INTRODUCTION TO SPATIAL VERIFICATION

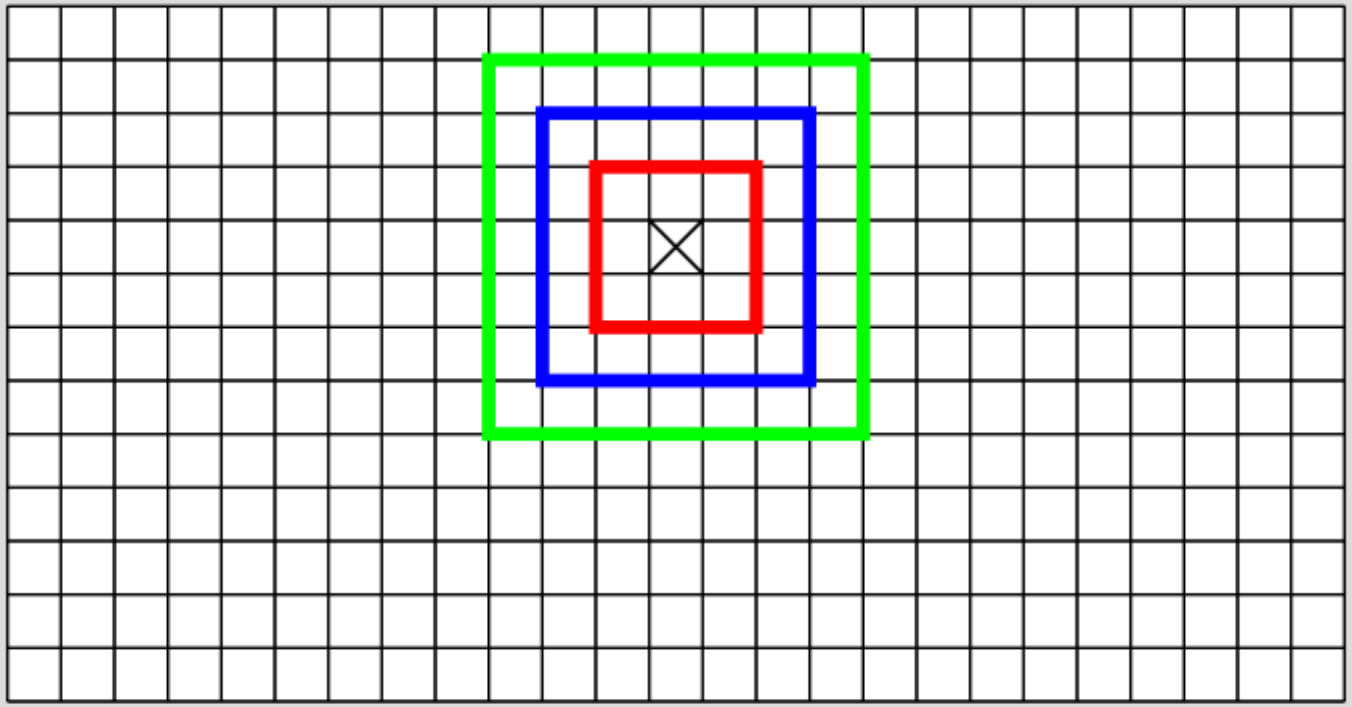
Randy Bullock

Categories of Spatial Verification

- Filtering
  - Neighborhood
  - Scale Separation
- Displacement
  - Field Deformation
  - Features

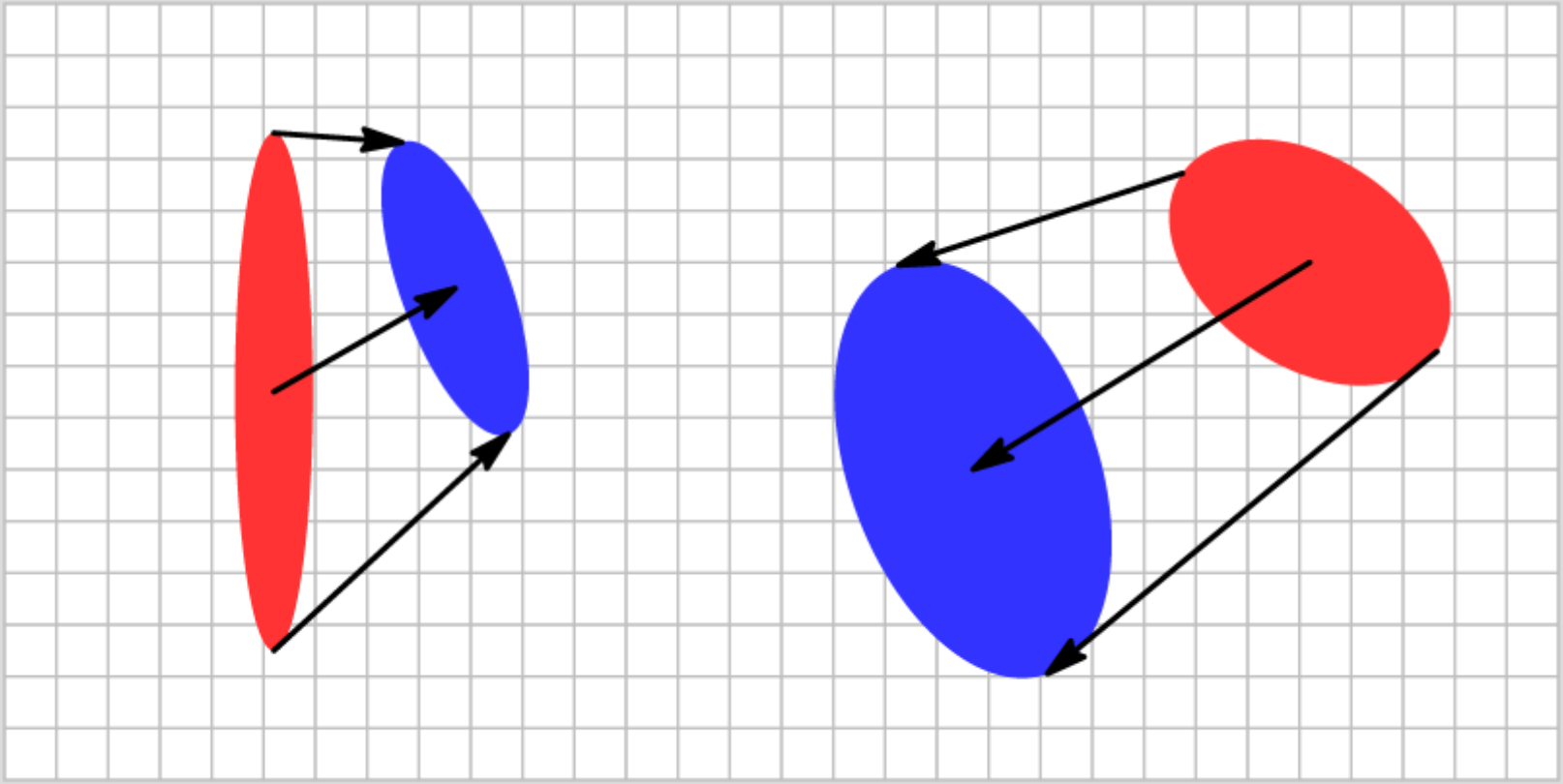
Neighborhood Methods

For each grid box, look in some neighborhood of that box, combine the values according to some rule, and assign the combined value to that grid box.



Possibilities: maximum, median, average.

Field Morphing



Small changes good, large changes bad.

Result is vector field defined on grid.

Field Deformation

Tries to verify large-scale features instead of grid points.

Examples:

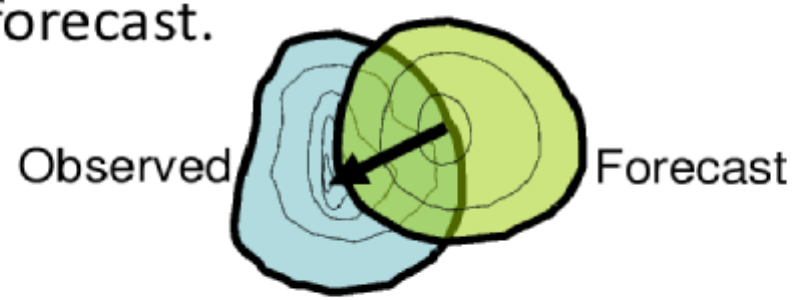
- Field Morphing
- Image Warping

  Result of these is vector field on grid
- Image Metrics

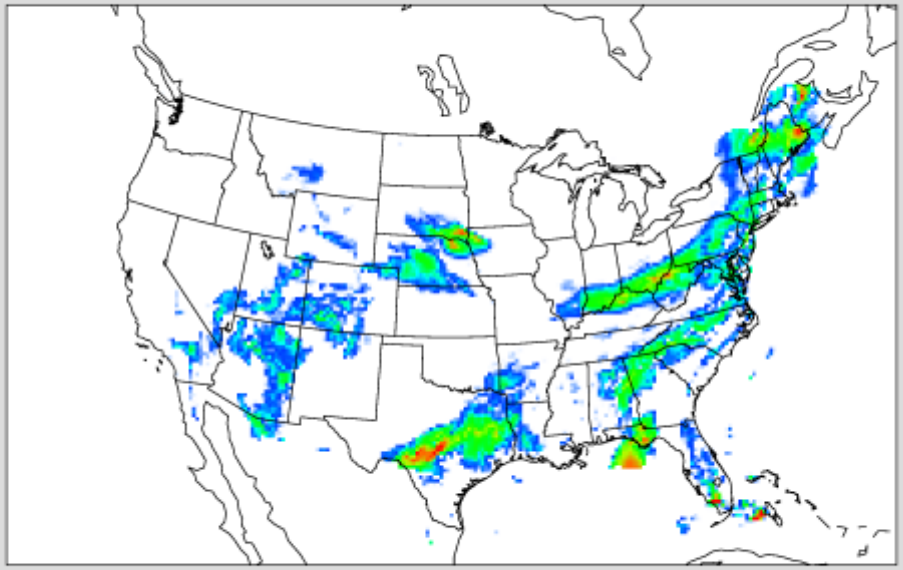
Feature-based approach (CRA)

Ebert and McBride, *J. Hydrol.*, 2000

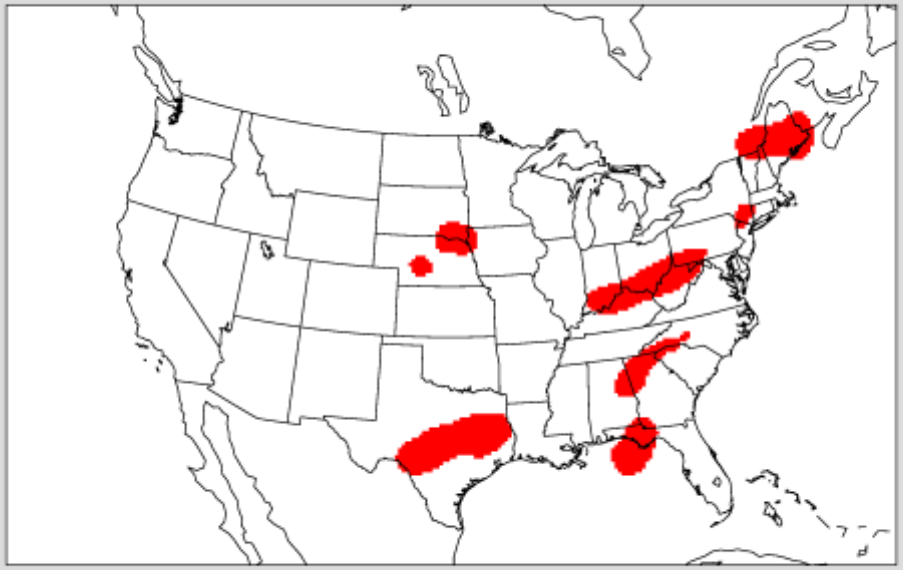
- Define entities using (user defined) threshold (Contiguous Rain Areas)
- Horizontally translate the forecast until a *pattern matching* criterion is met:
  - minimum total squared error between forecast and observations
  - maximum correlation
  - maximum overlap
- The displacement is the vector difference between the original and final locations of the forecast.



What are Objects?



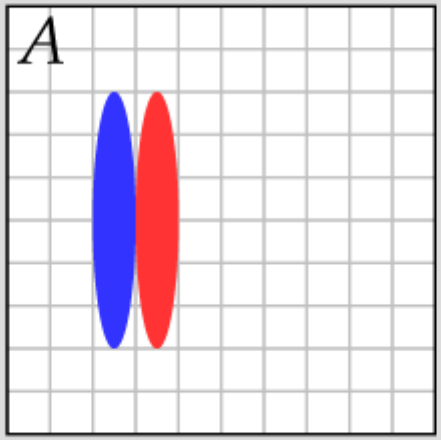
Raw Field

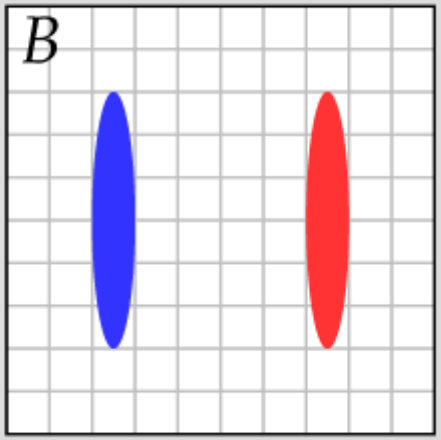


Object Field

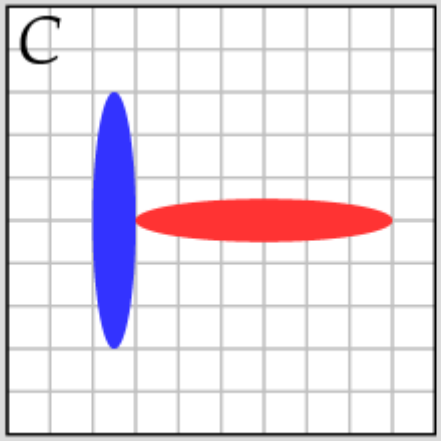
Objects are Regions of Interest

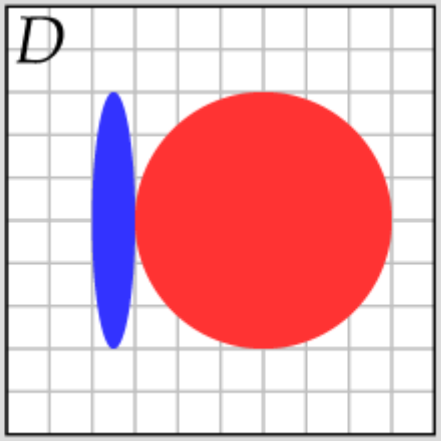
Background and Motivation

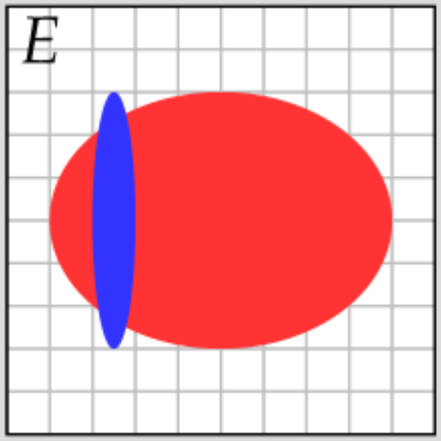




Score	A-D	E
Correlation Coefficient	0.0	0.2
Probability of Detection	0.0	0.9
False Alarm Ratio	1.0	0.9
Hanssen-Kuipers	0.0	0.7
Gilbert Skill Score	0.0	0.1



