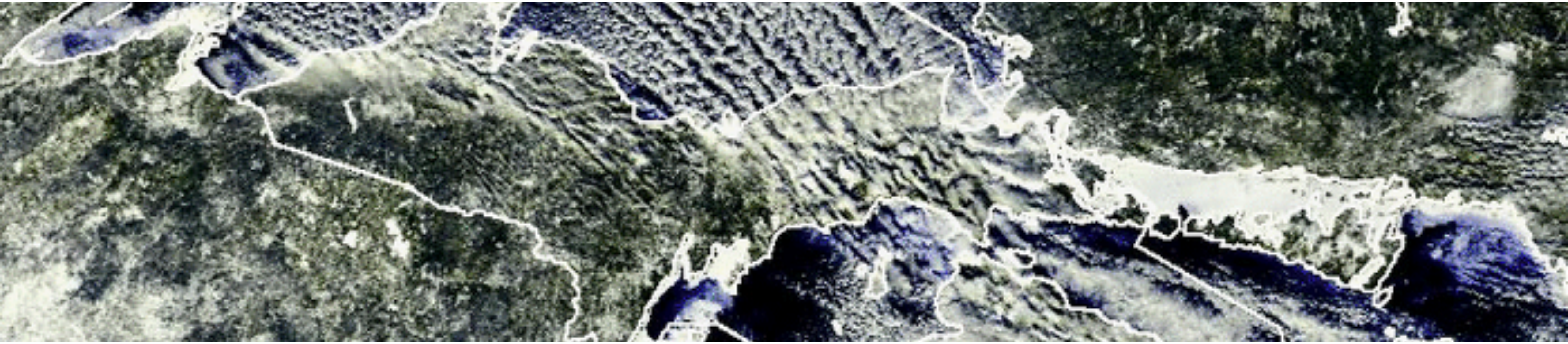
Preliminary Applications of the Geometry-Sensitive Ensemble Mean (GEM) for Lake-Effect Snowbands



Sam Ng 2014

Jon Seibert, Steven Greybush, Jia Li, Zhuomin Zhang, Fuqing Zhang

Motivation

- Address the issue of "information overload" in ensemble forecasts
  - Create new forms of ensemble mean visualization that are easy to interpret and apply
- Demonstrate a new method for creating ensemble mean images that retains more useful information than the arithmetic mean
  - Considers underlying geometry of reflectivity imagery
  - Flexible enough to handle many cloud types

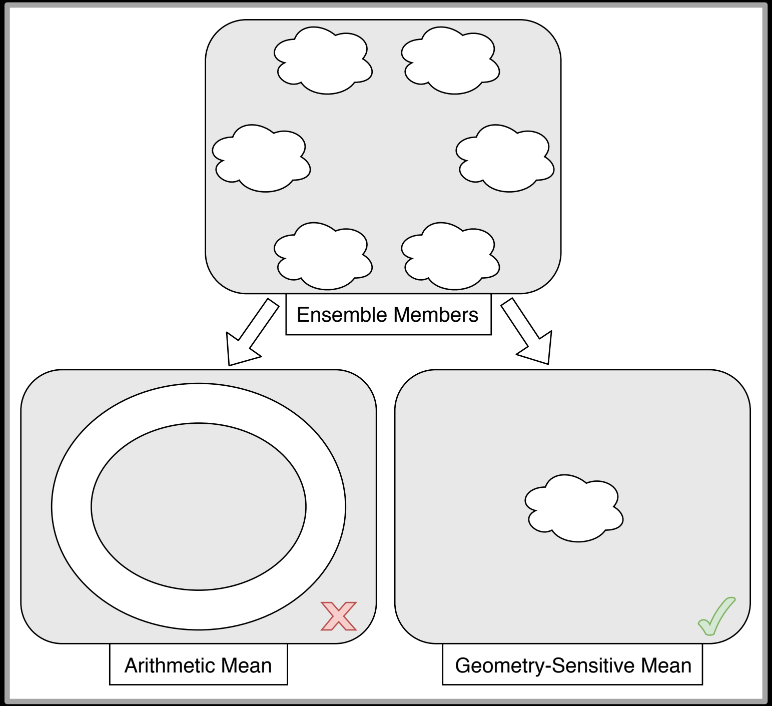
Introduction

- To combat information overload, a better method of visualizing ensemble mean images is needed
  - More readily interpretable than spaghetti diagrams, postage stamps, etc.
  - More representative than smeared-out pure arithmetic or weighted means

