



# Environmental Niche Modeling in Kepler

Chad Berkley for Dan Higgins  
National Center for Ecological Analysis and Synthesis  
UC Santa Barbara

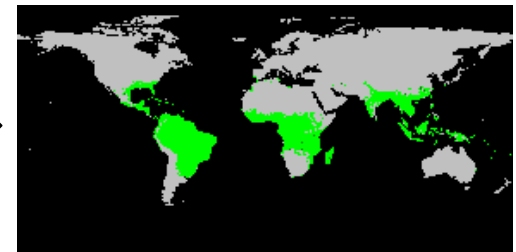
<http://www.kepler-project.org>

<http://seek.ecoinformatics.org>





- [illegible]



# GARP

- Stands for “Genetic Algorithm for Ruleset Production”
- Algorithm created by David Stockwell
- Uses genetic algorithm to create “rules” based on input data
- Statistics are then used to find the probability of each rule
- Used for ENM to predict species locations based on input data
  - Only one of many genetic algorithms that could be used for ENM

# GARP Modeling Workflows

jar:file:/C:/kepler-1.0.0alpha8/build...  
File Help

**Ready-to-run Workflows**

| Workflow   | Description   |
|--|---|
| <a href="#">Ptolemy Introduction</a>               | A link to the Ptolemy Intro                         |
| <a href="#">ENM (GARP) Workflows</a>               | A link to a collection of ENM                       |
| <a href="#">Using R in Kepler</a>                  | A link to a collection of w                         |
| <a href="#">EML2 Simple Plot Example</a>           | A workflow to test the EN                           |
| <a href="#">Promotor Identification Workflow</a>   | A workflow that tests a m web service query actor a |
| <a href="#">GEON Mineral Classifier</a>            | A workflow for modal cla                            |
| <a href="#">GEON Map Workflow</a>                  | GEON geology map integri                            |
| <a href="#">Discrete Logistic</a>                  | Single species Discrete Lo                          |
| <a href="#">Lotka-Volterra Predator Prey Model</a> | LV Predator Prey Model                              |
| <a href="#">Lotka-Volterra Predator</a>            |   |

jar:file:/C:/kepler-1.0.0alpha8/build... Jemy/configs/kepler/ENM\_Workflows.html  
File View Help

## Ecological Niche Modeling (GARP) Workflows

A number of Ecological Niche Modeling (GARP) workflows have been created using Kepler. Most are just pieces of a final, complete workflow.. Brief descriptions of these 'samples' are given below

|  |  |
|--|--|
| <a href="#">1. Baseline 3-Actor GARP - Browser Display</a>                 | This is the basic starting point for many of the ENM workflows. It consists of Desktop GARP code wrapped into 3 actors with inputs from static (pre-defined) files that are already in the exact formats required by GARP (e.g. layers in *.raw format). The final occurrence distribution map is displayed using a browser.   |
| <a href="#">2. Baseline 3-Actor GARP - ImageJ Display</a>                  | This workflow is the same as the previous one except the actor for displaying the resulting map is now the ImageJ actor, based on the ImageJ image processing system developed at NIH. This actor allows the user to carry out a variety of manipulations on the resulting image.  |
| <a href="#">3. GARP with Occurrence Data from Digir/Ecogrid</a>            | This workflow gets the GARP occurrence data from a DarwinCore data search through the Ecogrid. It illustrates how Kepler/SEEK technology can supply dynamic data sources.  |
| <a href="#">4. GARP with Occurrence Data and Layer Integration</a>         | This example extends the previous model by adding the dynamic integration of a set of geographic layers that are in *.ASC grid format. In this case, the ASC grids all have the same extent and grid sizes. These layers are feed through an actor which converts them to the raw files needed for GARP and automatically creates the *.dxi file which summarized the layers for GARP. |
| <a href="#">5. GARP with Occurrence Data and Ecogrid Layer Integration</a> | This example obtains occurrence data and one layer of ASC spatial data from the Ecogrid. The spatial data layer is regrided and integrated with other ASC layers to create the *.RAW and *.dxi files needed by GARP. This example thus illustrates how one can add to spatial data with new source. (Note: spatial data is low resolution to minimize download and run times.)         |
| <a href="#">6. GDAL Input and ReProjection</a>                             | This workflow shows an example of using the GDAL system for input of specialized GIS format files and reprojection of spatial data into desired forms.   |
| <a href="#">7. GARP Omission Commission Example</a>                        | The example calculates omission and commission values from the output of a GARP prediction map and a 'test set' of occurrence longitudes and latitudes. Results are written to a local file as a table.  |