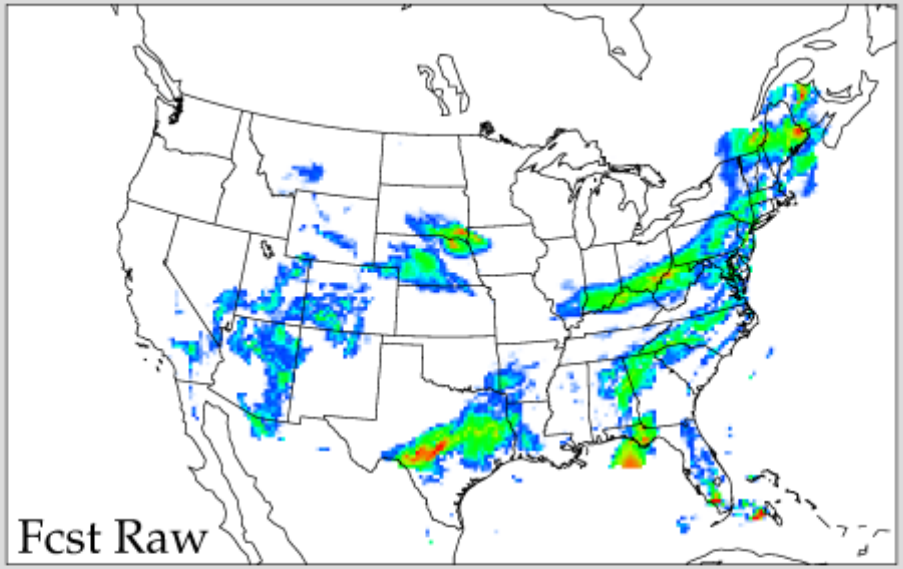


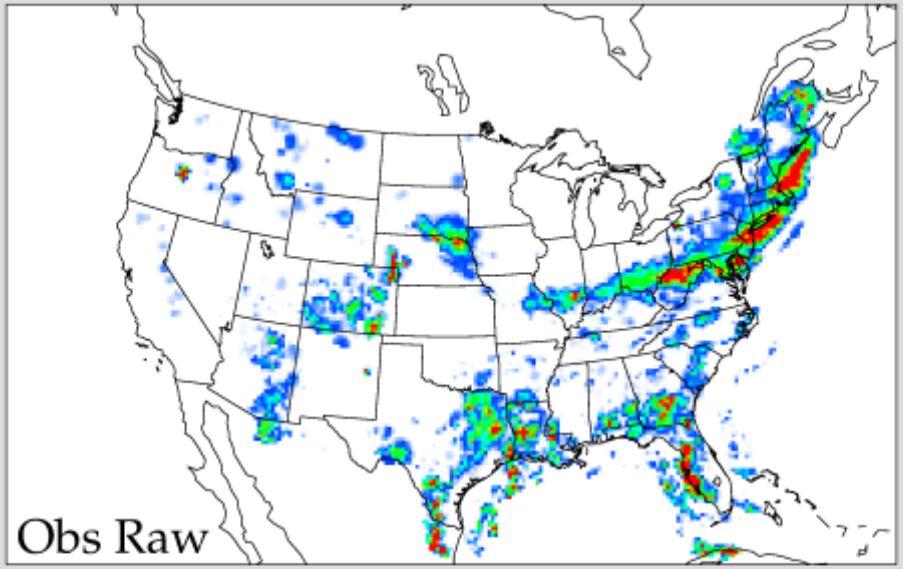


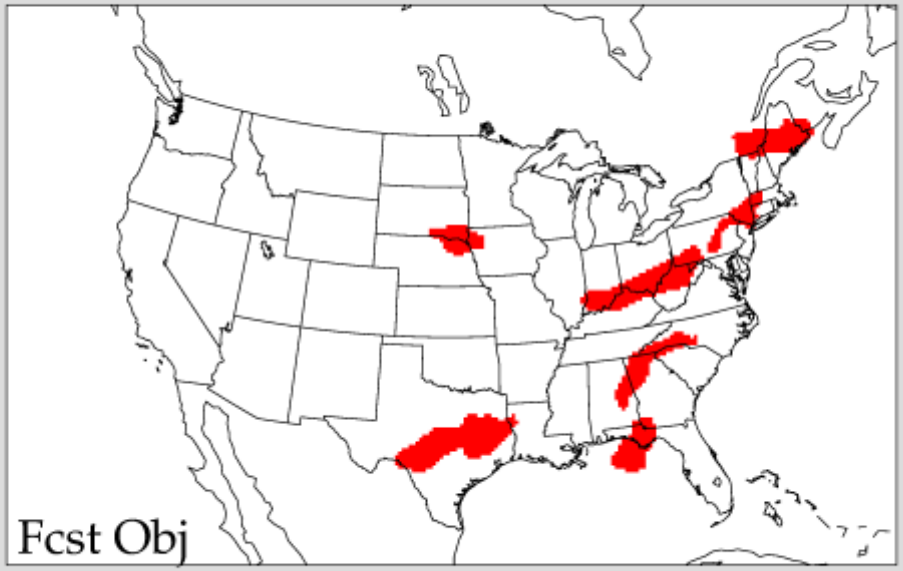
Forecast Observation

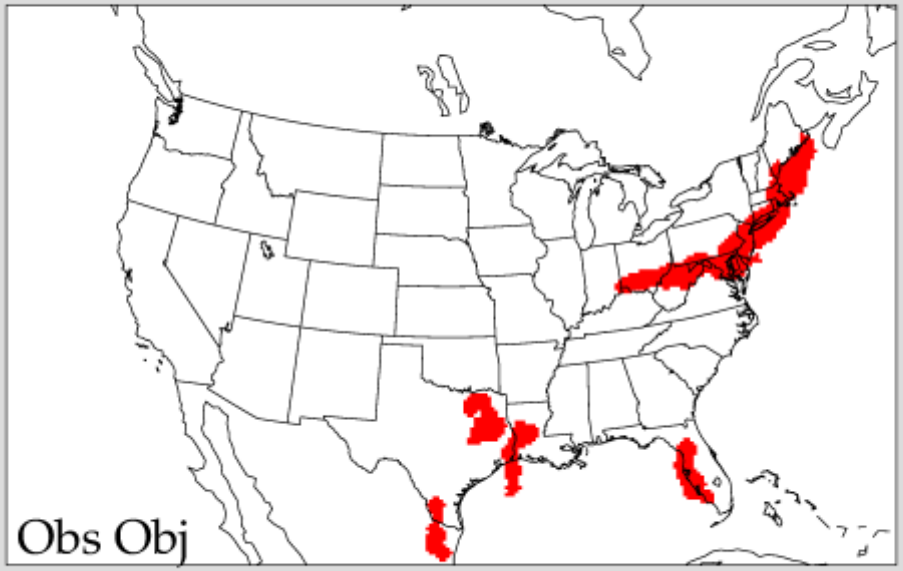
Which forecast is better?

MODE Example

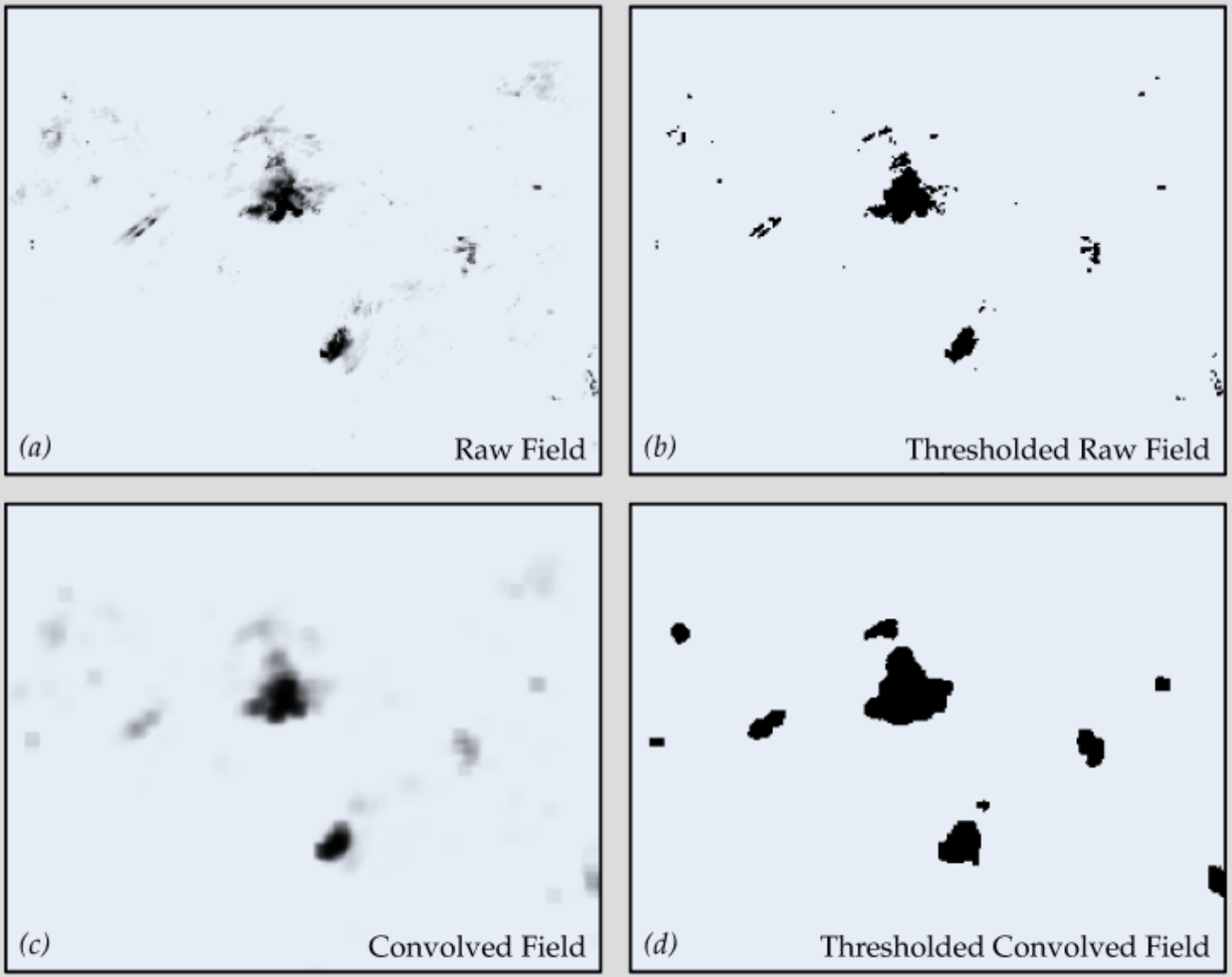




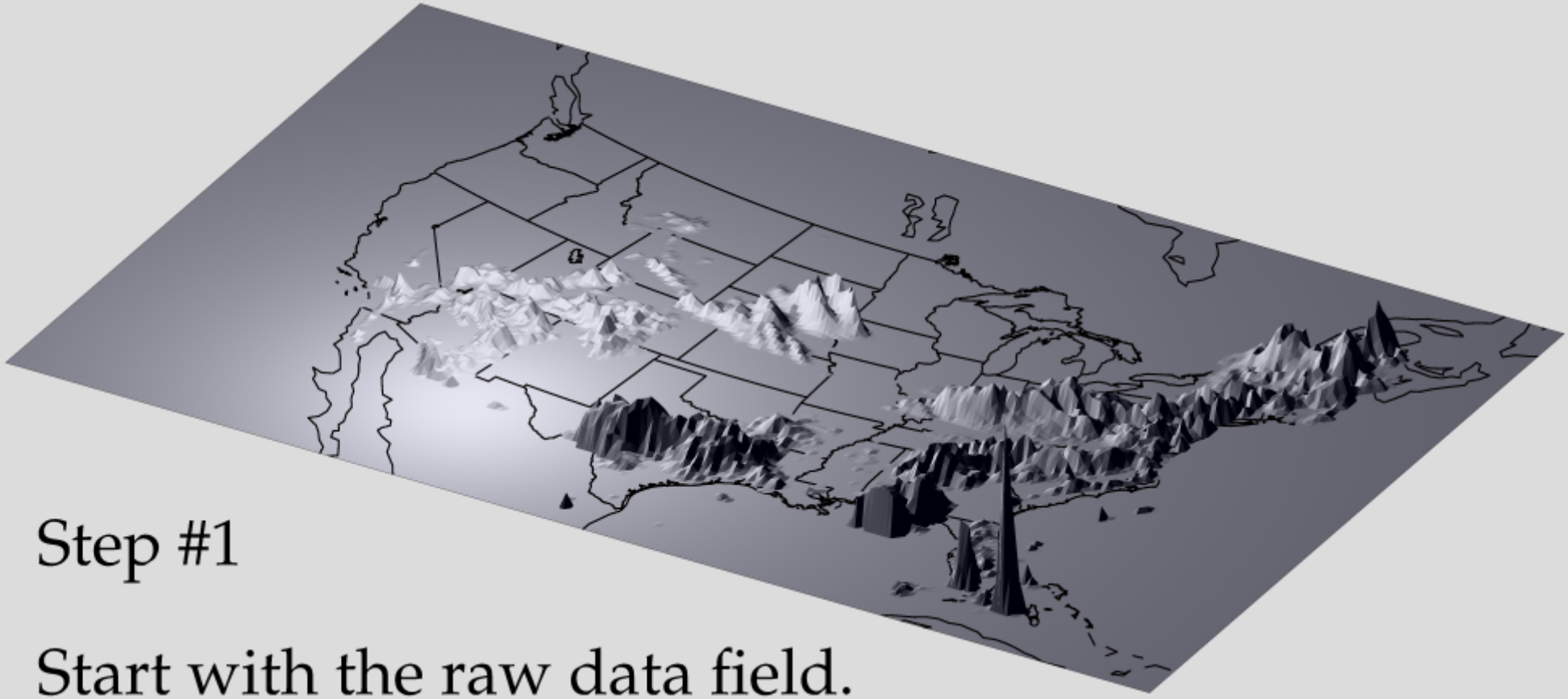




Just Threshold Raw Data?



Four Step Process for Resolving Objects

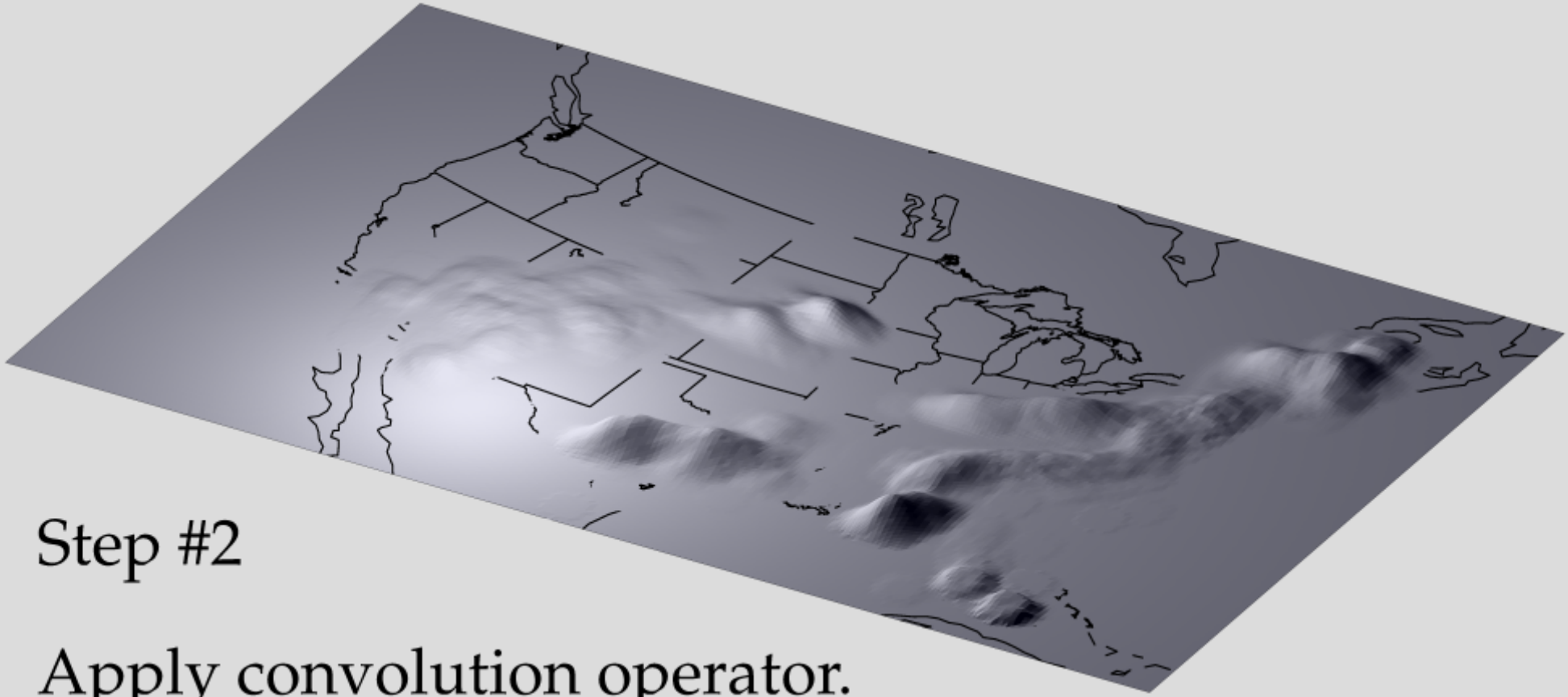


Step #1

Start with the raw data field.

In this case, a precipitation field.

Four Step Process for Resolving Objects

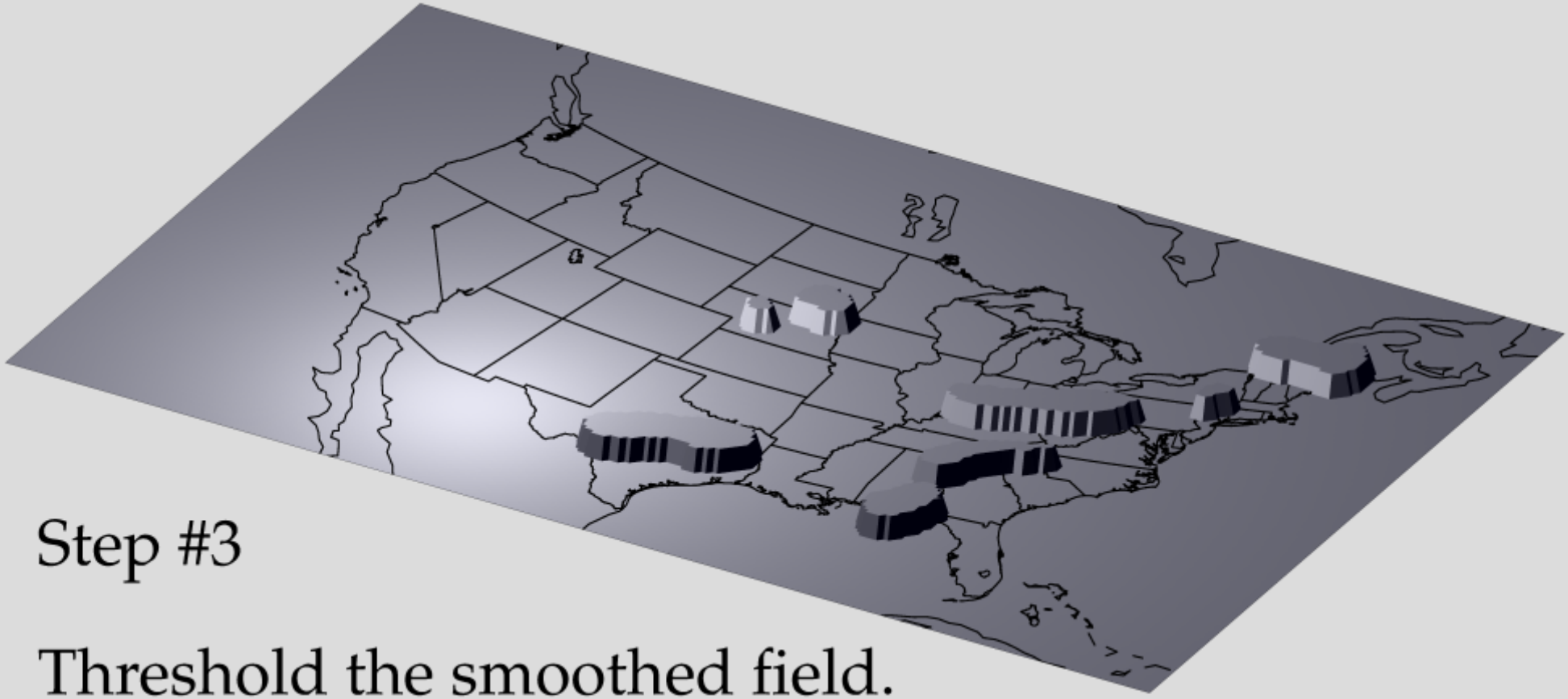


Step #2

Apply convolution operator.