- New method: create 3 distinct types of ensemble mean image: in-sample, probability-based, and Gaussian-based
  - All 3 employ a calculated ensemble centroid to generate the final image

Introduction

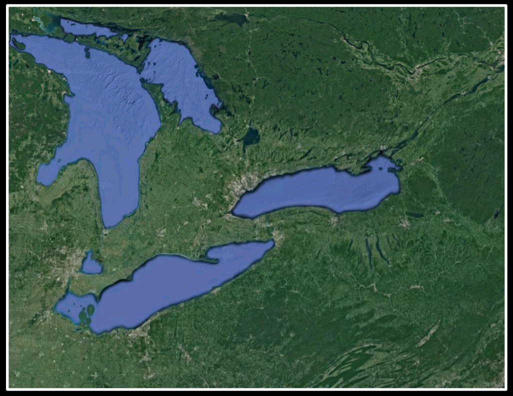
- To combat information overload, a better method of visualizing ensemble mean images is needed
  - More readily interpretable than spaghetti diagrams, postage stamps, etc.
  - More representative than smeared-out pure arithmetic or weighted means



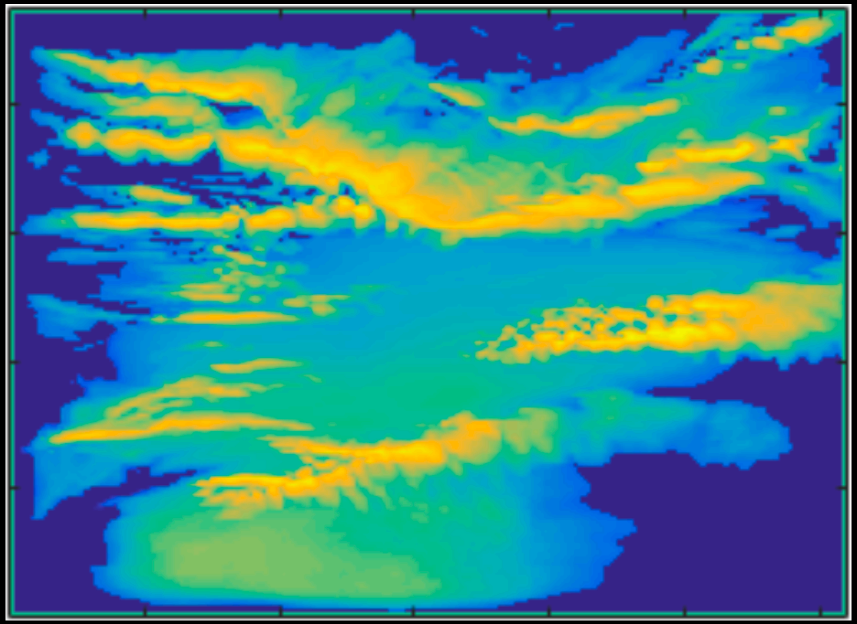
- New method: create 3 distinct types of ensemble mean image: in-sample, probability-based, and Gaussian-based
  - All 3 employ a calculated ensemble centroid to generate the final image