# Conceptual GARP Model



## Acquire Data

DiGIR Species  
Location List

Environmental GIS  
Layers

Hydro1K Precip GIS  
Layers

Climate Change GIS  
Layers

## Spatial Data Manipulation

Spatial Integration

Convex Hull Creation

## GARP

Presampling

Actual GARP  
execution

Prediction

## Statistical Sampling

Omission/Commission  
Calculations

Statistically choose  
the best subsets



# Tools for data acquisition

- Ecogrid
  - DiGIR - Museum collection data
  - SRB - large GIS layers
- Local files/Kepler file reader



# Longitude, Latitude Table from Digir Ecogrid Search

Unnamed

File View Edit Graph Debug Help

Components Data

SDF Director

Search

mephitis

Search Reset Cancel Sources

mephitis

Display

..Display

File Help

|          |         |
|----------|---------|
| -91.9053 | 30.0869 |
| -91.6158 | 30.7606 |
| -92.1935 | 30.3857 |
| -91.1841 | 30.4045 |
| -92.0342 | 30.273  |
| -92.2794 | 30.8    |
| -92.9773 | 30.247  |
| -92.6567 | 30.2221 |
| -92.0195 | 30.3977 |

disphragis mephitis  
luteimonas mephitis  
mephitis  
mephitis macroura  
mephitis macroura macroura  
mephitis macroura milleri  
mephitis mephitis  
mephitis mephitis avia  
mephitis mephitis elongata  
mephitis mephitis estor  
mephitis mephitis hudsonica  
mephitis mephitis holzneri  
mephitis mephitis holzneri  
mephitis mephitis hudsonica  
mephitis mephitis hudsonica ?  
mephitis mephitis hudsonicus  
mephitis mephitis major  
mephitis mephitis mephitis