This is basically a smoothing operation.

Four Step Process for Resolving Objects

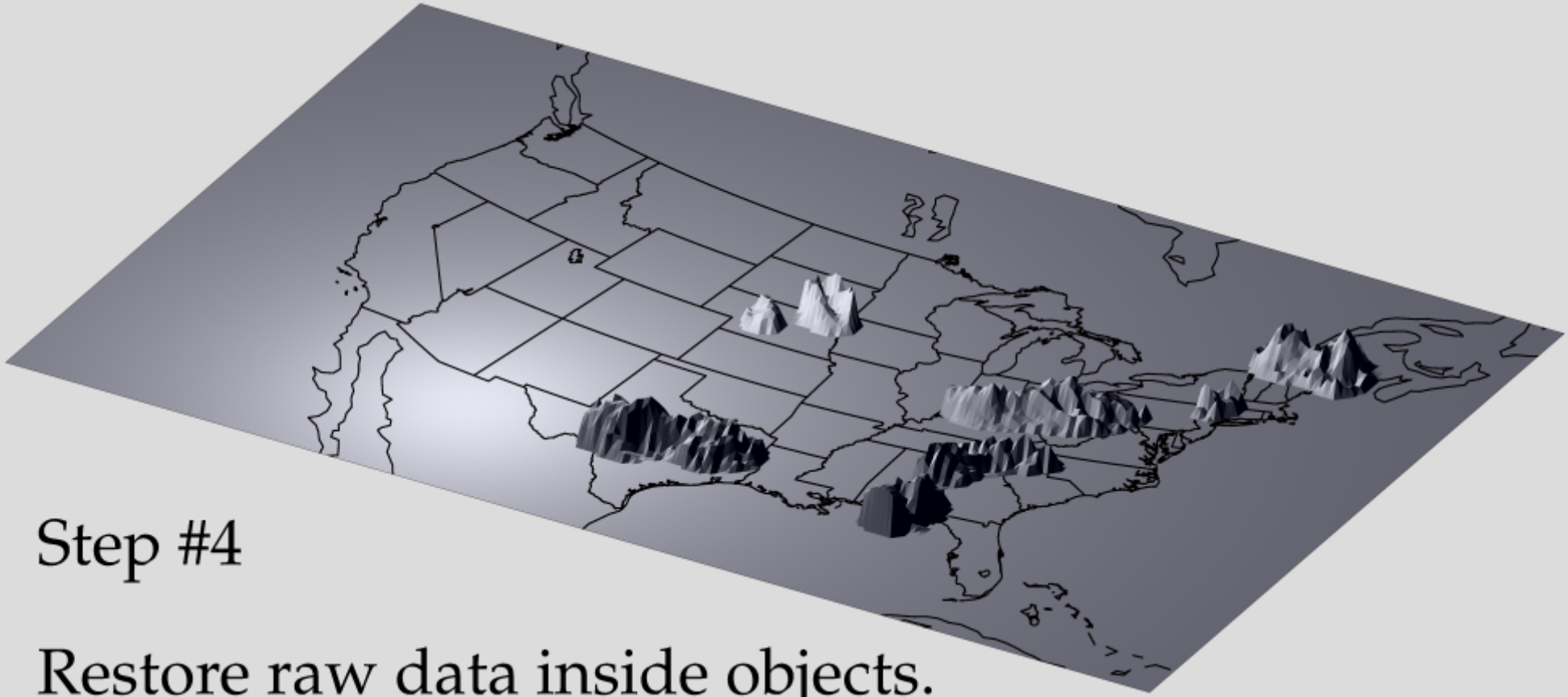


Step #3

Threshold the smoothed field.

This produces an on/off mask field.

Four Step Process for Resolving Objects

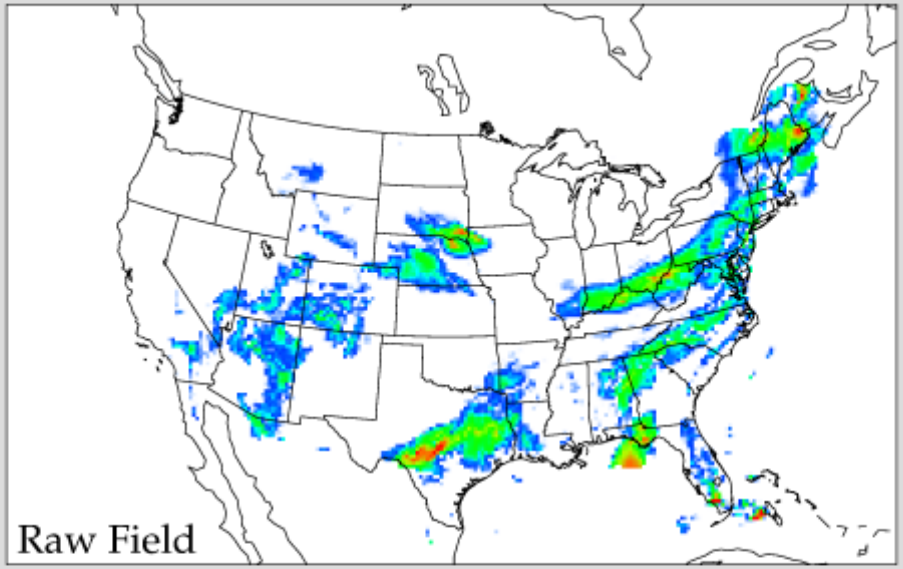


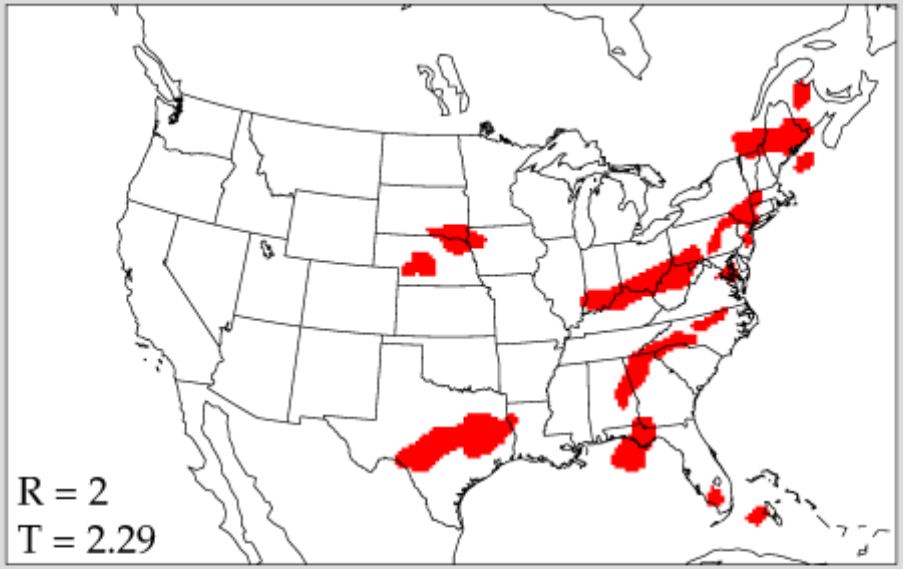
Step #4

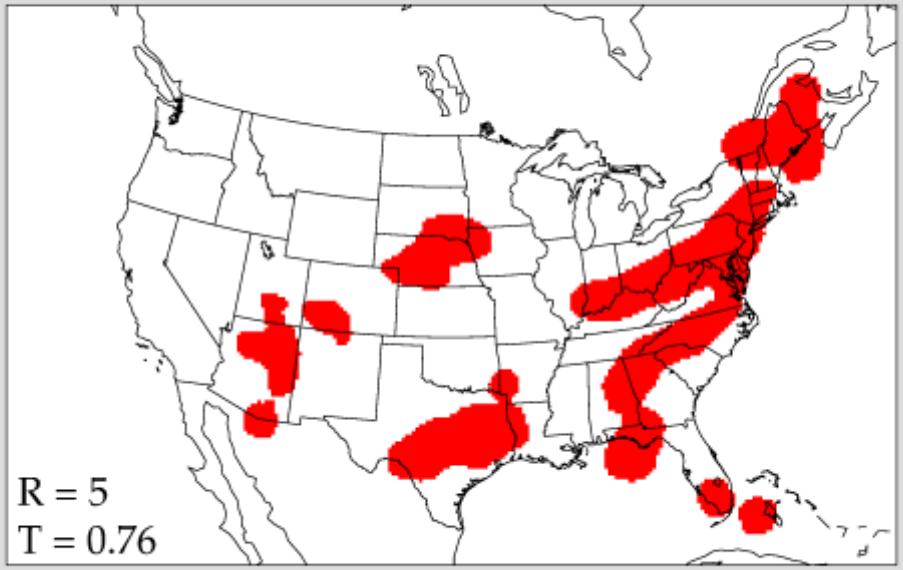
Restore raw data inside objects.

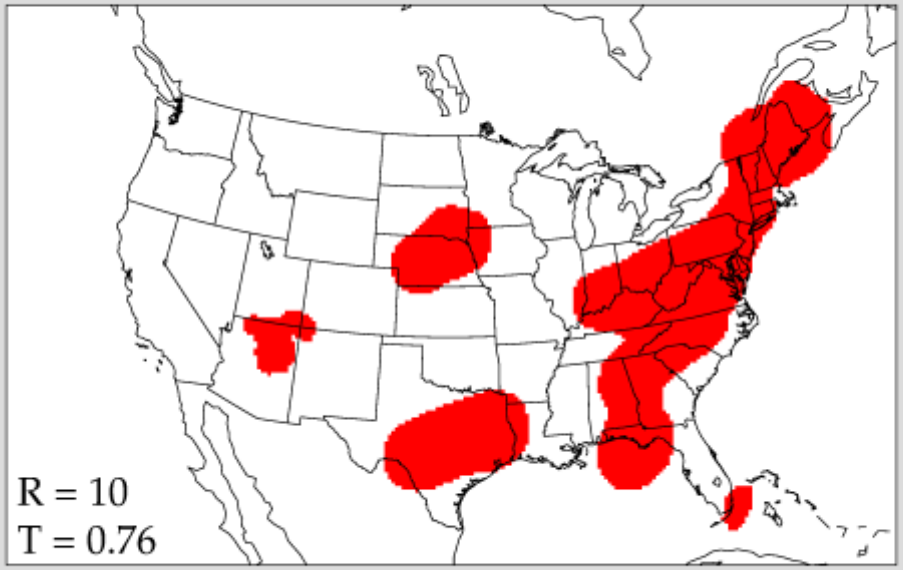
This gives us our objects for this field.

Note the Effect of Changing the Object-Definition Parameters

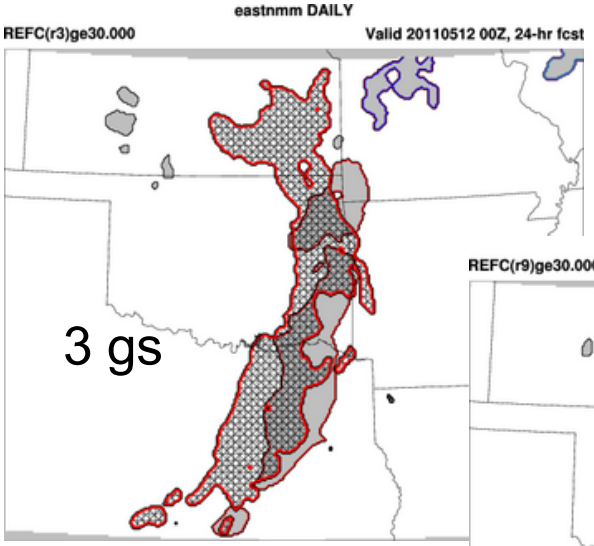






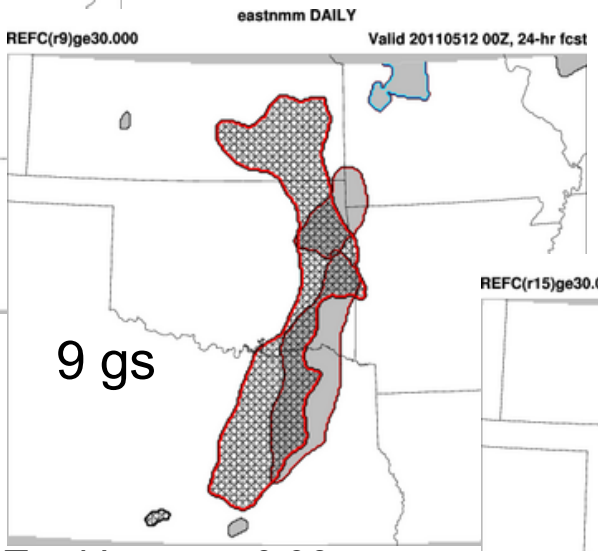


Example – REFC > 30 dBZ – Impact of smoothing radius

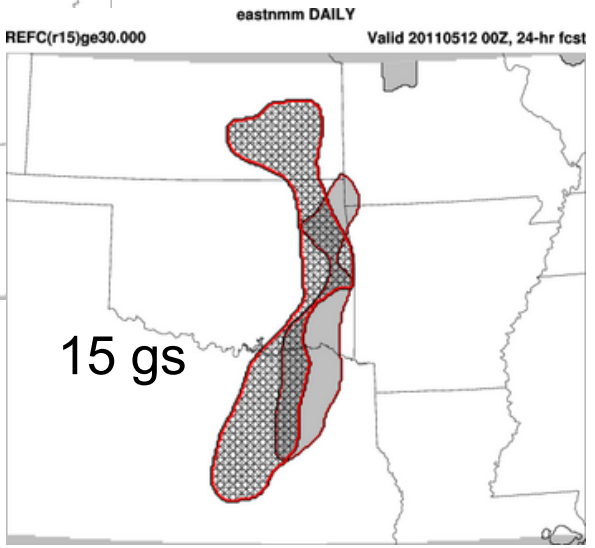


Total Interest: 0.96
Area Ratio: 0.57
Centroid Distance: 95km
P90 Intensity Ratio: 1.00

Convolution Radius Increases



Total Interest: 0.96
Area Ratio: 0.57
Centroid Distance: 94km
P90 Intensity Ratio: 1.02



Total Interest: 0.96
Area Ratio: 0.53
Centroid Distance: 92km
P90 Intensity Ratio: 1.04

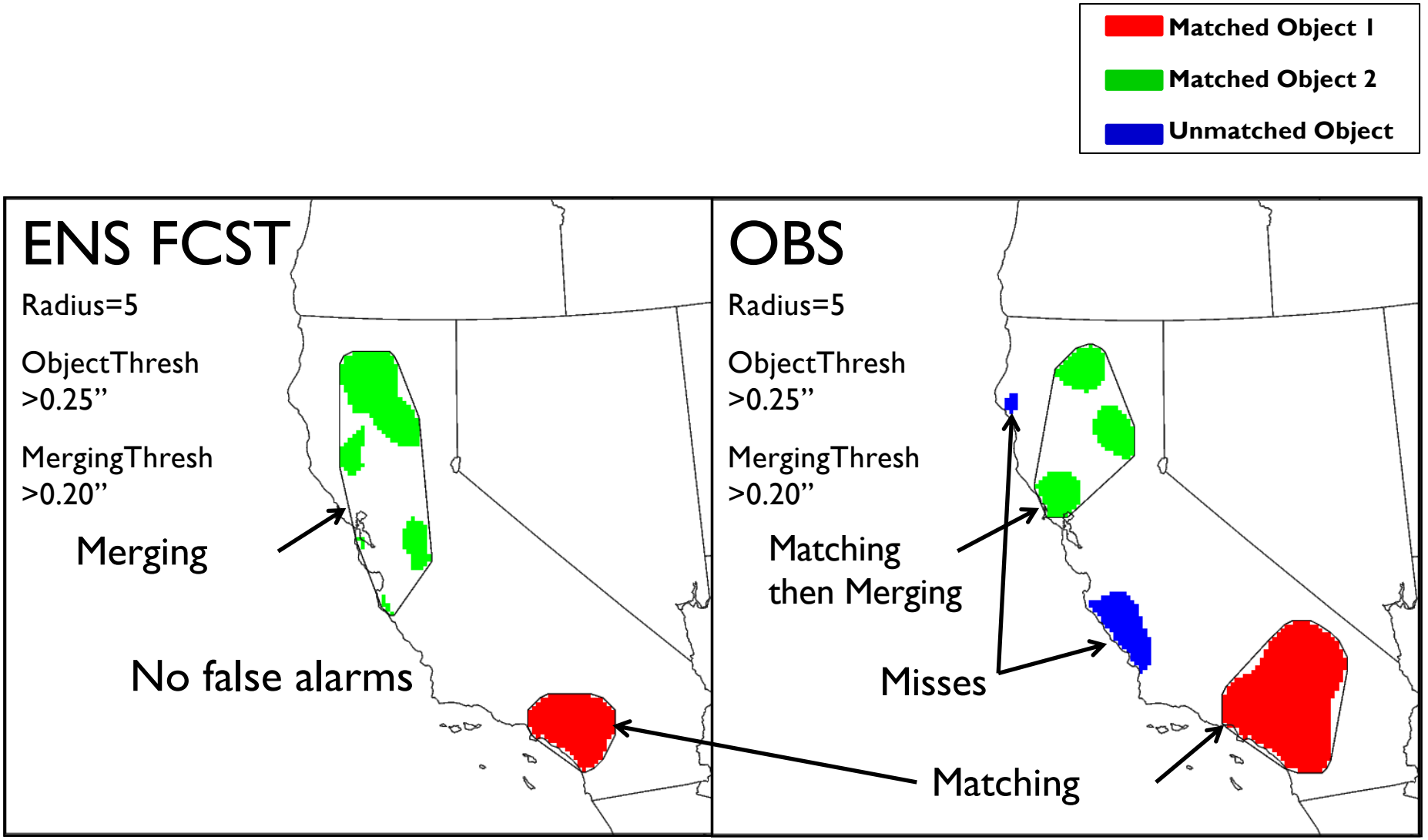
Matching & Merging

Merging: Associating objects in the same field.