Source Imagery

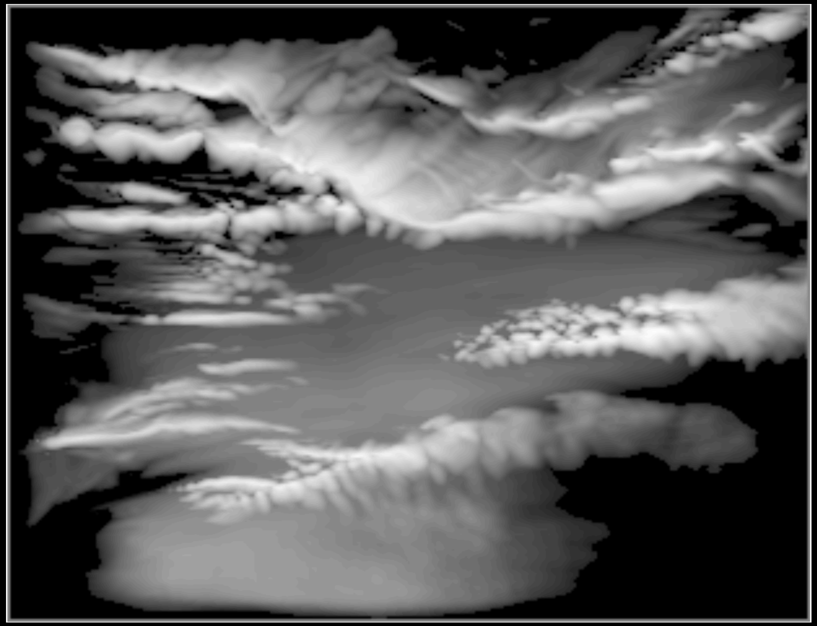
- Two ensembles of 30 WRF model-generated Lake Effect Snowband radar reflectivity images, converted to grayscale
- 3-km grid covering Lakes Huron, Erie, and Ontario
- Valid at 2013-12-11, 17 UTC



