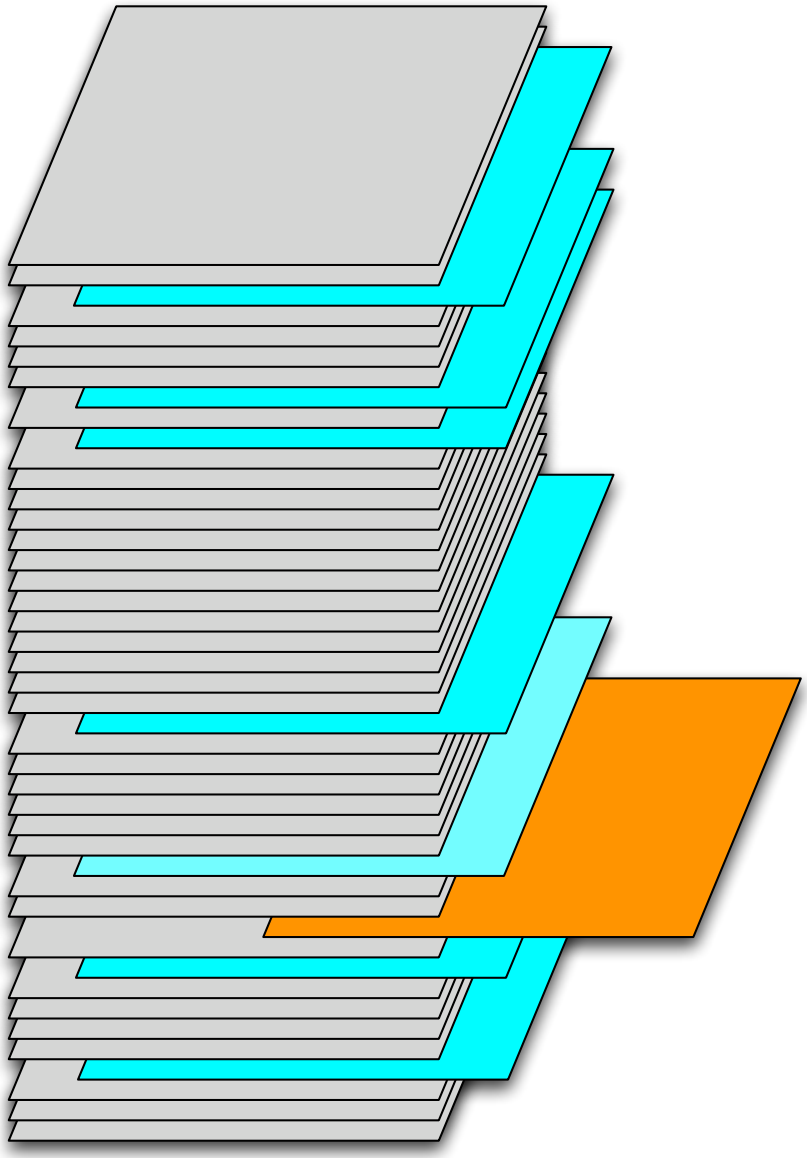
AVHRR



HYCOM

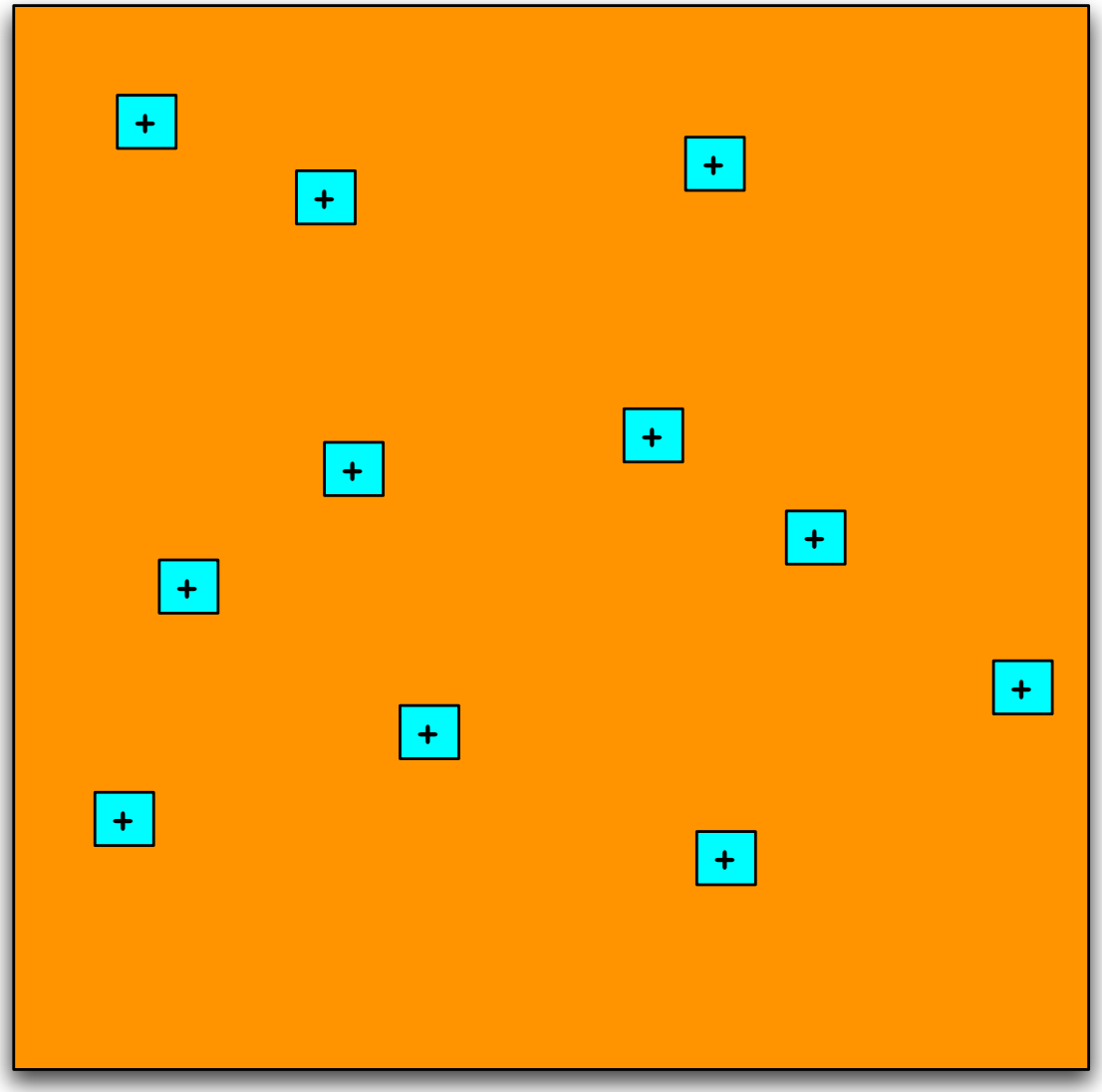
Consider One SST Field

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



Tiles Randomly Selected from One Image

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]