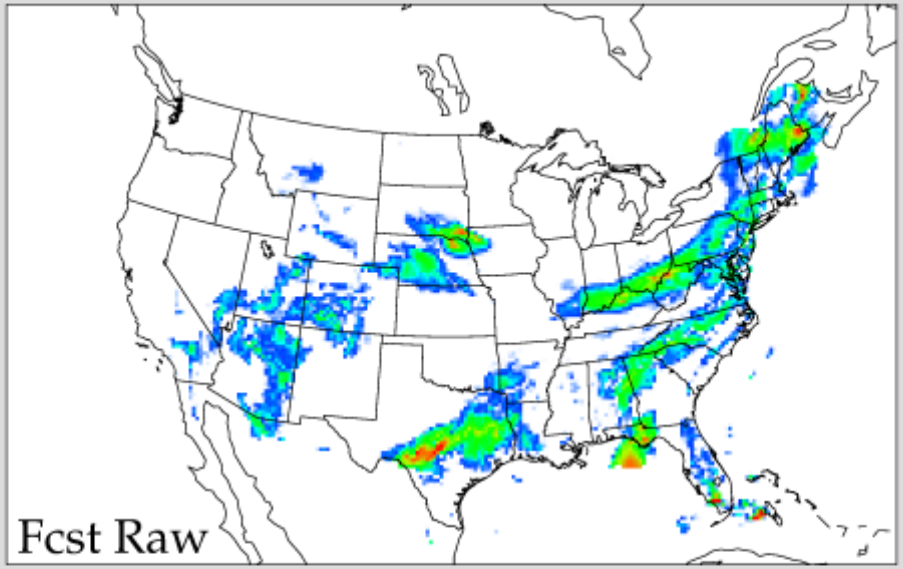
Matching: Associating objects in different fields.

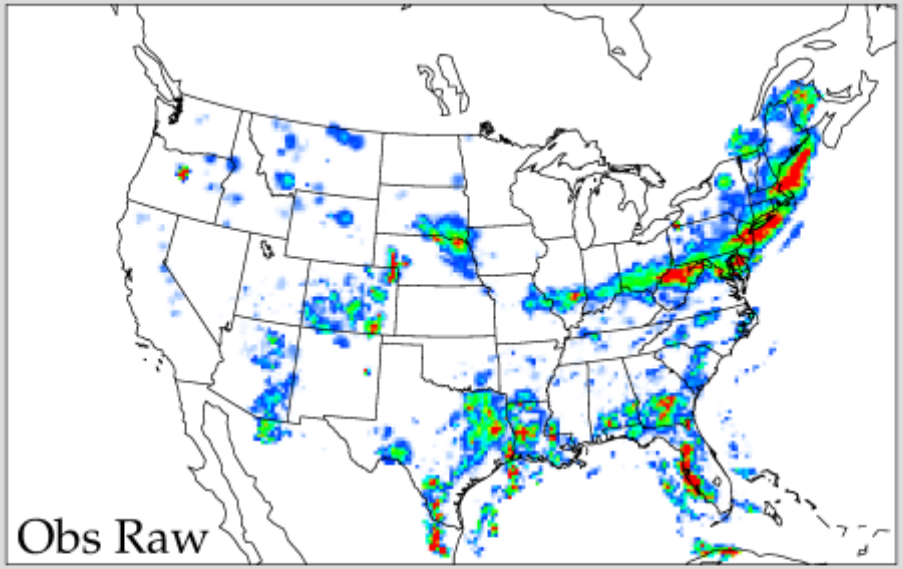
MODE does this using a Fuzzy-Logic engine.

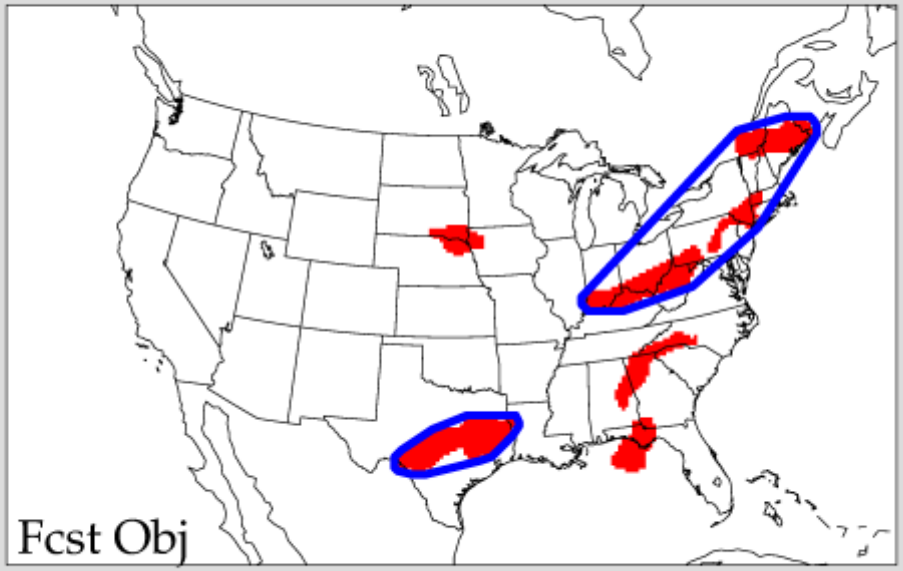
MODE Example

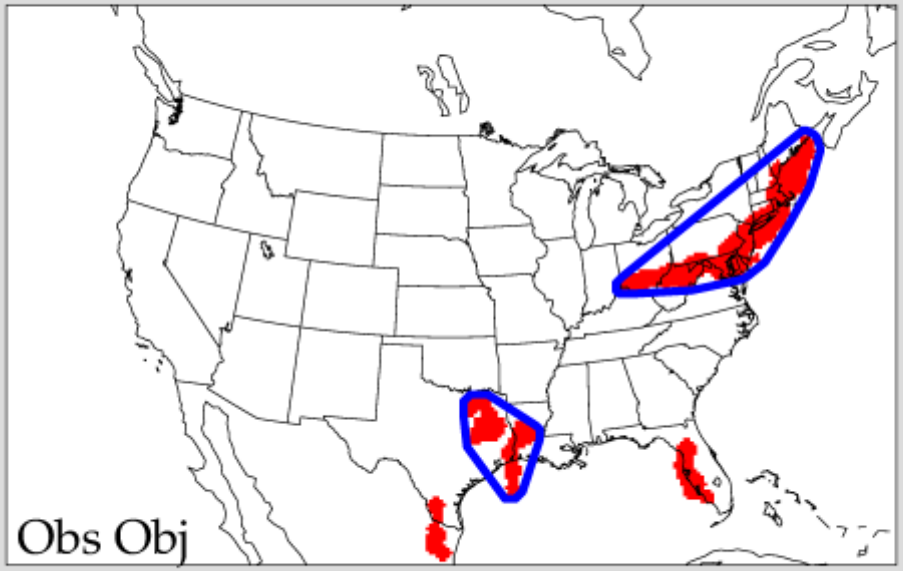


Example of Matching & Merging









Fuzzy Logic Intro

Attributes

Interest Maps

Confidence Maps

Weights

Total Interest

Object Attributes

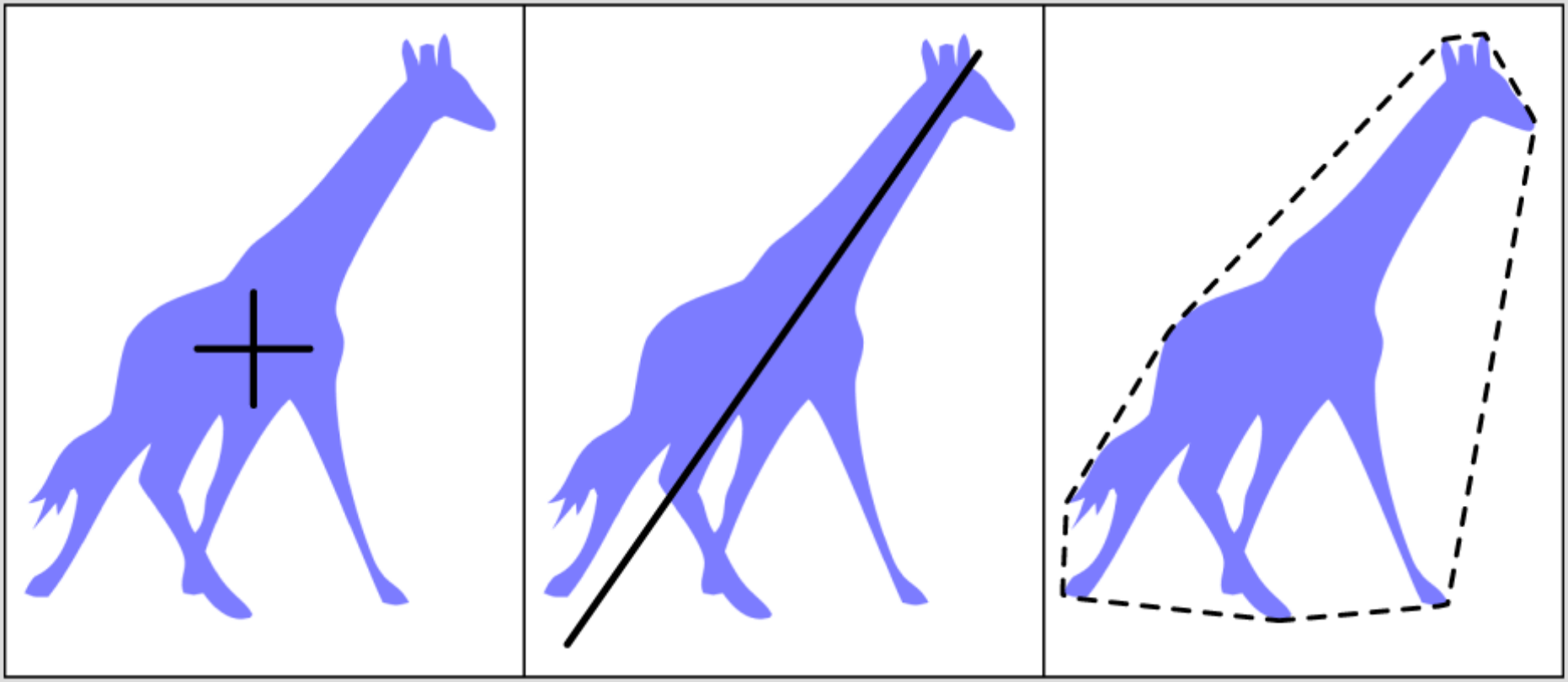
Single:

- Area
- Centroid
- Axis Angle
- Median Intensity
- Complexity
- Aspect Ratio
- Curvature

Pair:

- Centroid Distance
- Angle Difference
- Median Intensity Ratio
- Intersection Area
- Convex Hull Distance
- Boundary Distance
- Area Ratio