Google Earth

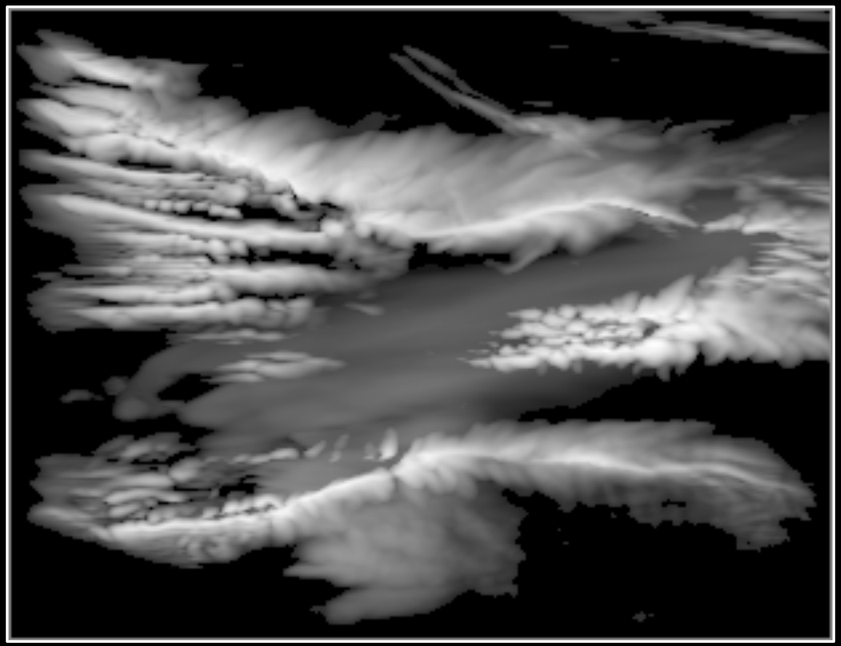


Z. Zhang 2018



Low-Intensity Thresholding

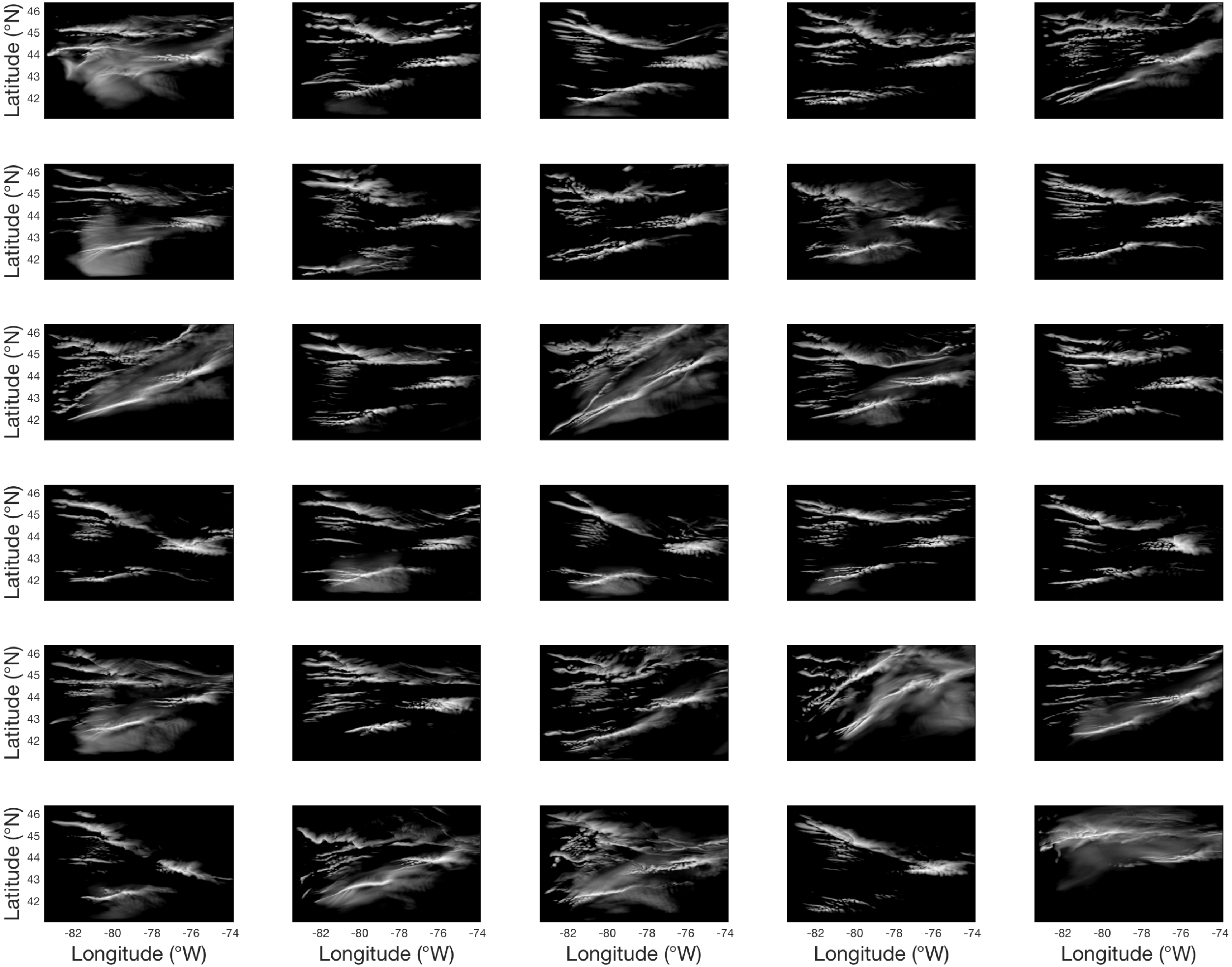
- Each source image in the ensemble is put through low-intensity thresholding, in which all regions of minimal reflectivity are set to zero.
  - The image is normalized by the ensemble maximum intensity (rather than image maximum) to preserve relative intensity relationships.
- Focuses on the high-precipitation cores of the model radar imagery, without being bogged down by areas of extremely light precipitation.