# Spatial Manipulation Tools in Kepler

- GIS tools
  - GDAL
  - GRASS
- Java tools
  - Specialized actors
  - ImageJ



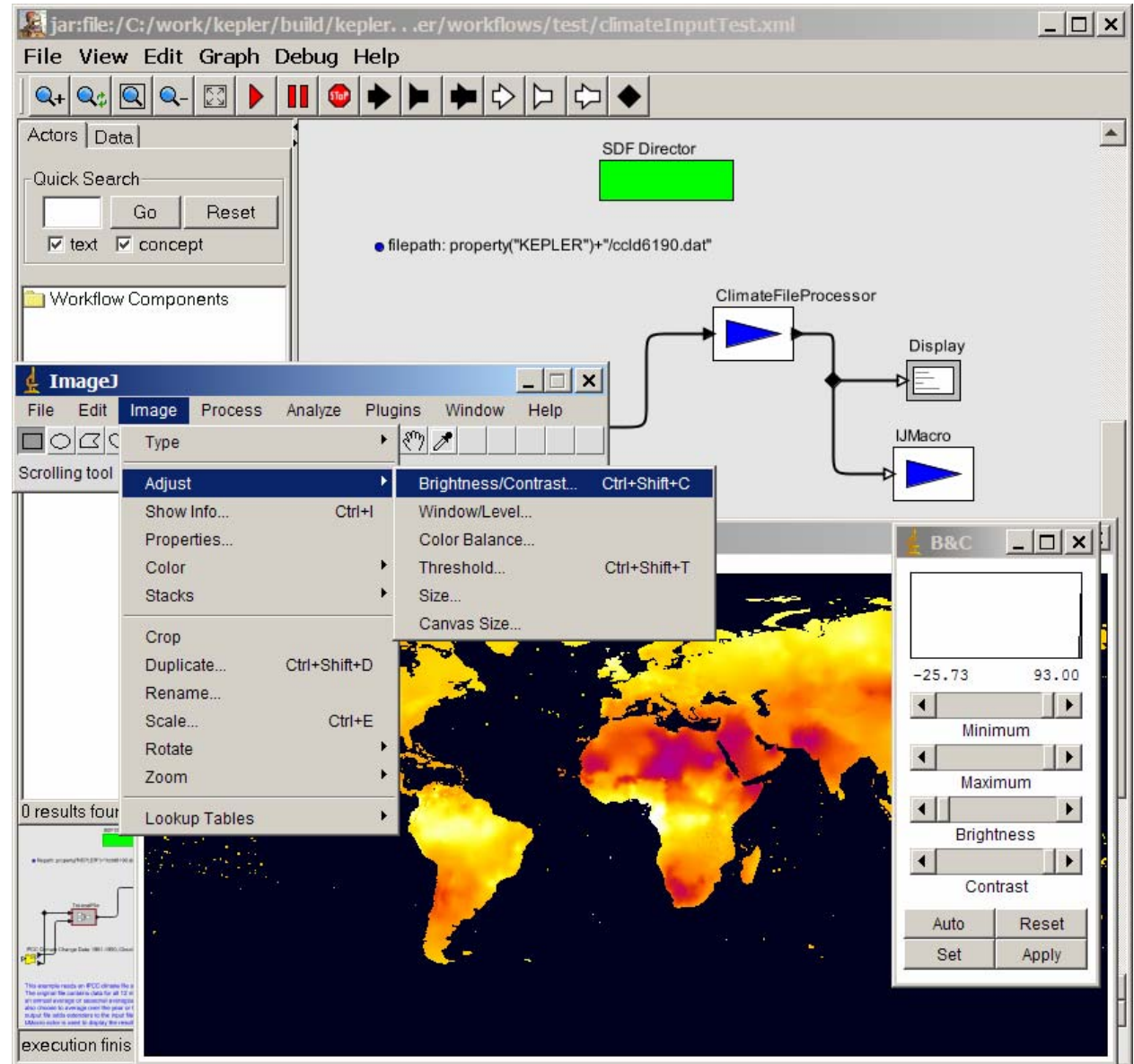


# Java Actors for Handling ASCII Grid Files

Java-based actors created to read and manipulate ASCII grid files