Progress to Date

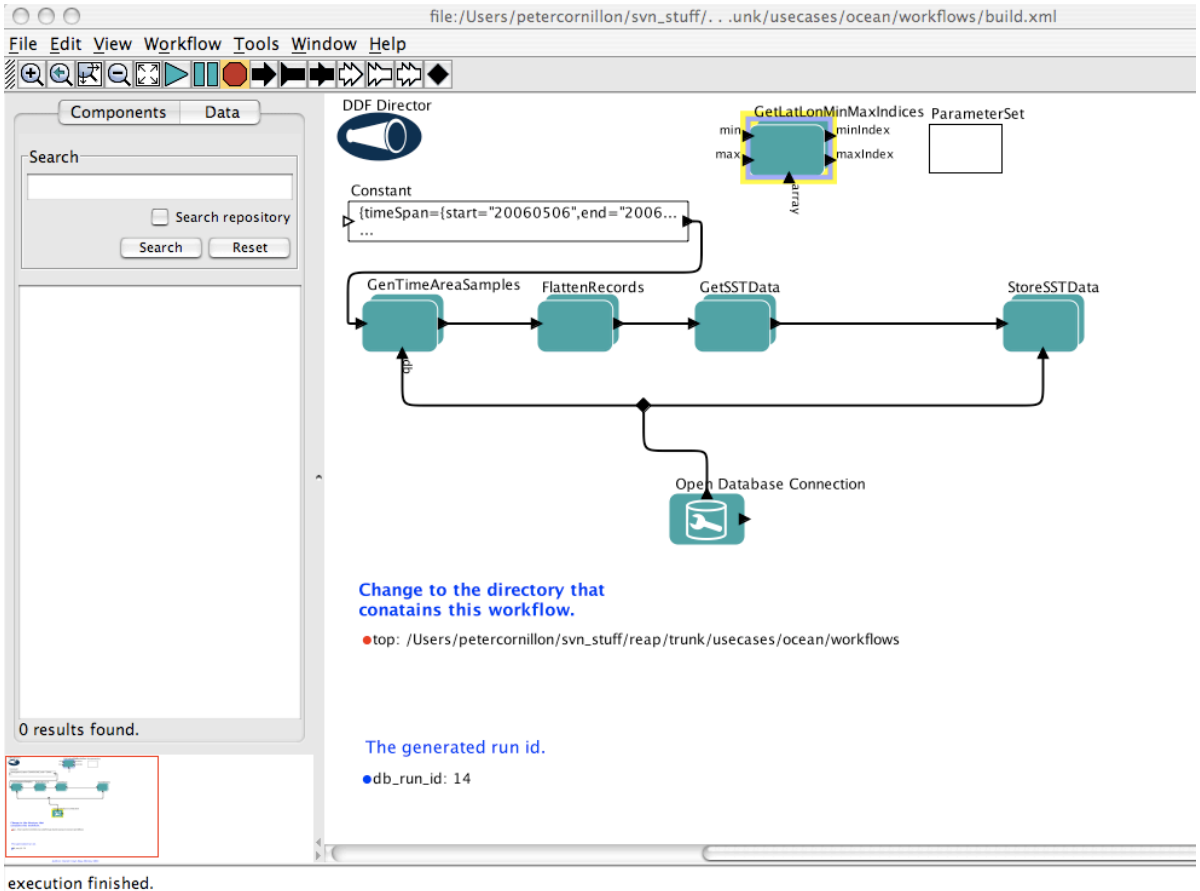
Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU,
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.

Progress to Date

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

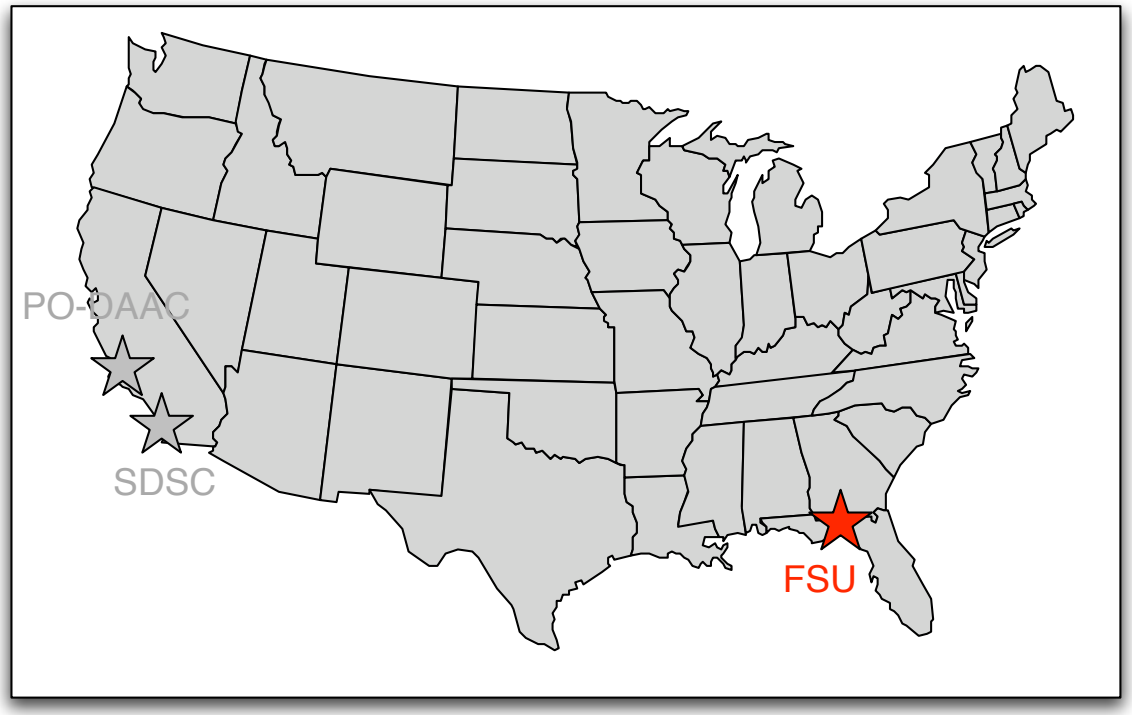
- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU.
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.