Original

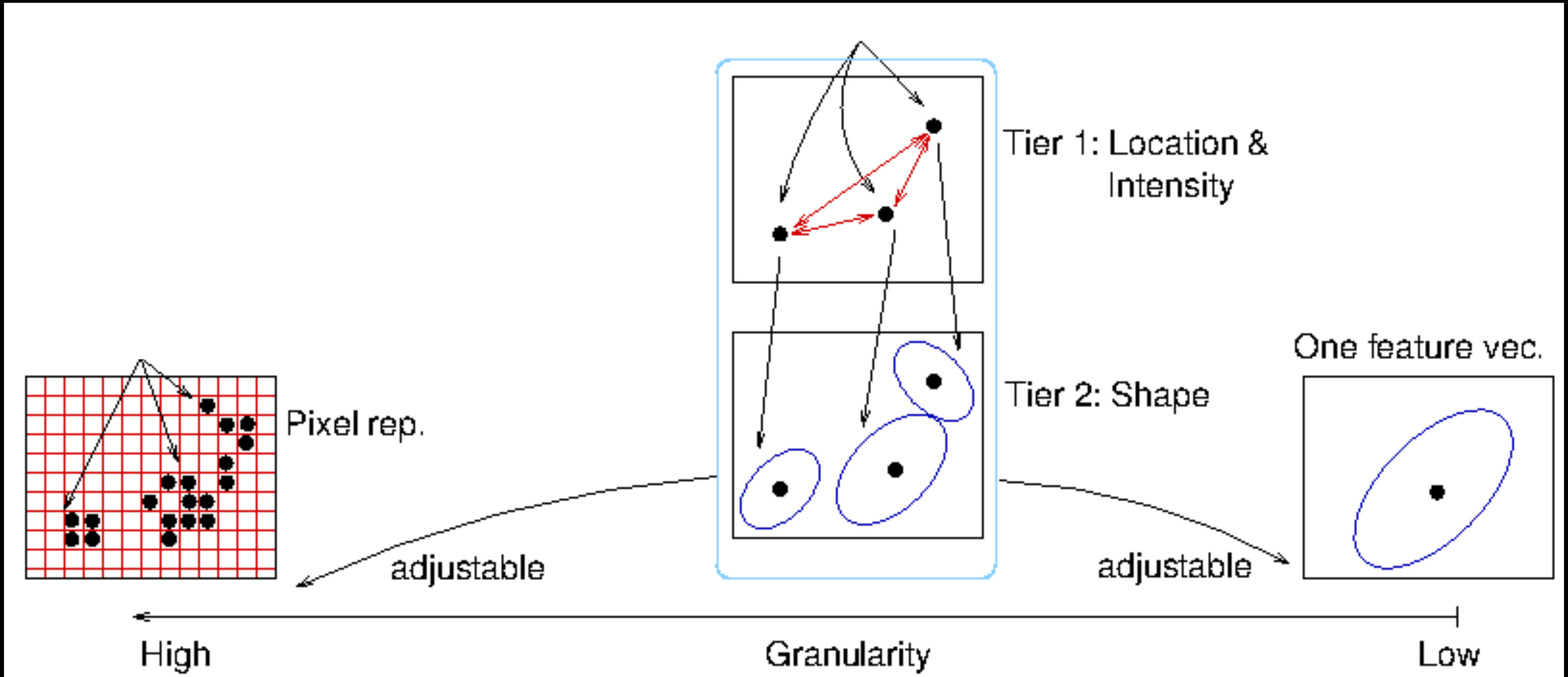
Thresholded

Source Image Ensemble: Thresholded



Two-Tier Cloud Signature (2TS)