ImageJ image processing package from NIH added as an actor to display and manipulate images

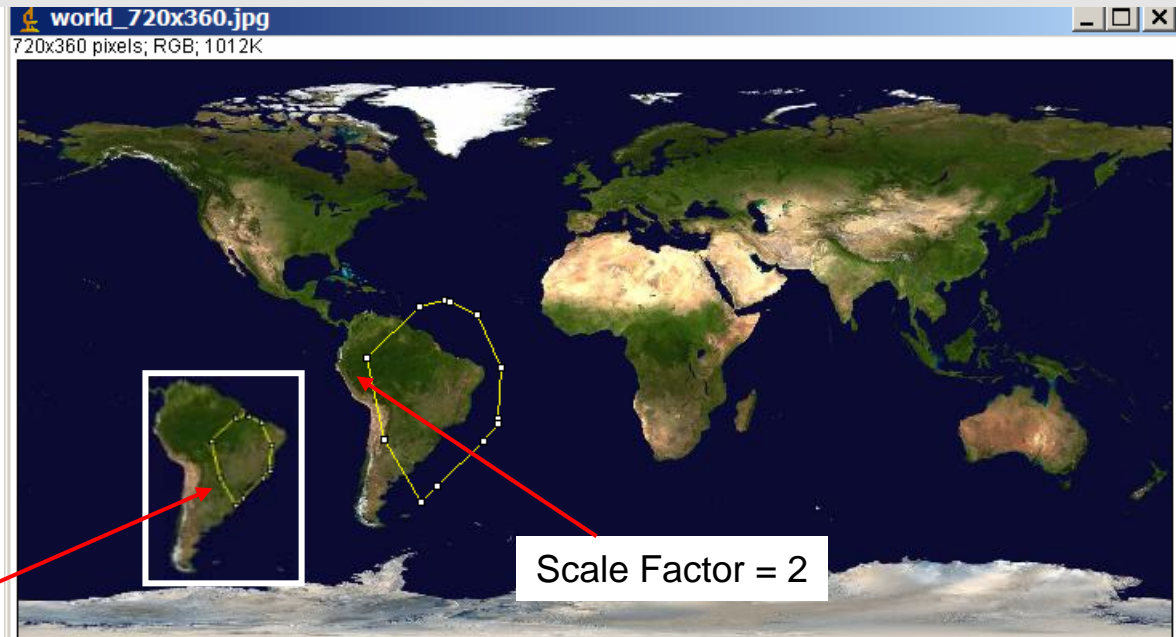
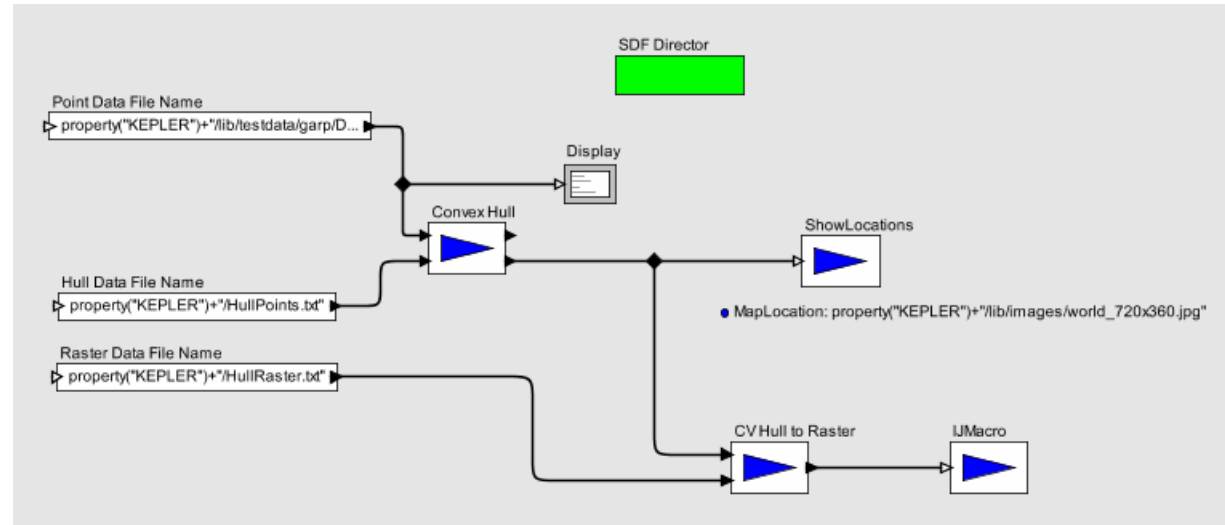
ImageJ has macro capabilities & numerous plug-ins