Example Single Attributes



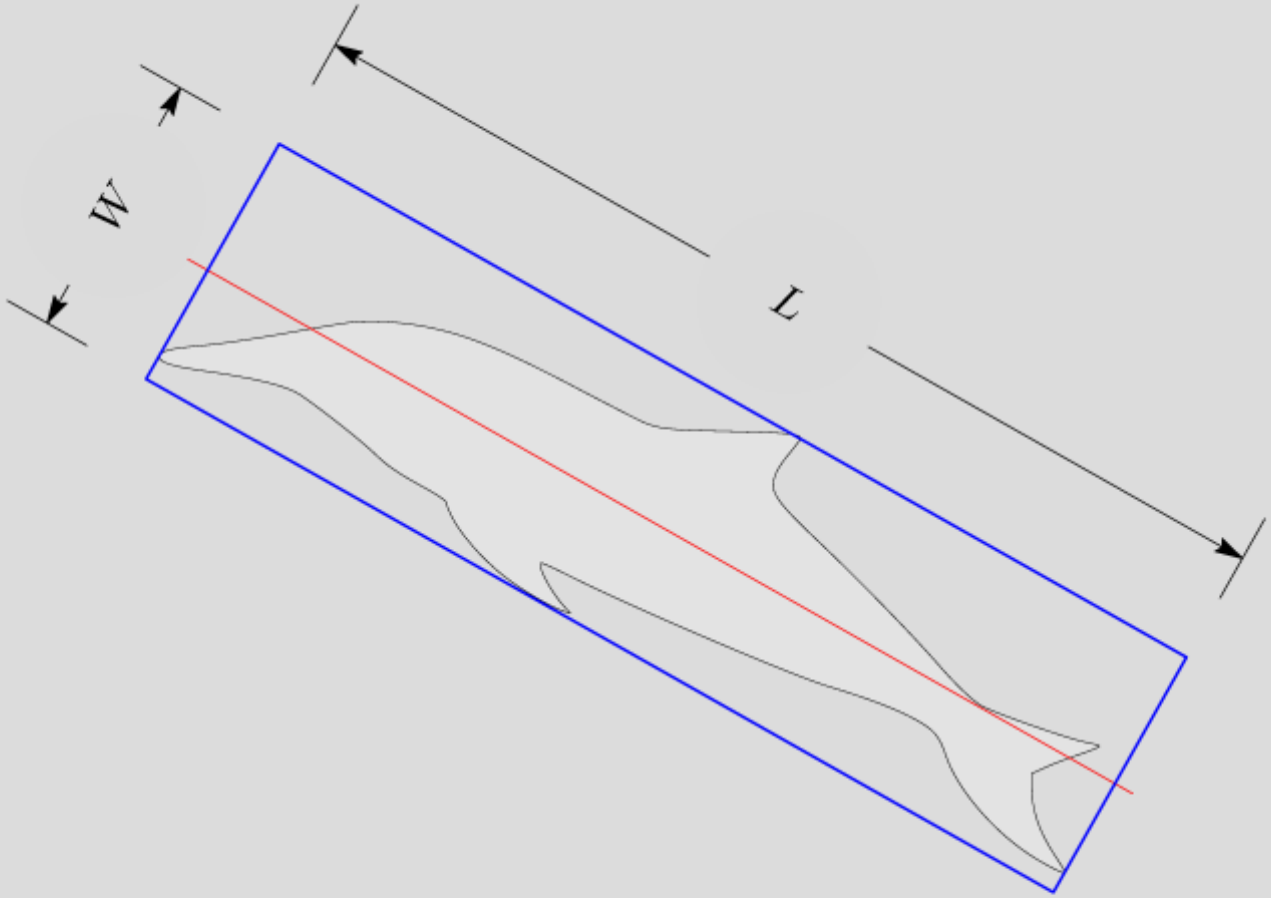
Centroid Axis Convex Hull

The Trouble with Least Squares

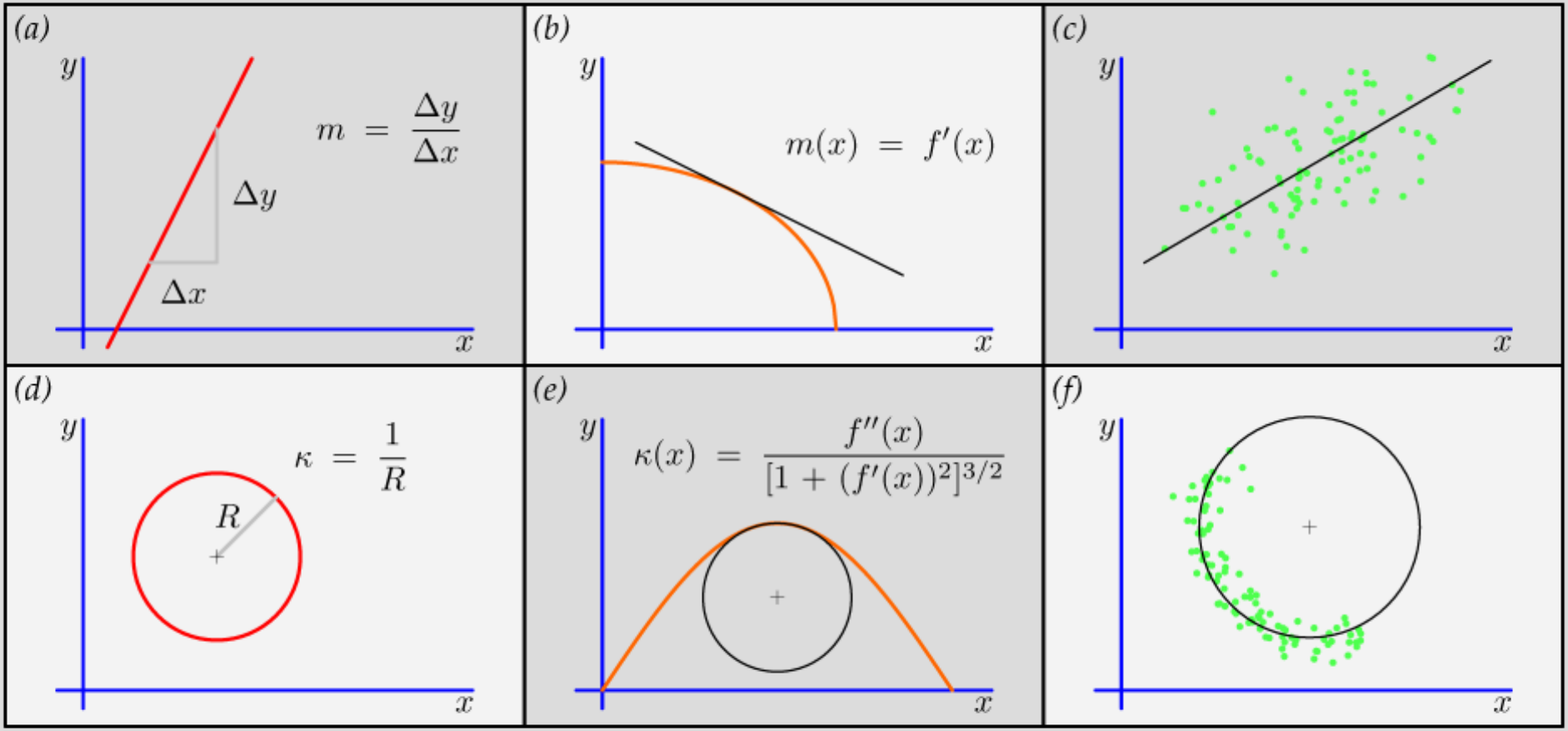
Axis

Least Squares

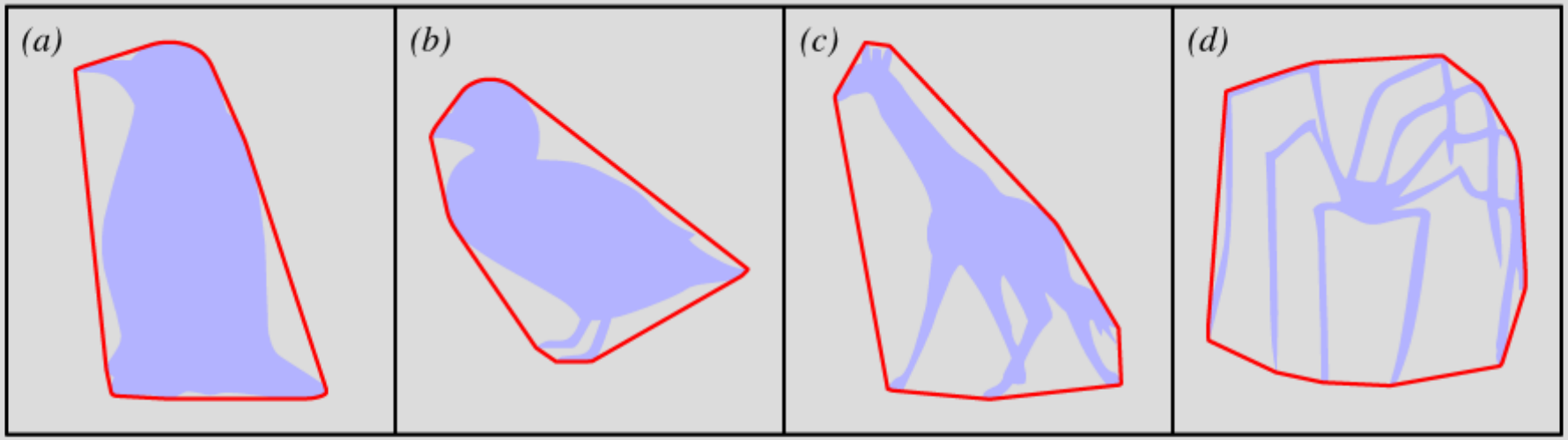
Aspect Ratio



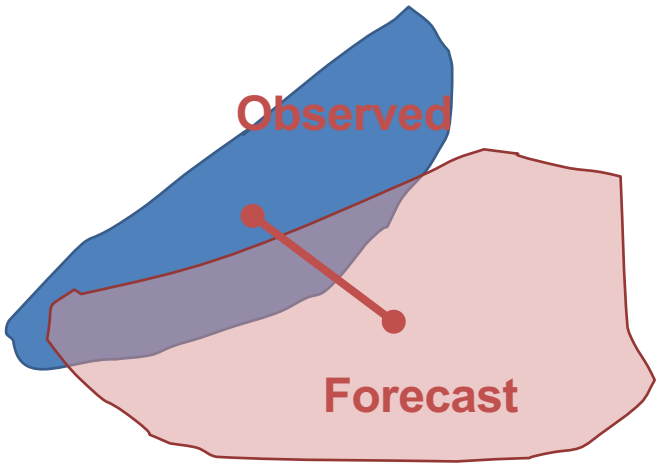
Curvature



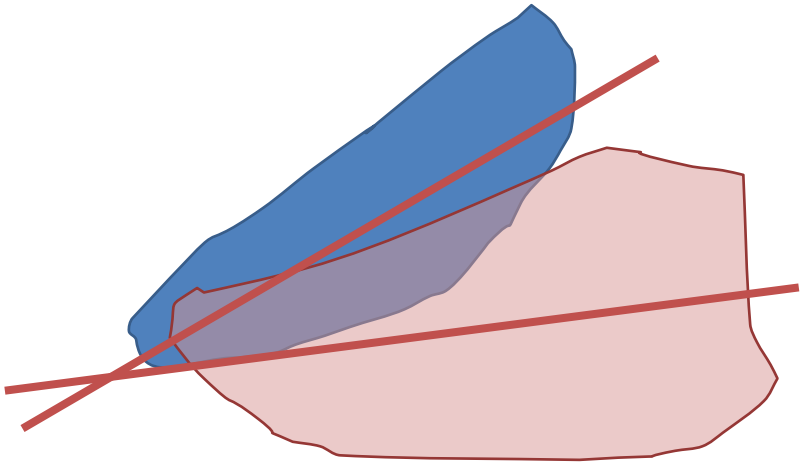
Complexity



Use of MODE Pair Attributes

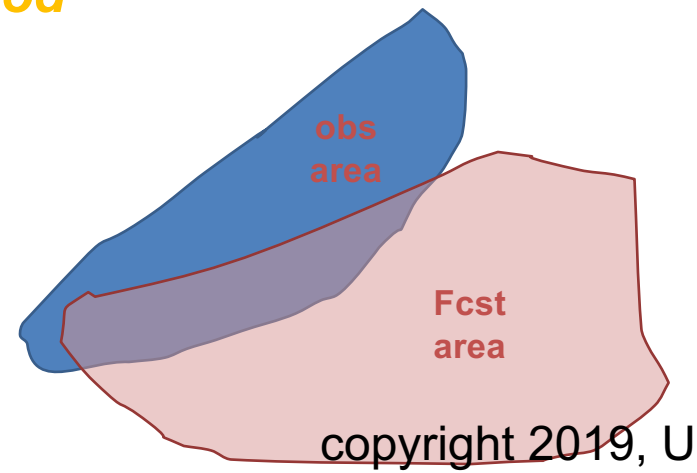


Centroid Distance: Quantitative measure of forecast spatial Displacement.
Small is good



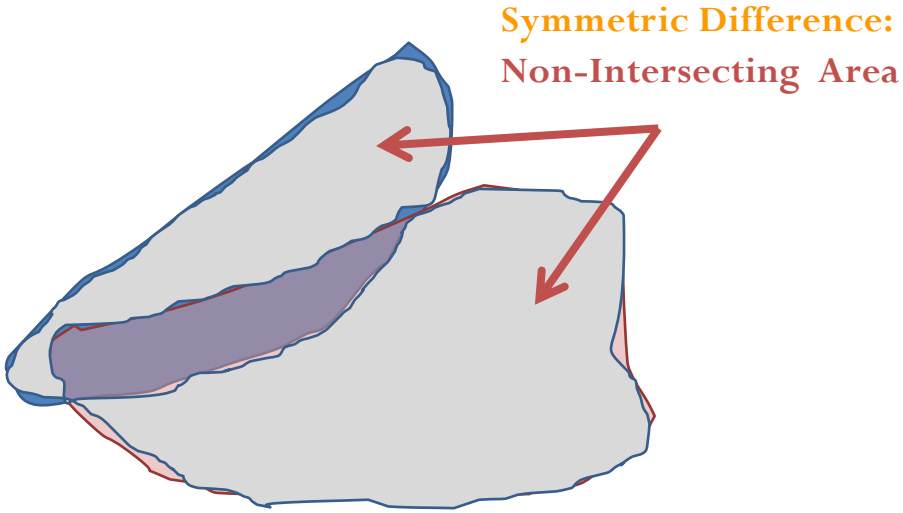
Axis Angle: For non-circular Objects, measure of orientation errors.
Small is good

Area Ratio = $\frac{\text{small Area}}{\text{Large Area}}$

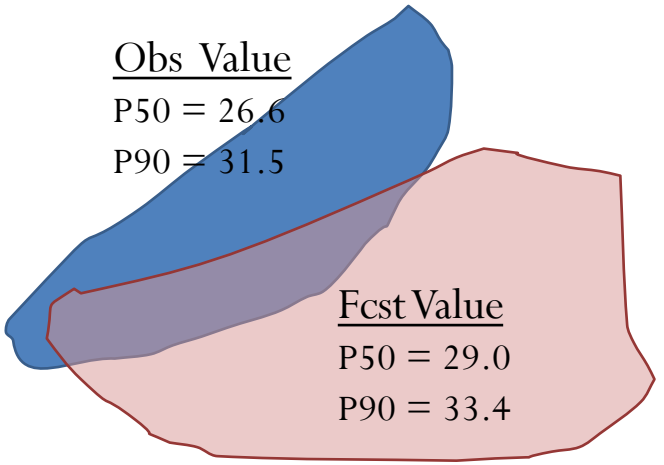


Area Ratio: Provides an objective measure of whether there is an over- or under-prediction of areal extent of forecast.
Close to 1 is good

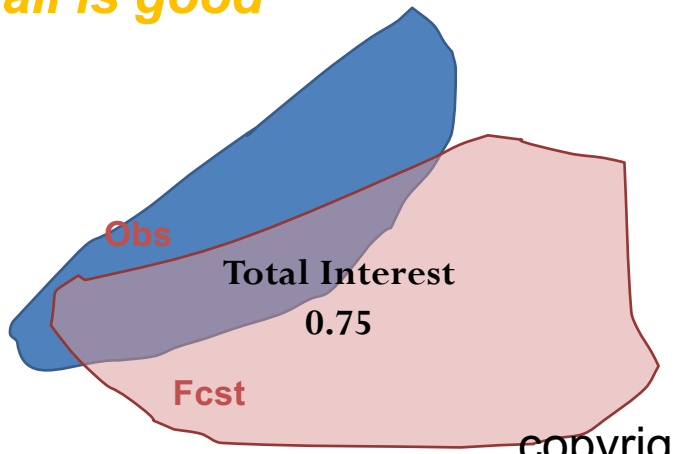
Use of MODE Pair Attributes



Symmetric Diff: Summary statistic for how well Forecast and Observed objects match.
Small is good

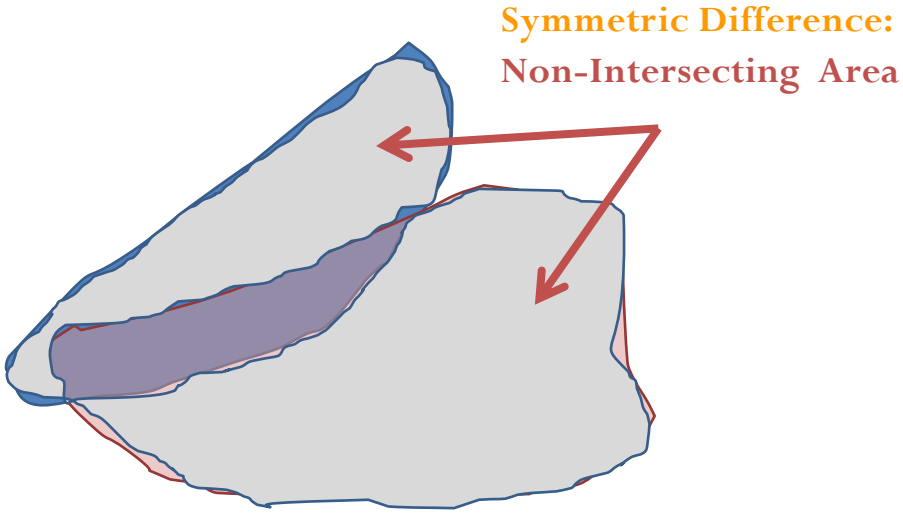


P50 | P90 Int: Objective measures of Median (50th percentile) and near-Peak (90th percentile) intensities in objects.
Ratio close To 1 is good

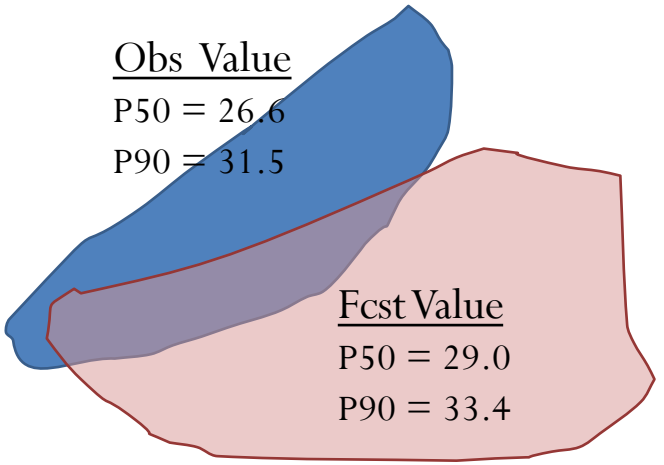


Total Interest: Summary statistic derived from fuzzy logic engine with user-defined Interest Maps for all these attributes plus some others.
Close to 1 is good

Use of MODE Pair Attributes

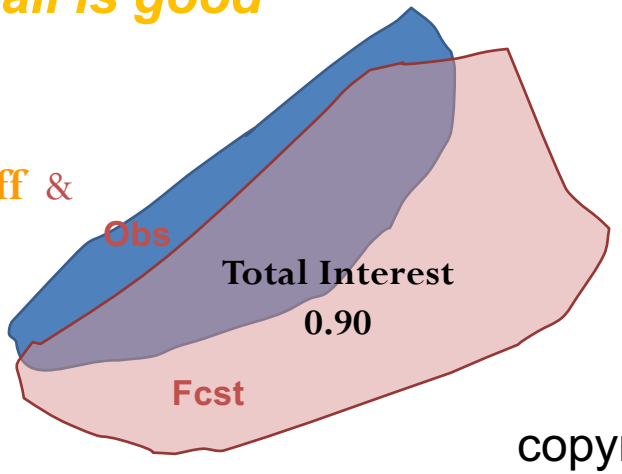


Symmetric Diff: Summary statistic for how well Forecast and Observed objects match.
Small is good



P50 | P90 Int: Objective measures of Median (50th percentile) and near-Peak (90th percentile) intensities in objects.
Ratio close To 1 is good

Angle_diff & **Sym_diff** less so **Total Int.** higher

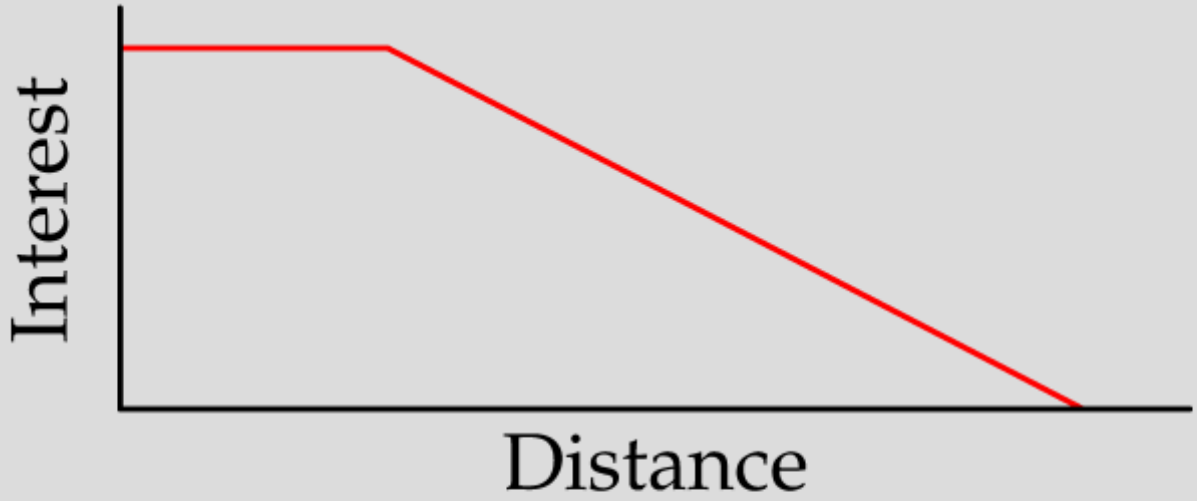


Total Interest: Summary statistic derived from fuzzy logic engine with user-defined Interest Maps for all these attributes plus some others.
Close to 1 is good