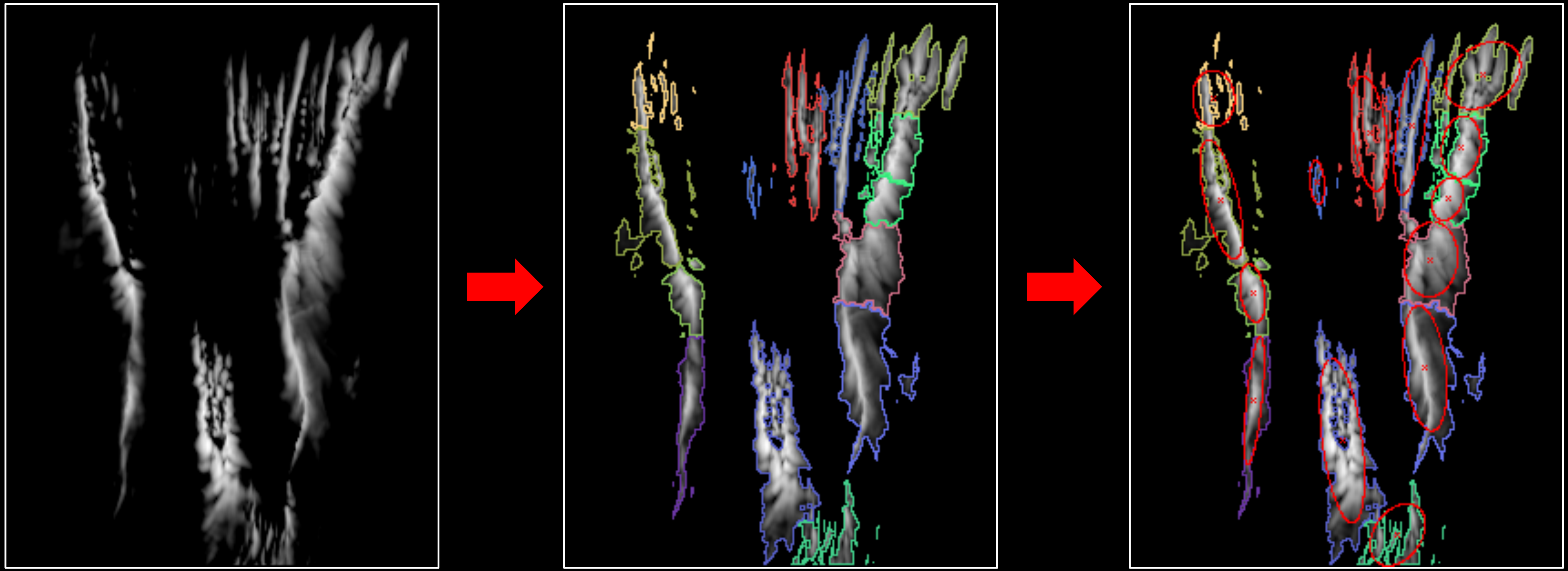
- Represents the geometric features of a reflectivity image as a set of “patches”:
  - 1st Tier: Coordinates for the centroid of each cloud patch, weighted by total patch intensity
  - 2nd Tier: Patch shapes, defined by the Gaussian distribution of the points in the patch



Li & Zhang 2018

Two-Tier Signature: Algorithm

1. Split pixels into small cells, cluster into cloud patches
2. For each patch, calculate total cloud intensity, weighted center, and weighted covariance matrix
    1. Generate 1st-tier (centroids) and 2nd-tier (Gaussian) signatures from those



Terminology

- **Support points** are the centroids of the calculated cloud patches- **support size** is how many there are: in essence, the granularity of the calculation.

- **Wasserstein distance (WD)**: The weighted sum of the distances between "support points" across all sets- the sum of the products of the distances and the points' differences in value (AKA Earth-Mover's Distance).

- **Barycenter:** Effective mean for Euclidean vectors- minimizes total squared Euclidean distance.

- **Wasserstein Barycenter (WB)**: The image (or distribution) such that the total WD to each ensemble member is minimized: the effective mean of the ensemble.