ImageJ - <http://rsb.info.nih.gov/ij/>

# Java Routines for Convex Hull Calculation and Rasterization

Java routines for the Convex Hull convert the polygon to a shape which can be 'scaled' in size





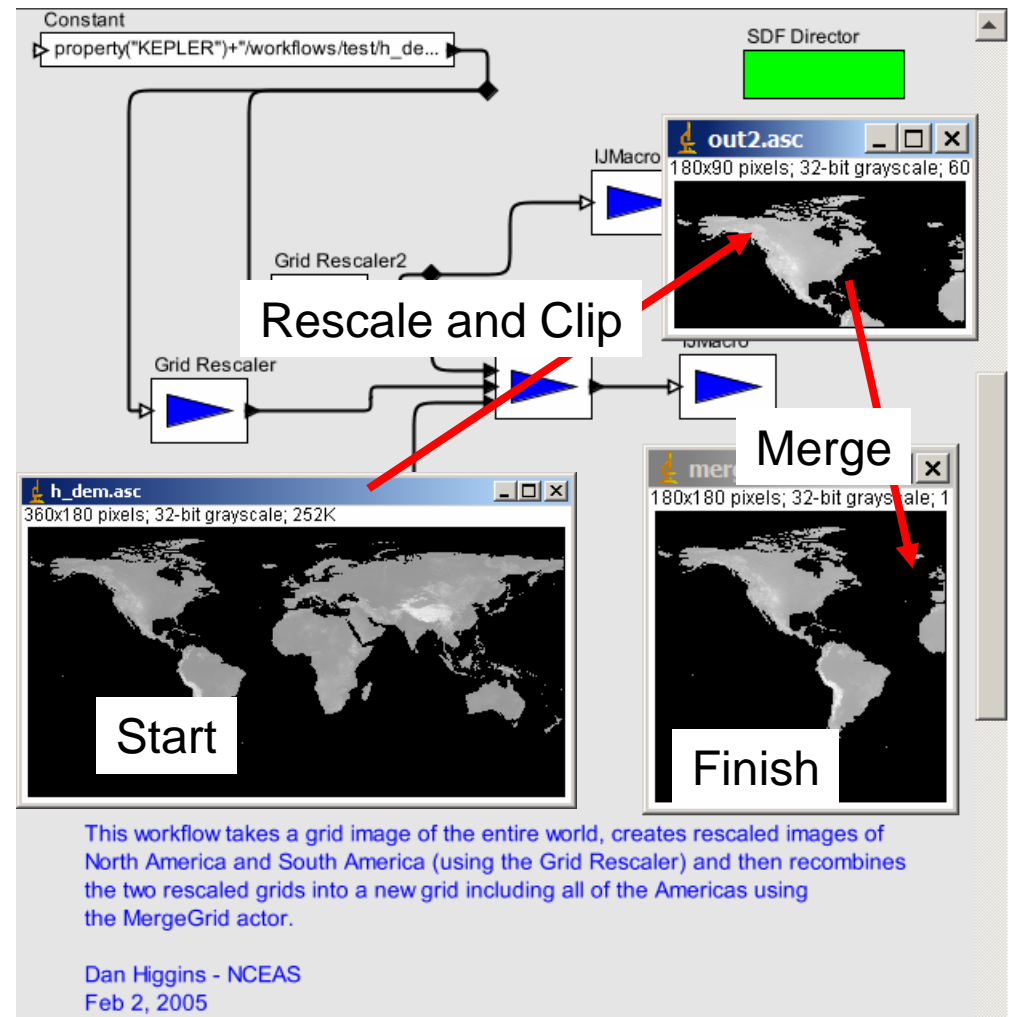
# Java Actors for Rescaling and Merging

Java actors can rescale  
ASCII grid files, changing  
resolution and extent

Both nearest-neighbor and  
Inverse-Distance-  
Weighted algorithms  
implemented

Disk-based code allows very  
large grids to be  
manipulated (i.e not  
limited by RAM)

Grid rasters can be  
'merged' with various  
operations on data in  
overlapping pixels