Interest Maps

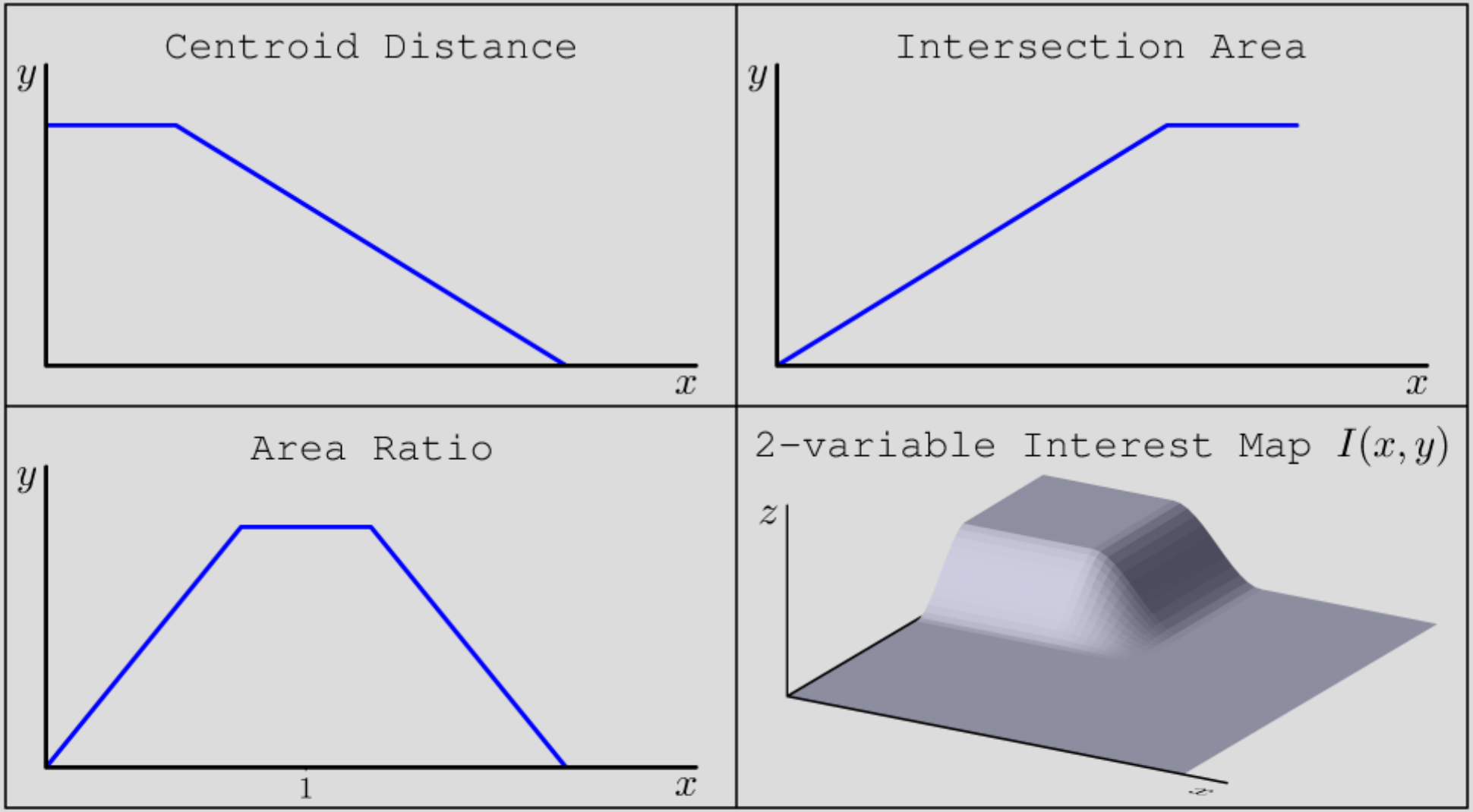
Map attributes to interest values.

Example: Centroid Distance



All interest maps can be changed in the config file.

Interest Maps

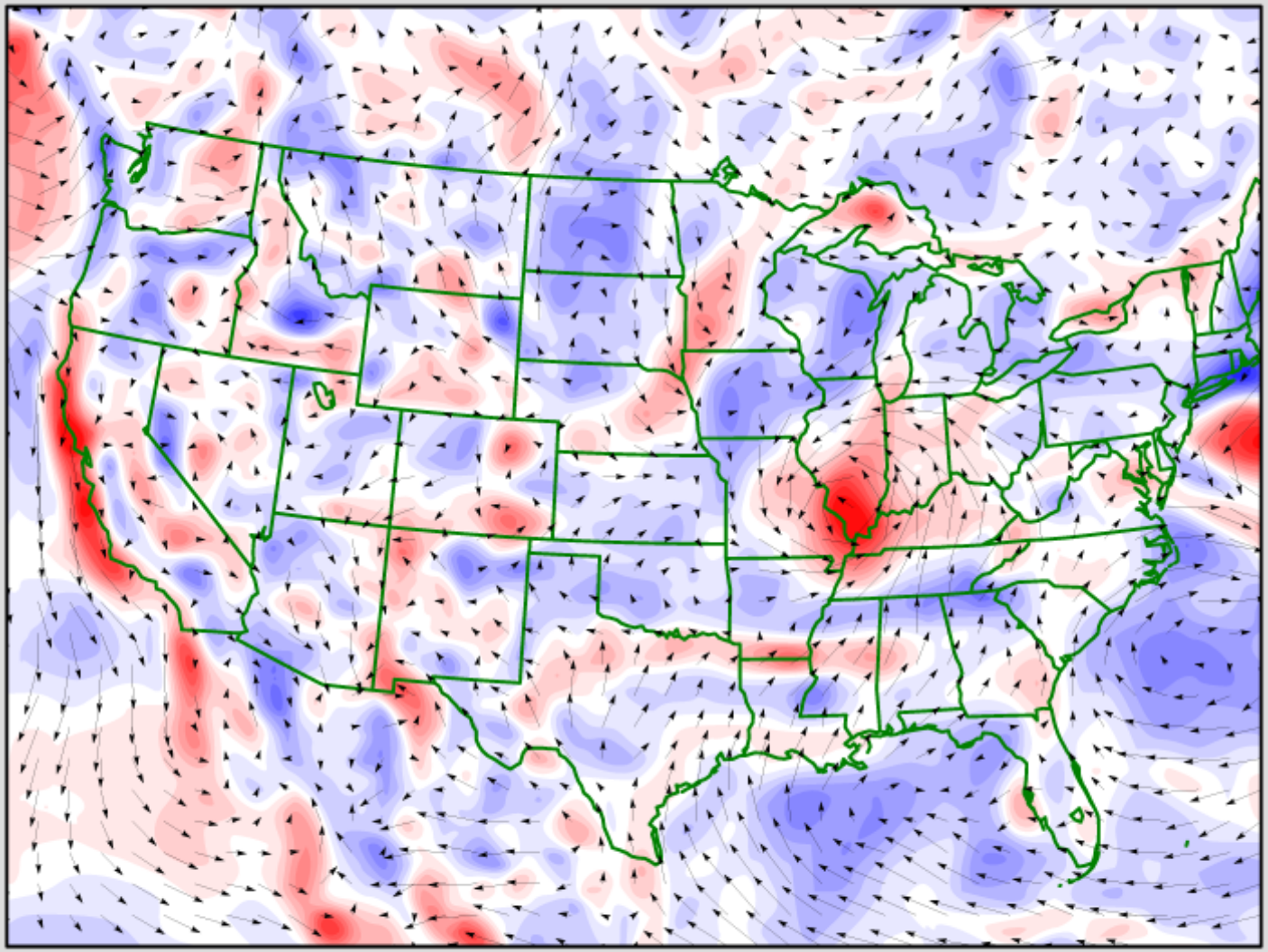


Weights

Express relative importance of different attributes in matching and merging.

All weights can be changed in the config file.

Other Objects



MODE Input and Usage

- Input Files: Gridded forecast and observation
  - GRIB1, GRIB2 (Unified Post-Processor, NCEP, other)
  - NetCDF (PCP-Combine, wrf_interp, CF-compliant)

- Usage: mode
  - fcst_file
  - obs_file
  - config_file
  - [-config_merge merge_config_file]
  - [-outdir path]
  - [-log file]
  - [-v level]

fcst_file	Gridded forecast file
obs_file	Gridded observation file
config_file	ASCII configuration file
-config_merge	Second configuration file for fuzzy engine merging
-outdir	Output directory to be used
-log	Optional log file
-v	Level of logging

MODE Config File

```
// Run all permutations of radius and threshold
//
quilt = TRUE;

//
// Forecast and observation fields to be verified
//
fcst = {
   field = {
      name  = "tasAnomDJF";
      level = "(0,*,*)";
   }

   file_type          = NETCDF_NCCF;
   censor_thresh      = [];
   censor_val         = [];
   conv_radius        = [2,4]; // in grid squares
   conv_thresh        = [<=-0.5,<=-0.55,<=-0.6,<=-0.75,>=0.50,>=0.55,>=0.6,>=0.75];
   vld_thresh         = 0.5;
   area_thresh        = NA;
   inten_perc_value   = 100;
   inten_perc_thresh  = NA;
   merge_thresh       = [>=1.25,>=1.25,>=1.25,>=1.25,>=1.25,>=1.25,>=1.25,>=1.25];
   merge_flag         = THRESH;
}
obs = fcst;
```

```
//
// Handle missing data
//
mask_missing_flag = BOTH;

//
// Match objects between the forecast and observation fields
//
match_flag = MERGE_BOTH;

//
// Maximum centroid distance for objects to be compared
//
max_centroid_dist = 800.0/grid_res;

////////////////////////////////////////////////////////////////////////

//
// Verification masking regions

mask = {
   grid      = "";
   grid_flag = NONE; // Apply to NONE, FCST, OBS, or BOTH
   poly      = "";
   poly_flag = NONE; // Apply to NONE, FCST, OBS, or BOTH
}
```

Config File

```
// Fuzzy engine interest functions
//
interest_function = {

   centroid_dist = (
      (             0.0, 1.0 )
      (  60.0/grid_res, 1.0 )
      ( 600.0/grid_res, 0.0 )
   );

   boundary_dist = (
      (             0.0, 1.0 )
      ( 400.0/grid_res, 0.0 )
   );

   convex_hull_dist = (
      (             0.0, 1.0 )
      ( 400.0/grid_res, 0.0 )
   );

   angle_diff = (
      (  0.0, 1.0 )
      ( 30.0, 1.0 )
      ( 90.0, 0.0 )
   );

   aspect_diff = (
      (  0.00, 1.0 )
      (  0.10, 1.0 )
      (  0.75, 0.0 )
   );

   corner   = 0.8;
```

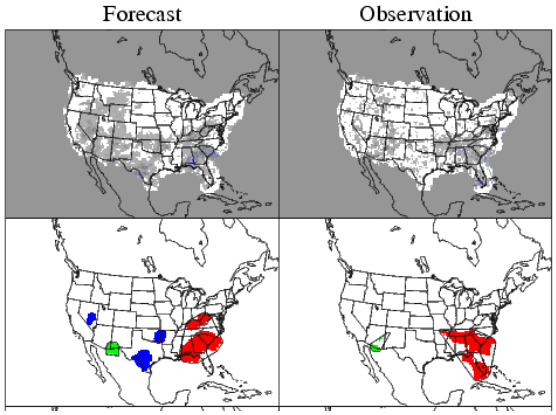
```
// Fuzzy engine weights
//
weight = {
   centroid_dist    = 2.0;
   boundary_dist    = 4.0;
   convex_hull_dist = 0.0;
   angle_diff       = 1.0;
   aspect_diff      = 0.0;
   area_ratio       = 1.0;
   int_area_ratio   = 2.0;
   curvature_ratio  = 0.0;
   complexity_ratio = 0.0;
   inten_perc_ratio = 0.0;
   inten_perc_value = 50;
}
```

```
//
// Total interest threshold for determining matches
//
total_interest_thresh = 0.5;

//
// Interest threshold for printing output pair information
//
print_interest_thresh = 0.0;
```

MODE Output

- PostScript
  - object pictures, definitions
  - matching/merging strategy
  - total interest for each object pair
- ASCII Text
  - attributes of simple, paired objects, clusters
    - size, shape, position, separation, total interest
  - verification scores (CSI, bias, etc.) for objects
- netCDF
  - gridded object fields
  - view with ncview

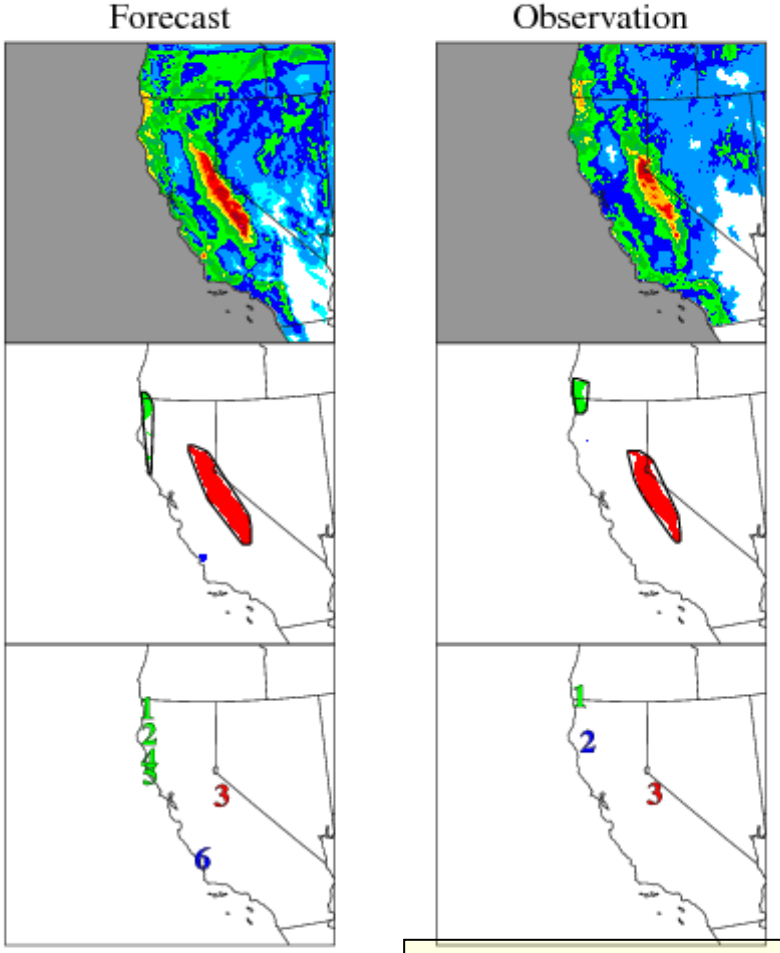


Fcst	Obs
4	1
6	4
2	1
6	3
5	4
5	3
6	2
5	2
1	3
1	2
4	2
1	1
5	1
7	2
6	1
2	3
3	3
4	3
7	1

```
//
// NetCDF matched pairs, PostScript,
// and contingency table output files
//
ps_plot_flag    = TRUE;
nc_pairs_flag   = {
   latlon       = TRUE;
   raw          = TRUE;
   object_raw   = TRUE;
   object_id    = TRUE;
   cluster_id   = TRUE;
   polylines    = TRUE;
}
ct_stats_flag   = TRUE;
```

MODE: APCP_24_A24_ENS_MEAN at A24 vs APCP_24 at A2

object pictures



Fcst	Obs	Interest
3	3	1.0000
1	1	0.9360
2	2	0.6436
5	2	0.6372
4	2	0.5085
2	1	0.4060
1	2	0.3871
4	1	0.3545
3	1	0.3422
6	3	0.3265
1	3	0.3141
4	3	0.2813
3	2	0.2719
2	3	0.2704
5	1	0.2406
5	3	0.2266
6	1	0.2203
6	2	0.1936

Total Interest of object pairs

Pairs above dashed line processed further

Field names
model description

	Forecast	Observation
Model:	hmt-ens-d01	
Field:	APCP_24_A24_ENS_MEAN	
Level:	A24	A24
Units:	kg/m^2	kg/m^2
Initial:	20110216 12:00:00	20110216 12:00:00
Valid:	20110217 12:00:00	20110217 12:00:00
Accum:	24:00:00	24:00:00

Weight of object attributes

Centroid/Boundary:	2.00	4.00
Convex Hull/Angle:	0.00	1.00
Area/Intersection Area:	4.00	4.00
Complexity/Intensity:	0.00	2.00
Total Interest Thresh:	0.70	

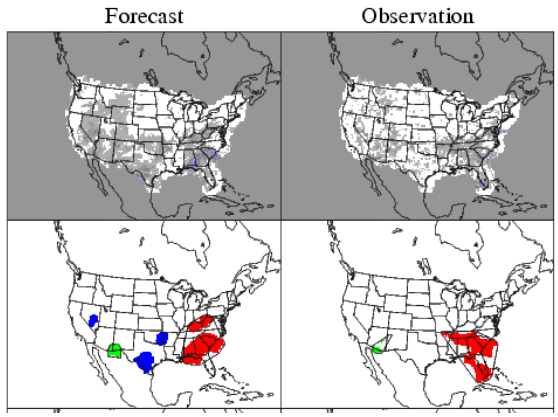
	Forecast	Observation
Mask M/G/P:	on/off/off	on/off/off
Raw Thresh:	>=0.00	>=0.00
Conv Radius:	2 gs	2 gs
Conv Thresh:	>=25.40	>=25.40
Area Thresh:	>=0 gs	>=0 gs
Inten Thresh:	p100>=0.00	p100>=0.00
Merge Thresh:	>=20.00	>=20.00
Merging:	thresh	thresh
Matching:	match/merge	
Simple/M/U:	6/5/1	3/2/1
Area:	696 gs	589 gs
Area M/U:	674/22	585/4
Cluster:	2	2
MMI:	0.6404	0.9360
MMI (F+O):	0.6436	