Progress to Date

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU.
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.



Progress to Date

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU.
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.

Progress to Date

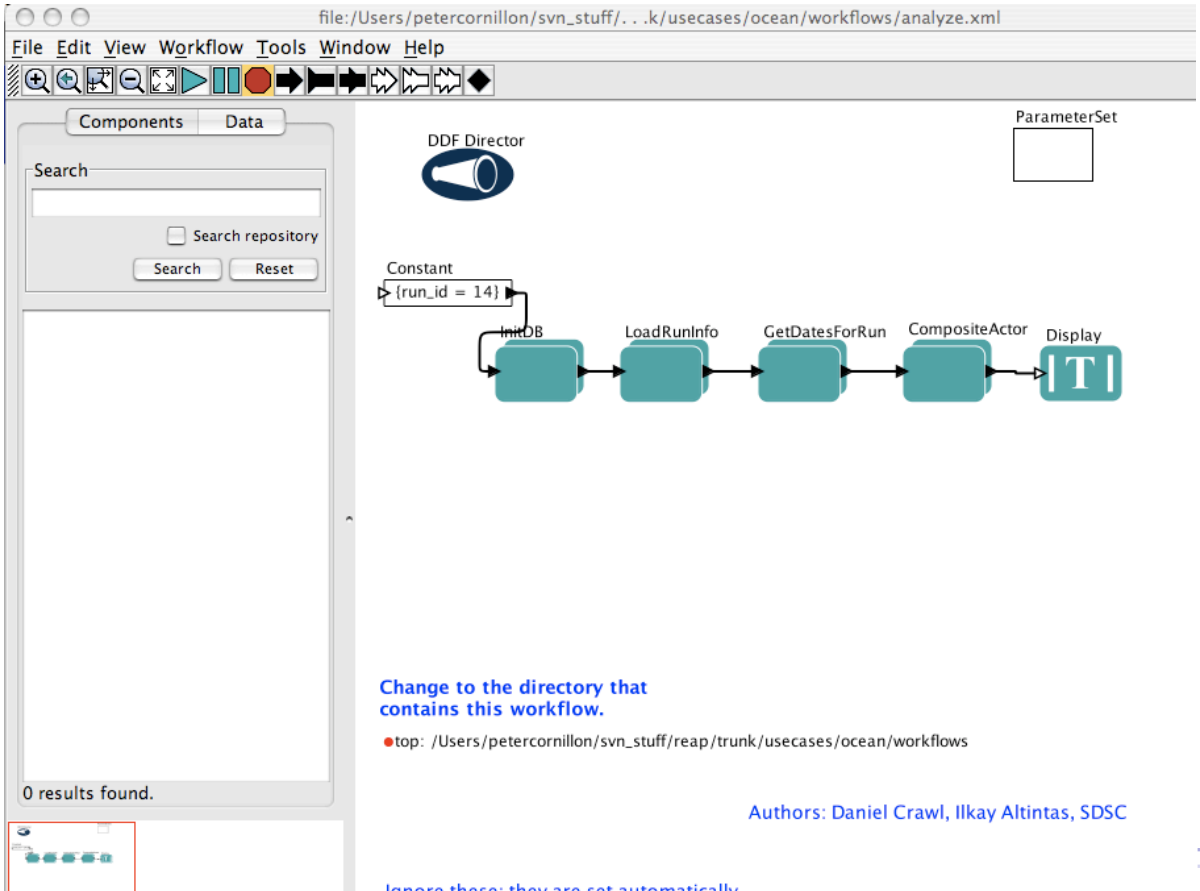
Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU.
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.

Progress to Date

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU.
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.



Progress to Date

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU.
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.

Progress to Date

Peter Cornillon[1], Daniel Crawl[2] and Ilkay Altintas[2]

- We have built basic workflows for the
  - Match-up data base construction. This work flow will:
    - Access randomly located tiles at randomly located times for HYCOM from FSU.
    - Access the corresponding tiles for Pathfinder from PO-DAAC (JPL),
    - Store the results in a database in San Diego.
  - A rudimentary analysis.
    - Retrieve the matchup data from San Diego, and
    - Perform some rudimentary stats on the data.