# GDAL - Geospatial Data Abstraction Library

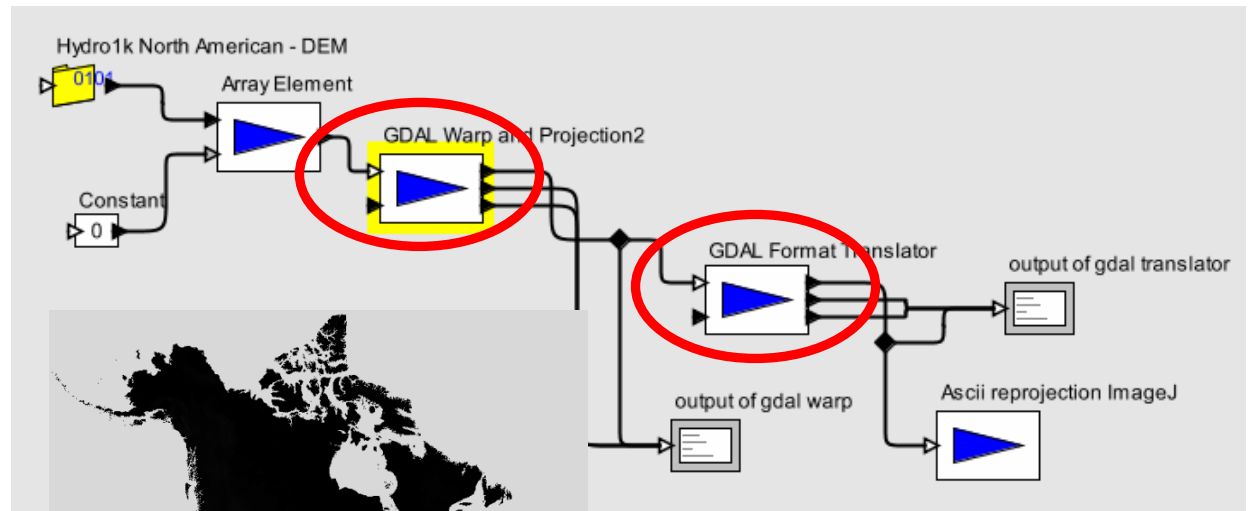
<http://www.remotesensing.org/gdal/>

GDAL translator  
library connected to  
Kepler actors

GDAL can input ~40  
different raster formats

Conversion between  
different projections  
possible

File format conversions  
also possible

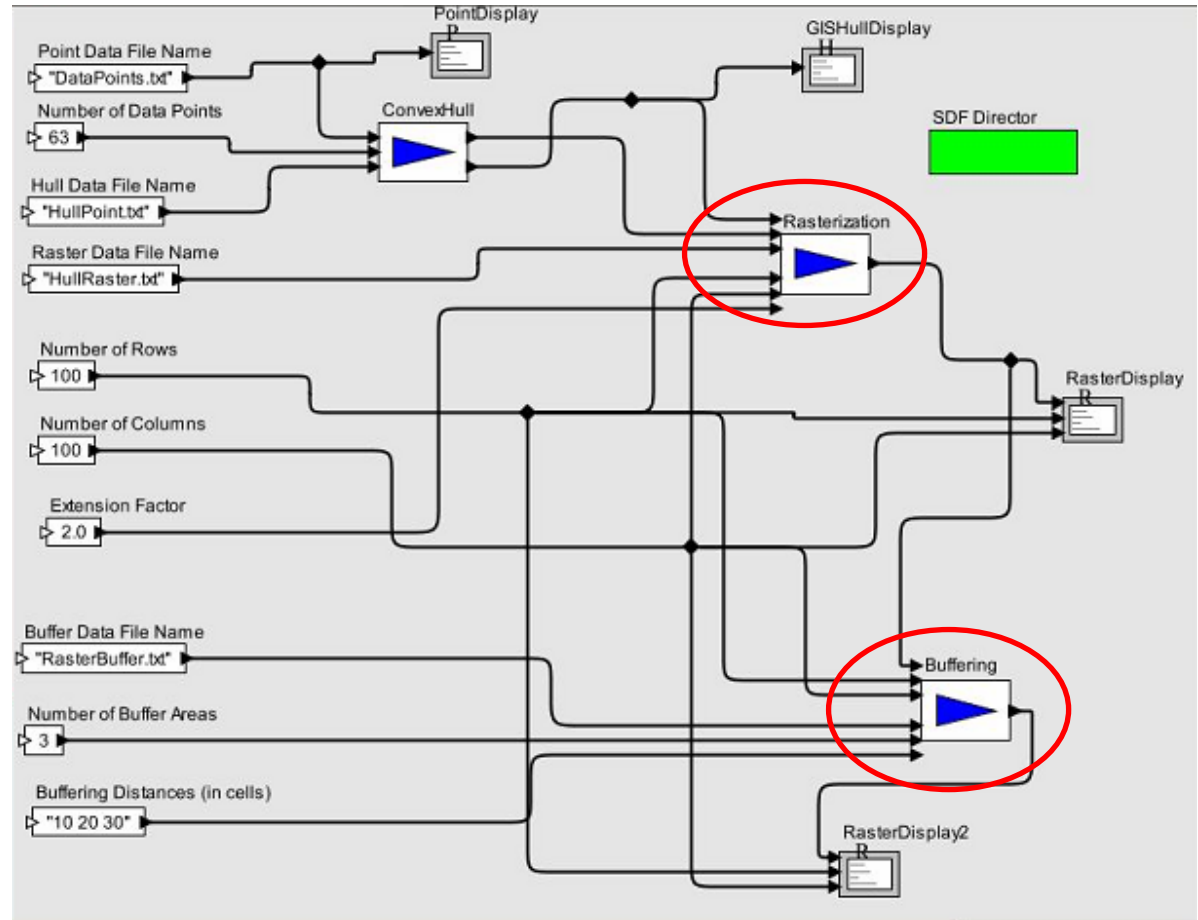




# GRASS - Geographic Resources Analysis Support System (<http://grass.itc.it/>)

## Functionality

- Polygon  
Rasterization
- Raster  
Buffering





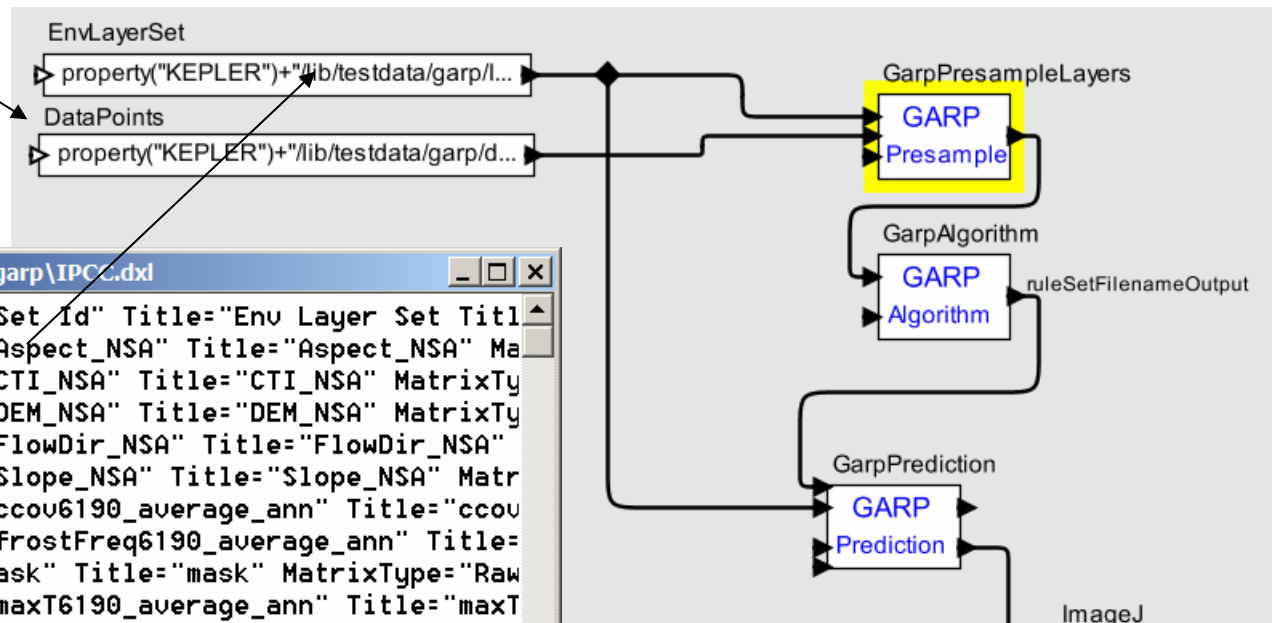
# GARP Actors in Kepler

- GARP Presample
  - Presamples data points based on preset sampling criteria
- GARP Algorithm
  - Runs the actual GARP algorithm
- GARP Prediction
  - Predicts spatial points based on the output “rules” of GARP Algorithm



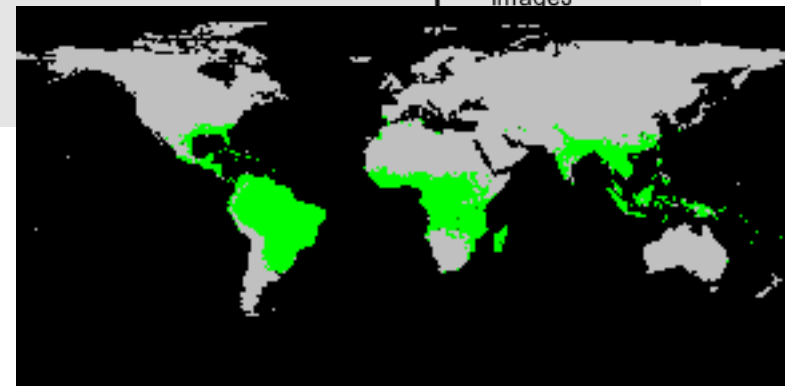
# Single GARP Calculation Workflow

Species locations  
(longitude, latitude)



```

C:\kepler-1.0.0alpha8\lib\testdata\garp\IPCC.dxl
<EnvLayerSet Id="Env Layer Set Id" Title="Env Layer Set Titl
<EnvLayer Type="Layer" Id="Aspect_NSA" Title="Aspect_NSA" Ma