Definition of objects

- smoothing radius
- intensity threshold
- area threshold
- matching and/or merging
- # and area of objects
- **M**edian **M**ax. **I**nterest (**MMI**)

MODE Output

- PostScript
  - object pictures, definitions
  - matching/merging strategy
  - total interest for each object pair
- ASCII Text
  - attributes of simple, paired objects, clusters
    - size, shape, position, separation, tota interest
  - verification scores (CSI, bias, etc. for objects
- netCDF
  - gridded object fields
  - view with ncview




```
//
// NetCDF matched pairs, PostScript,
// and contingency table output files
//
ps_plot_flag     = TRUE;
nc_pairs_flag    = {
   latlon        = TRUE;
   raw           = TRUE;
   object_raw    = TRUE;
   object_id     = TRUE;
   cluster_id    = TRUE;
   polylines     = TRUE;
}
ct_stats_flag    = TRUE;
```

ASCII Output

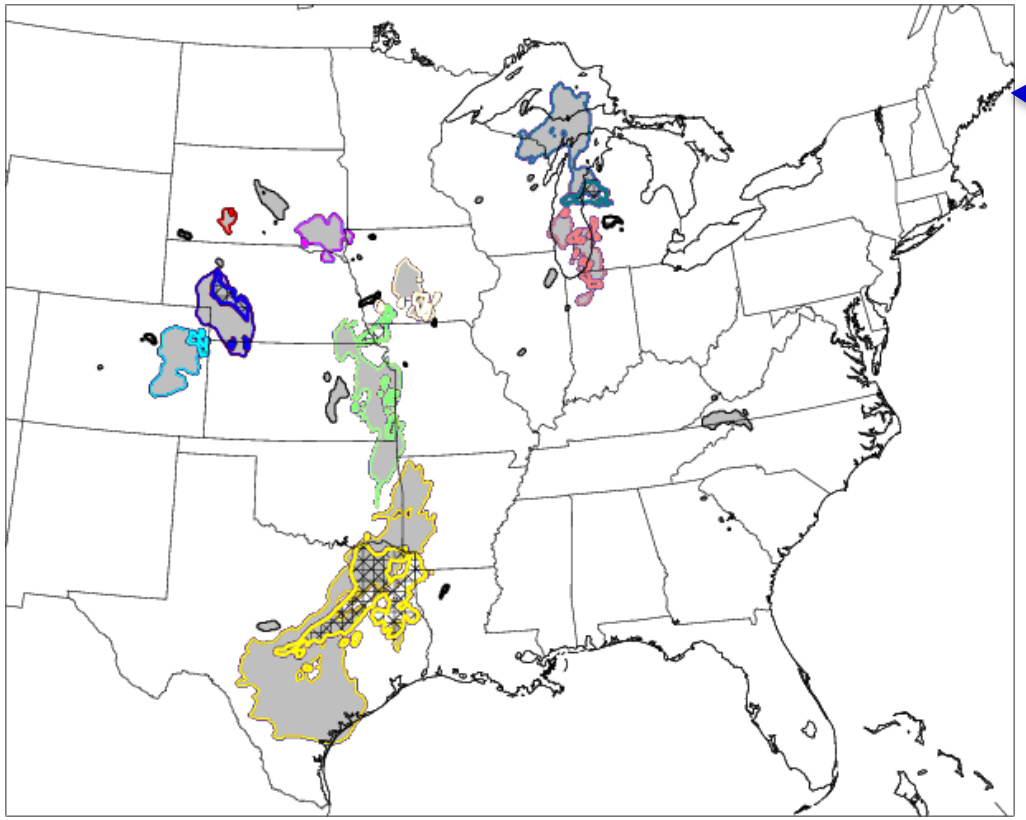
Object Attribute file (*_obj.txt)

- Header with fields names and object definition info
- Object ID and Category
- Simple Object Attributes
  - Simple Obj. Centroid info, Length, Width, Area, etc…
- Matched Pair/Composite information
  - Centroid Distance, Angle Difference, Symmetric Difference, etc…
- NA's for not relevent output

Contingency Table Stat file (*_cts.txt)

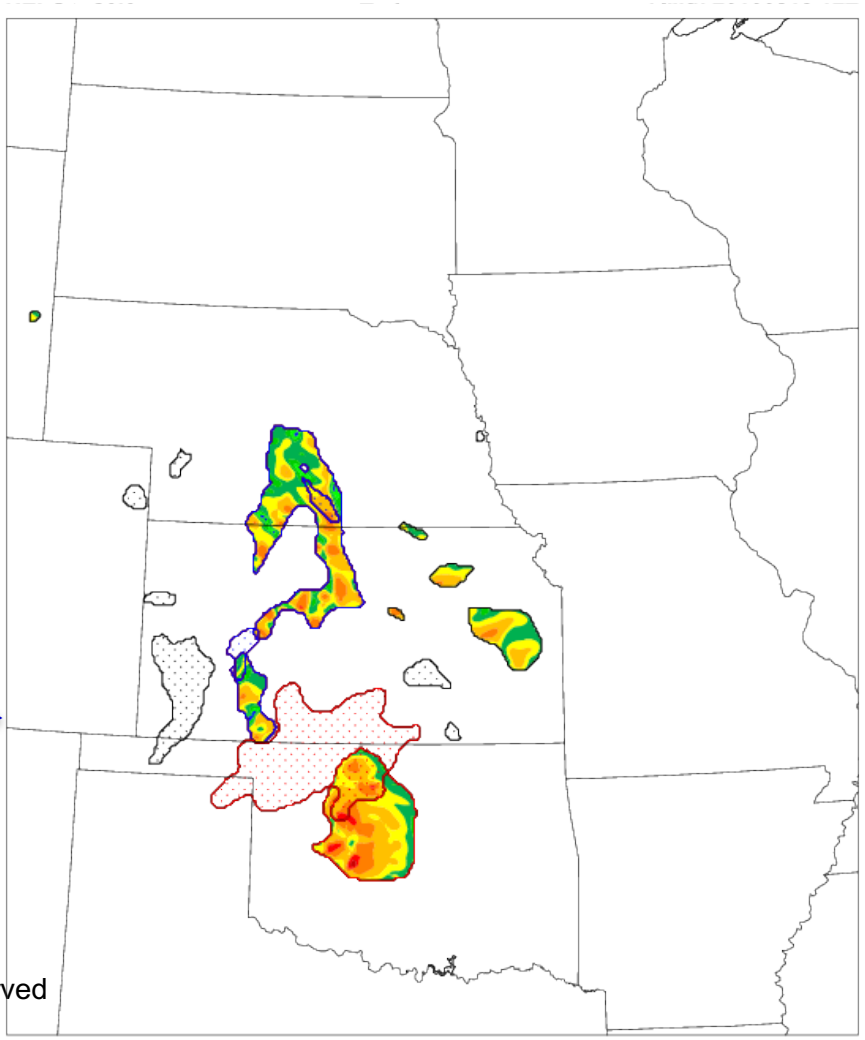
- Header with fields names and object definition info
- Contingency Table counts
  - hits, false alarms, misses and correct negs (FY|FN_OY|ON notation)
- Contingency Table statistics such
  - BASER, FBIAS, GSS, CSI, PODY, FAR etc…

How netCDF could be used



Employ a different plotting approach to show matched clusters

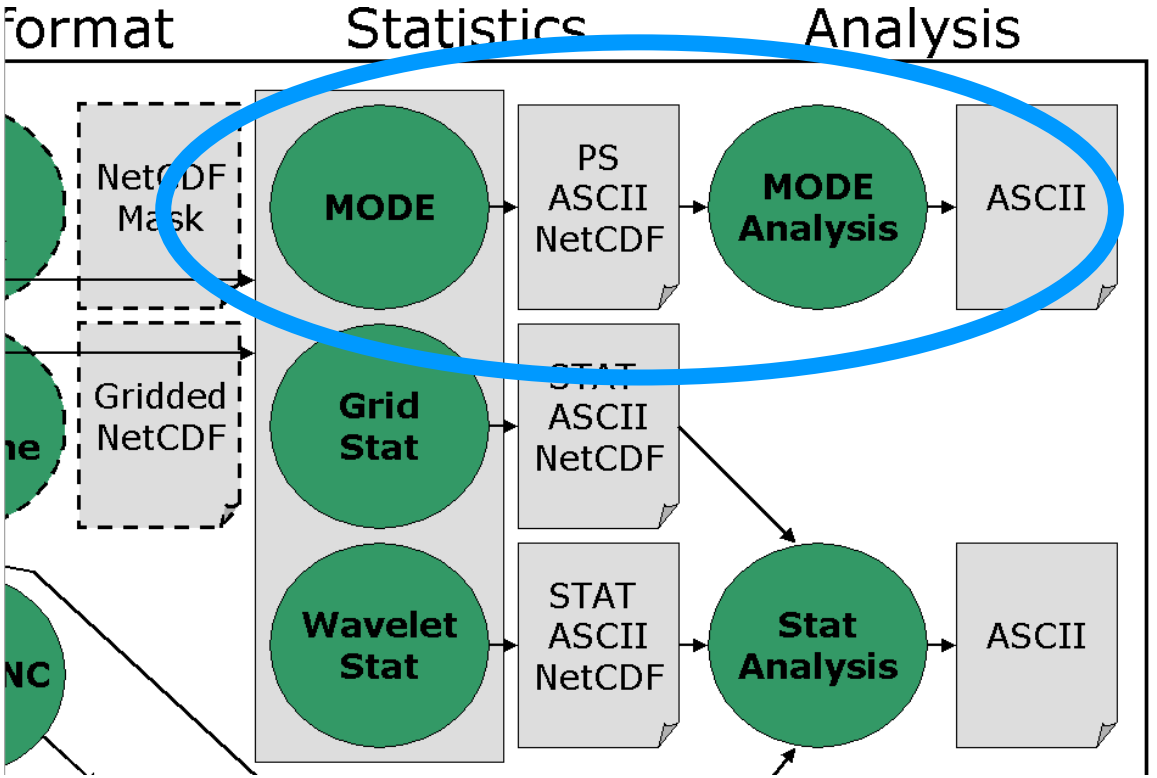
Display actual intensities inside objects (in this case Reflectivity)



Plots generated using NCL

MODE Analysis Tool

- mode_analysis



MODE_Analysis Usage

Usage: ***mode_analysis***

-lookin path

-summary or -bycase

[-column name]

[-dump_row filename]

[-out filename]

[-log filename]

[-v level]

[-help]

[MODE FILE LIST]

[-config config_file]
or [MODE LINE OPTIONS]

MODE LINE OPTIONS

Object Toggles

-fcst versus ***-obs***
Selects lines pertaining to forecast objects or observation objects

-single versus ***-pair***
Selects single object lines or pair lines

-simple versus ***-cluster***
Selects simple object lines or cluster

-matched versus ***-unmatched***
Selects matched simple object lines or unmatched simple object lines.

Other Options (each option followed by value)

-model, -fcst|obs_thr , -fcst_var , etc…

-area_min|max, -intersection_area_min|max , etc…

-centroid_x_min|max , -centroid_y_min|max,
-axis_ang_min|max, -int10_min|max,
-centroid_dist_min|max, -angle_diff_min|max, etc…

MODE Analysis Tool -*summary* Example

Command Line

```
mode_analysis -summary \
   -lookin mode_output/wrf4ncep/40km/ge03.\
   -fcst    -cluster \
   -area_min 100 \
   -column centroid_lat -column centroid_lon \
   -column area \
   -column axis_ang \
   -column length
```

Provides summary statistics for Forecast Clusters with minimum area of 100 grid-sq for the specified MODE output columns

Output

```
Total mode lines read =  393
Total mode lines kept =   17
```

Field	N	Min	Max	Mean	StdDev	P10	P25	P50	P75	P90	Sum
centroid_lat	17	31.97	46.24	38.65	3.81	33.89	36.13	38.54	40.12	43.99	657.00
centroid_lon	17	-103.89	-85.20	-96.32	5.91	-103.15	-102.65	-96.26	-93.95	-86.78	-1637.49
area	17	180.00	8393.00	2955.06	2246.49	624.80	1206.00	2662.00	3958.00	5732.20	50236.00
axis_ang	17	-88.63	85.66	12.62	64.35	-70.77	-63.86	35.04	74.37	79.24	214.60
length	17	25.25	234.76	124.41	60.99	48.85	65.37	116.67	169.37	204.57	2114.90

MODE Analysis Tool -*bycase* Example

Command Line

```
mode_analysis -bycase -lookin mode_output/wrf4ncep/40km/ge03. -single -simple
```

Output

```
Total mode lines read =  393

Total mode lines kept =  141
```

Fcst Valid Time	Area Matched	Area Unmatched	# Fcst Matched	# Fcst Unmatched	# Obs Matched	# Obs Unmatched
Apr 26, 2005 00:00:00	3210	1046	2	4	1	1
May 13, 2005 00:00:00	8892	9320	2	19	1	2
May 14, 2005 00:00:00	16994	4534	7	4	5	3
May 18, 2005 00:00:00	6057	852	3	2	2	1
May 19, 2005 00:00:00	1777	1624	1	5	2	1
May 25, 2005 00:00:00	8583	928	4	2	4	2
Jun 1, 2005 00:00:00	12456	2657	5			
Jun 3, 2005 00:00:00	7561	102	11			
Jun 4, 2005 00:00:00	11464	5715	6			

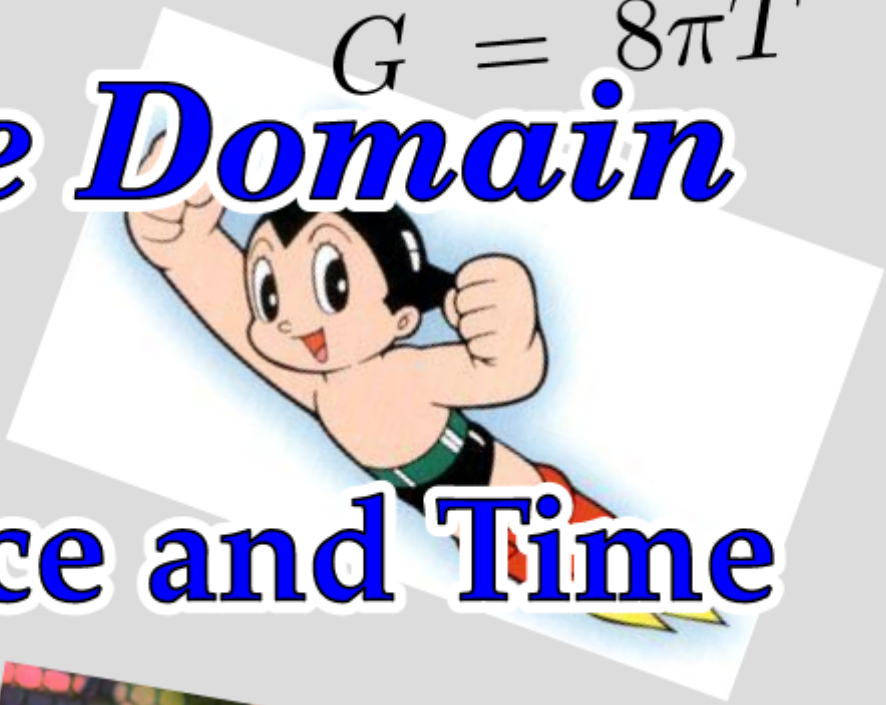
Provides tallied information for all Simple Objects for each case in directory