<EnvLayer Type="Layer" Id="CTI_NSA" Title="CTI_NSA" MatrixTy
<EnvLayer Type="Layer" Id="DEM_NSA" Title="DEM_NSA" MatrixTy
<EnvLayer Type="Layer" Id="FlowDir_NSA" Title="FlowDir_NSA"
<EnvLayer Type="Layer" Id="Slope_NSA" Title="Slope_NSA" Matr
<EnvLayer Type="Layer" Id="ccov6190_average_ann" Title="ccov
<EnvLayer Type="Layer" Id="frostFreq6190_average_ann" Title=
<EnvLayer Type="Mask" Id="mask" Title="mask" MatrixType="Raw
<EnvLayer Type="Layer" Id="maxT6190_average_ann" Title="maxT
<EnvLayer Type="Layer" Id="minT6190_average_ann" Title="minT
<EnvLayer Type="Layer" Id="precip6190_average_ann" Title="pr
<EnvLayer Type="Layer" Id="radiance6190_average_ann" Title="
<EnvLayer Type="Layer" Id="trange6190_average_ann" Title="tr
<EnvLayer Type="Layer" Id="vaporpress6190_average_ann" Title
<EnvLayer Type="Layer" Id="wetdayfreq6190_average_ann" Title
<EnvLayer Type="Layer" Id="wind6190_average_ann" Title="wind
  
```





# Statistical Post-processing

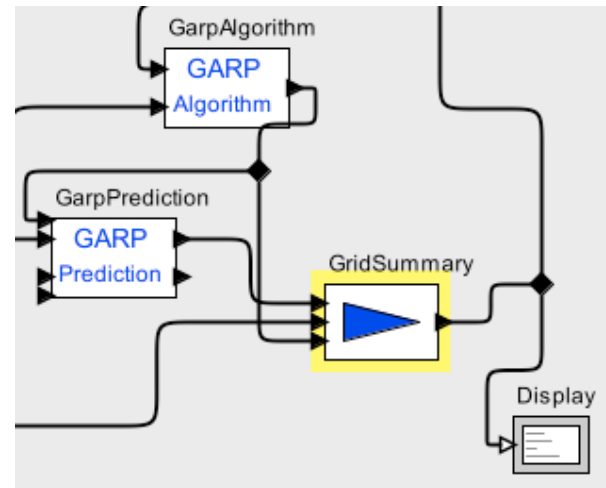
- Calculate omission/commission errors
- Select best subsets based on omission/commission





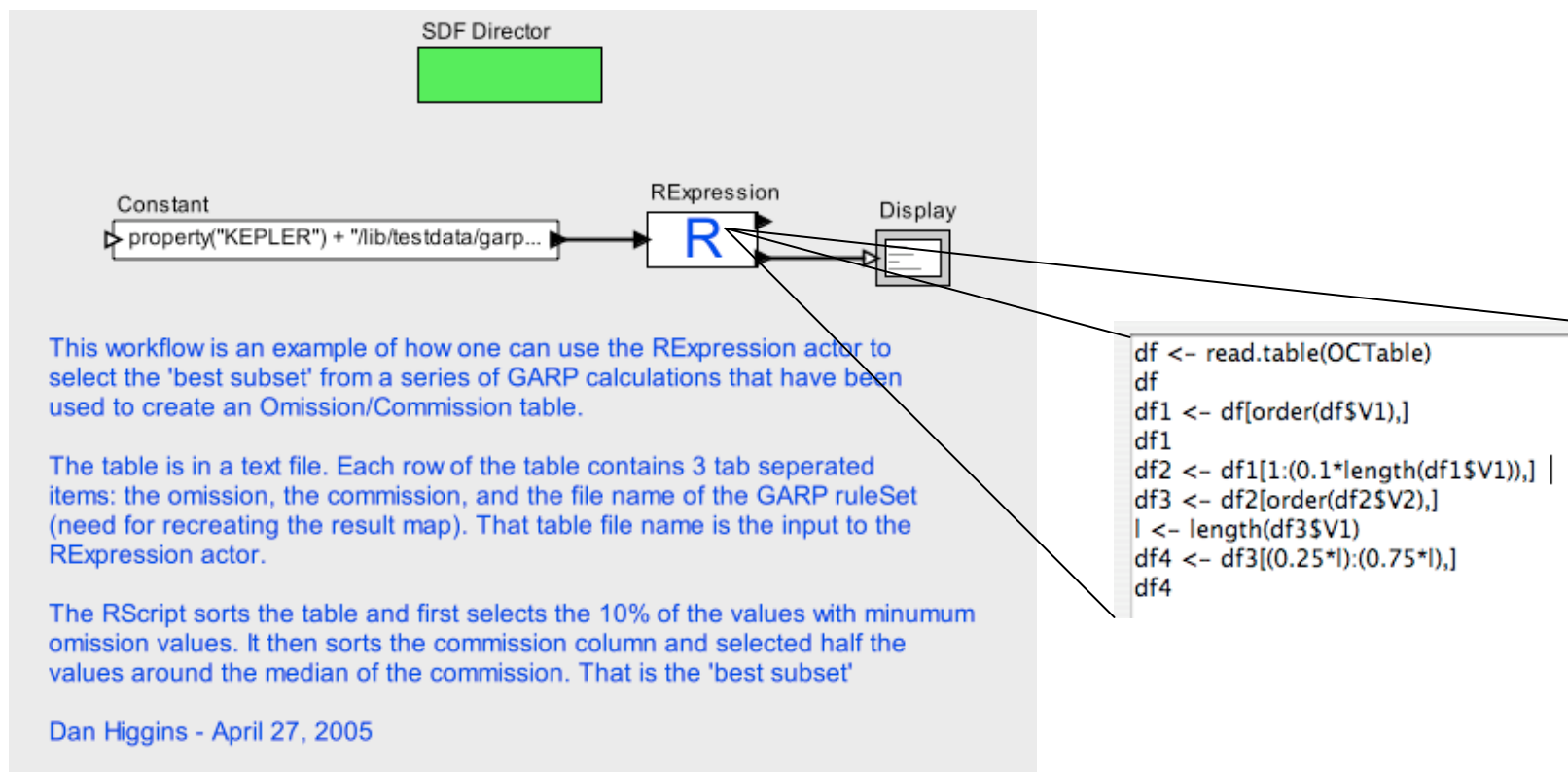
# Omission/Commission Calculation

- GridSummary actor can calculate omission/commission errors from GARP output
- GridSummary is a specialized actor written in Java



# Creating a Best Subset

- R can be used for best subset selection
- Choose the GARP predictions with the lowest omission and commission rates





# Barriers to Implementation

- Diverse input file formats
- Agreement on statistical methods
- Understanding the “true” workflow behind ENM/GARP
- Complex data flow
- Prerequisite “massaging” of data
- We have currently implemented pieces of the entire ENM workflow, but not all the pieces will fit together