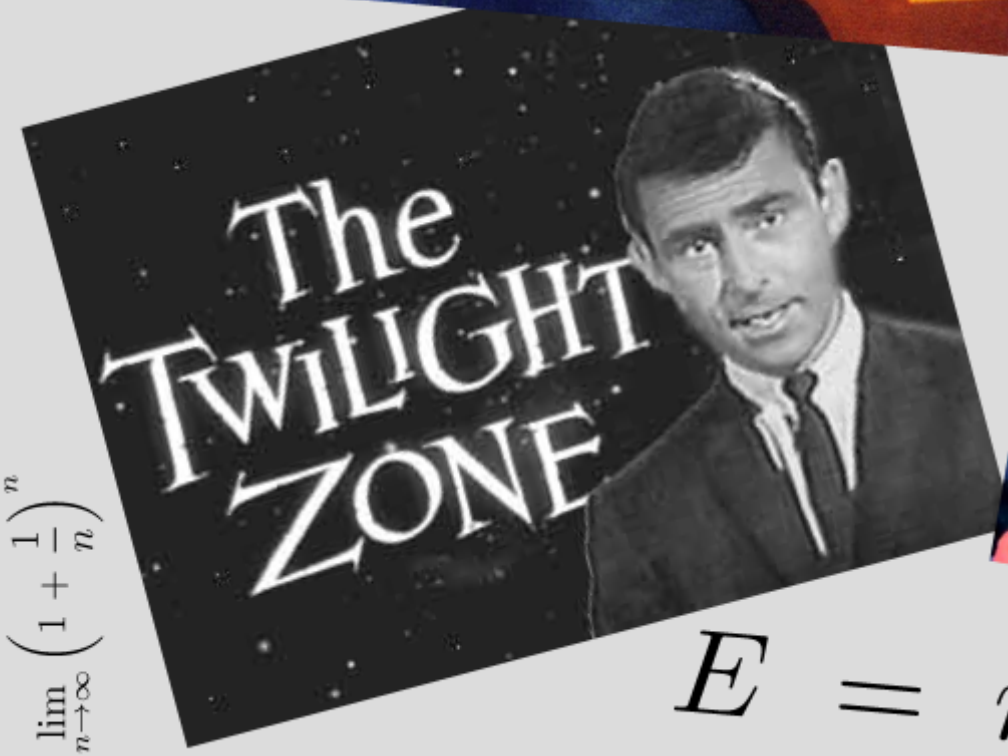
MODE Technical Note

https://opensky.ucar.edu/islandora/object/technotes%3A546

MODE Time Domain

Objects in Space and Time



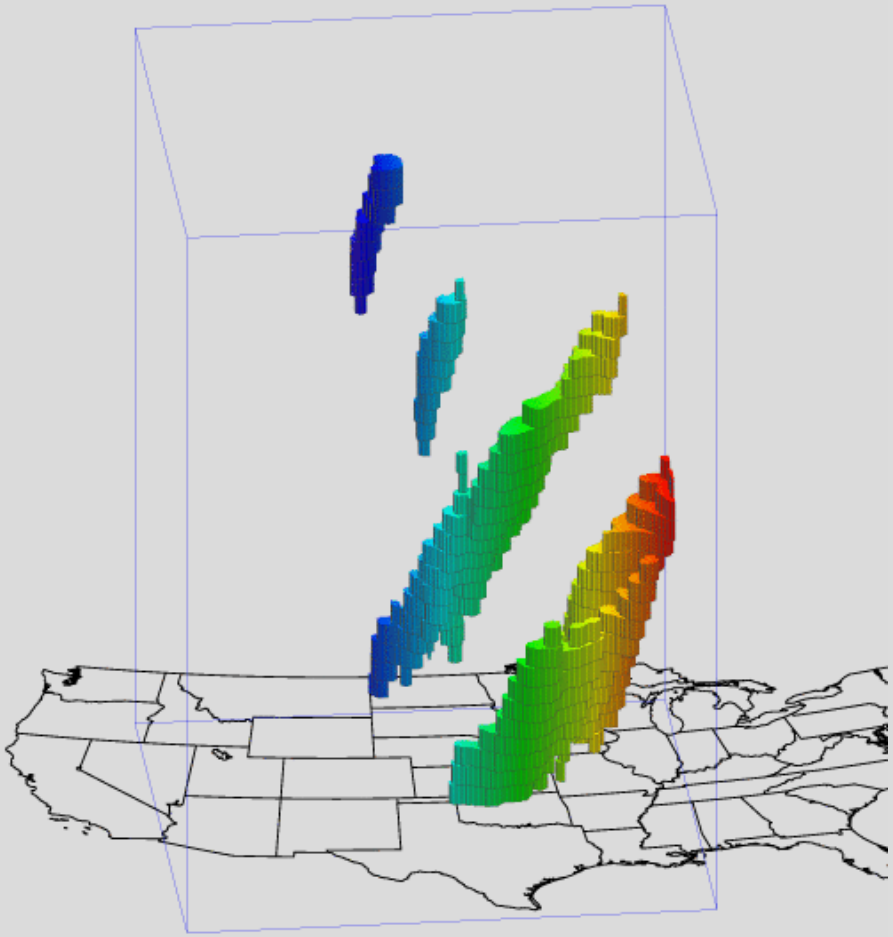


$$G = 8\pi T$$

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

$$E = mc^2$$

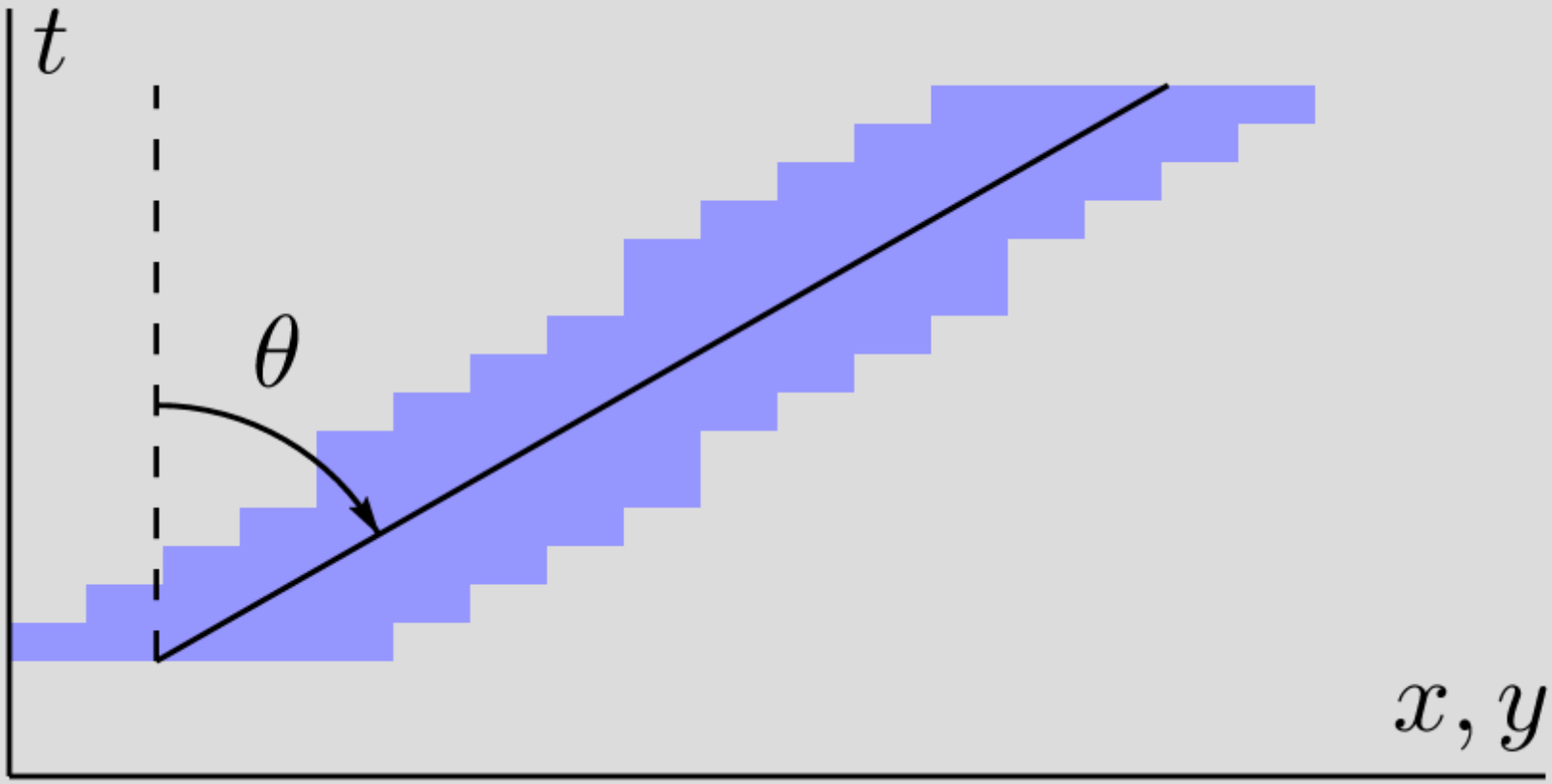
3D Objects



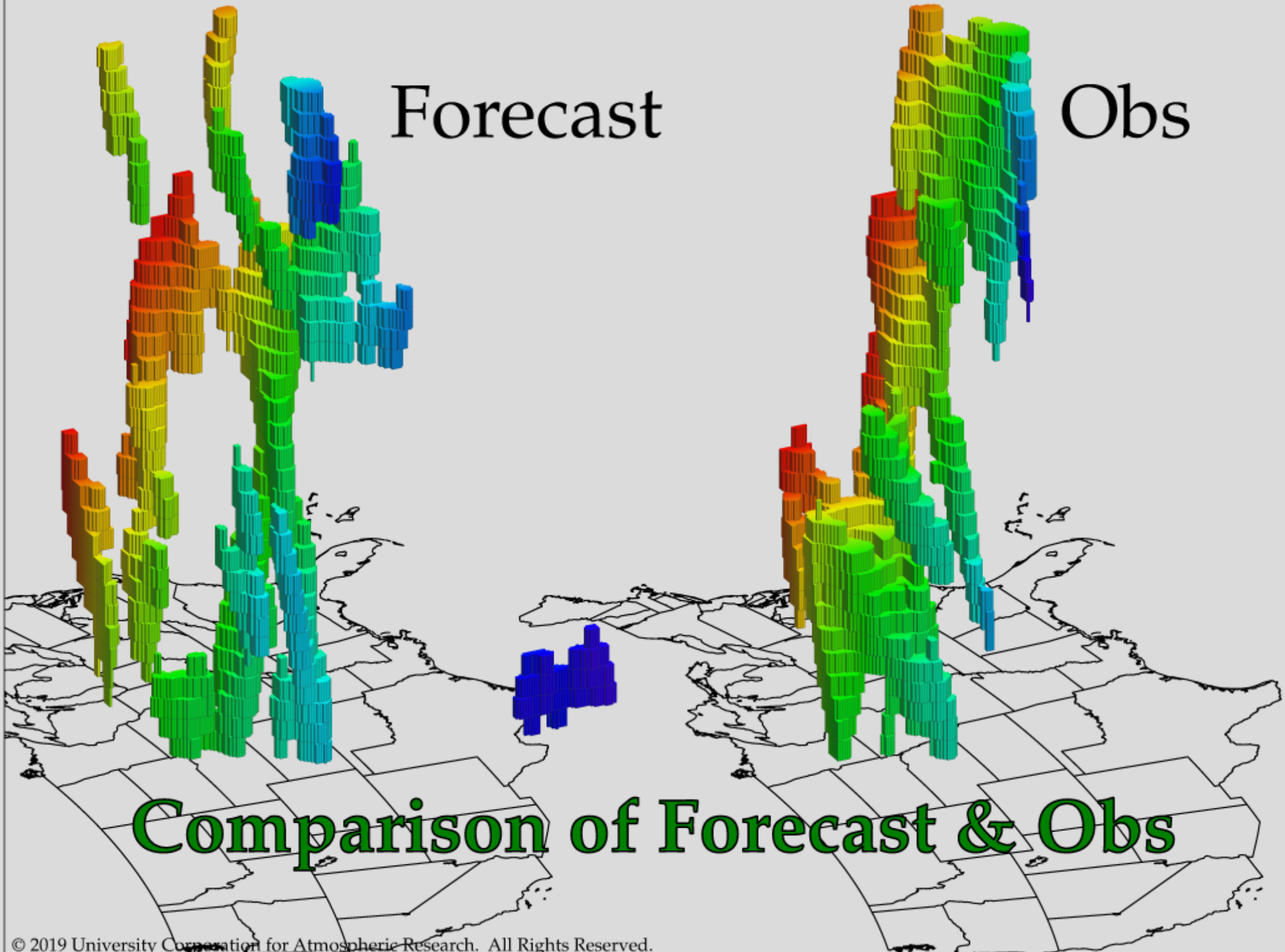
June 13, 2002
IHOP Precip Data

Vertical Dimension
is Time

Object velocity is related to the tilt angle of the axis



The inclination of the axis from the vertical (time) direction increases with object speed.



Forecast Errors

2D MODE

Location Errors

Intensity Errors

Shape Errors

Size Errors

Orientation Errors

MTD

Timing Errors

Velocity Errors

Duration Errors

Buildup & Decay

MODE-TD Usage

- Usage: mtd
  -fcst file_1 ... file_n | file_list
  -obs file_1 ... file_n | file_list
  -config config_file
  [-outdir path]
  [-log file]
  [-v level]

```
mtd -fcst fcst_files/*.grb \
-obs obs_files/*.grb \
-config MTDConfig_default \
-outdir out_dir/mtd \
```

```
bin/mtd \
-fcst data/sample_fcst/2005080700/wrfprs_ruc13_03.tm00_G212 \
data/sample_fcst/2005080700/wrfprs_ruc13_06.tm00_G212 \
data/sample_fcst/2005080700/wrfprs_ruc13_09.tm00_G212 \
data/sample_fcst/2005080700/wrfprs_ruc13_12.tm00_G212 \
data/sample_fcst/2005080700/wrfprs_ruc13_15.tm00_G212 \
data/sample_fcst/2005080700/wrfprs_ruc13_18.tm00_G212 \
data/sample_fcst/2005080700/wrfprs_ruc13_21.tm00_G212 \
data/sample_fcst/2005080700/wrfprs_ruc13_24.tm00_G212 \
-obs data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080703V_03A.nc \
data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080706V_03A.nc \
data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080709V_03A.nc \
data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080712V_03A.nc \
data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080715V_03A.nc \
data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080718V_03A.nc \
data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080721V_03A.nc \
data/tutorial/sample_obs/ST2ml_3h/sample_obs_2005080724V_03A.nc \
-outdir tutorial/out/mtd \
-config tutorial/config/MTDConfig_tutorial \
-v 2
```

MODE-TD Input/Output

- Input Files
  - Gridded forecast and observation files
    - GRIB1 output of Unified Post-Processor (or other)
    - GRIB2 from NCEP (or other)
    - NetCDF from PCP-Combine, wrf_interp, or CF-compliant

- Output:
  - Single attributes for 2D simple objects (_2d.txt)
    - Centroid lat/lon, axis angle, area, etc.
  - Single attributes for 3D composite objects (_3d_sc.txt) and 3D simple objects(_3d_ss.txt)
    - Centroid lat/lon, axis angle, volume, etc
  - Pair attributes for 3D composite objects (_3d_pc.txt) and 3D simple objects (_3d_ps.txt)
    - Axis differences, direction differences, volume ratio
  - Object NetCDF file (_obj.nc)

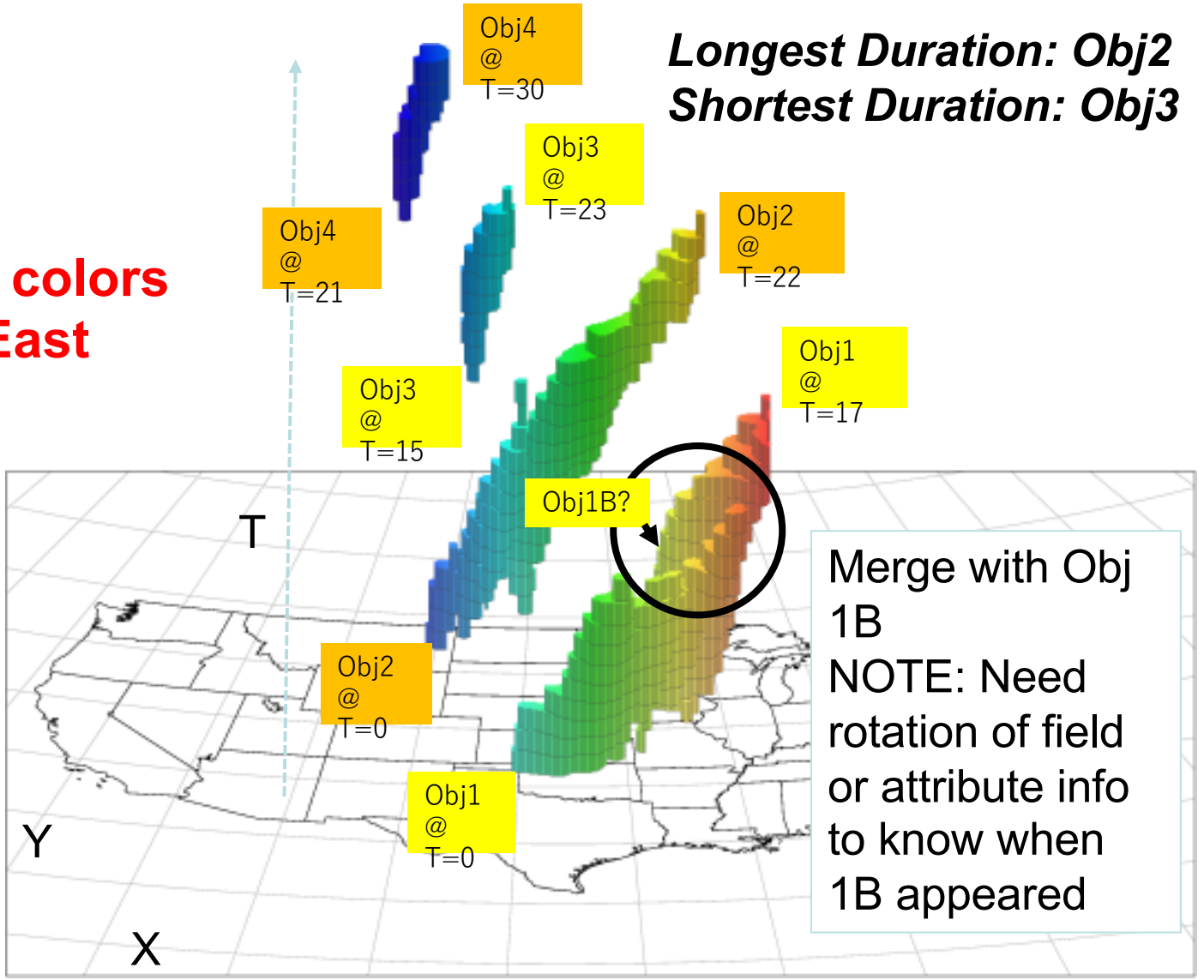
Interpretation

Cool colors – Warm colors
West – East

X-Y shows geographical location of objects

T shows timeline

Objects anchored to map are at time T=0



Longest Duration: Obj2
Shortest Duration: Obj3

Merge with Obj 1B
NOTE: Need rotation of field or attribute info to know when 1B appeared

Summary of MTD

- MODE 3D is a generalization of MODE 2D that incorporates the time dimension.
- Object matching and merging are done with a fuzzy-logic engine, as in the 2D case.
- Because one of the dimensions is not spatial, some object attributes have new interpretations.
- The addition of the time dimension produces a few new wrinkles.