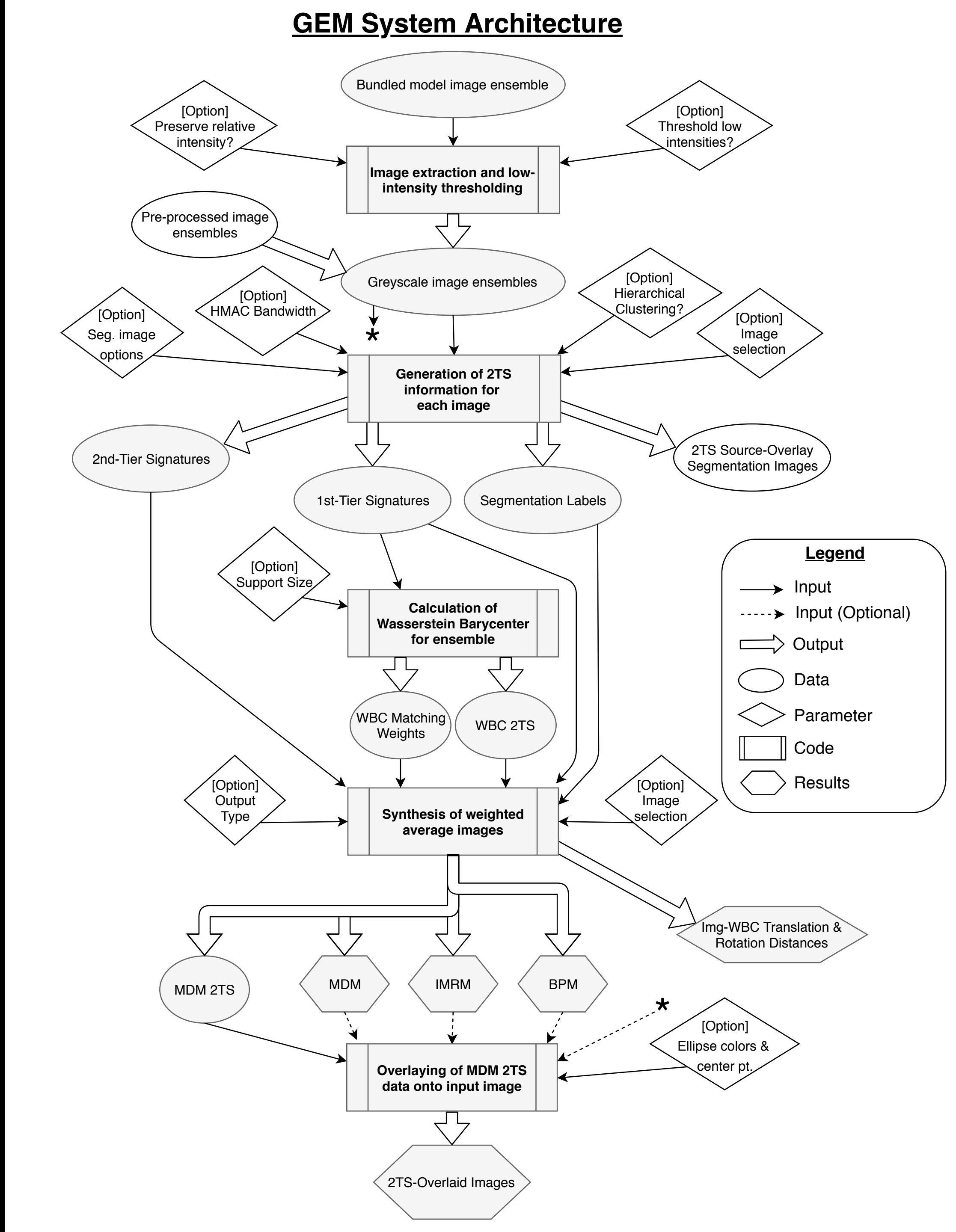
GEM Ensemble Mean Image Types

- **Mixture Density Mean (MDM) Image** A more physical representation of the 2TS, scaled by intesity totals and covariance. (“Gaussian-based”)

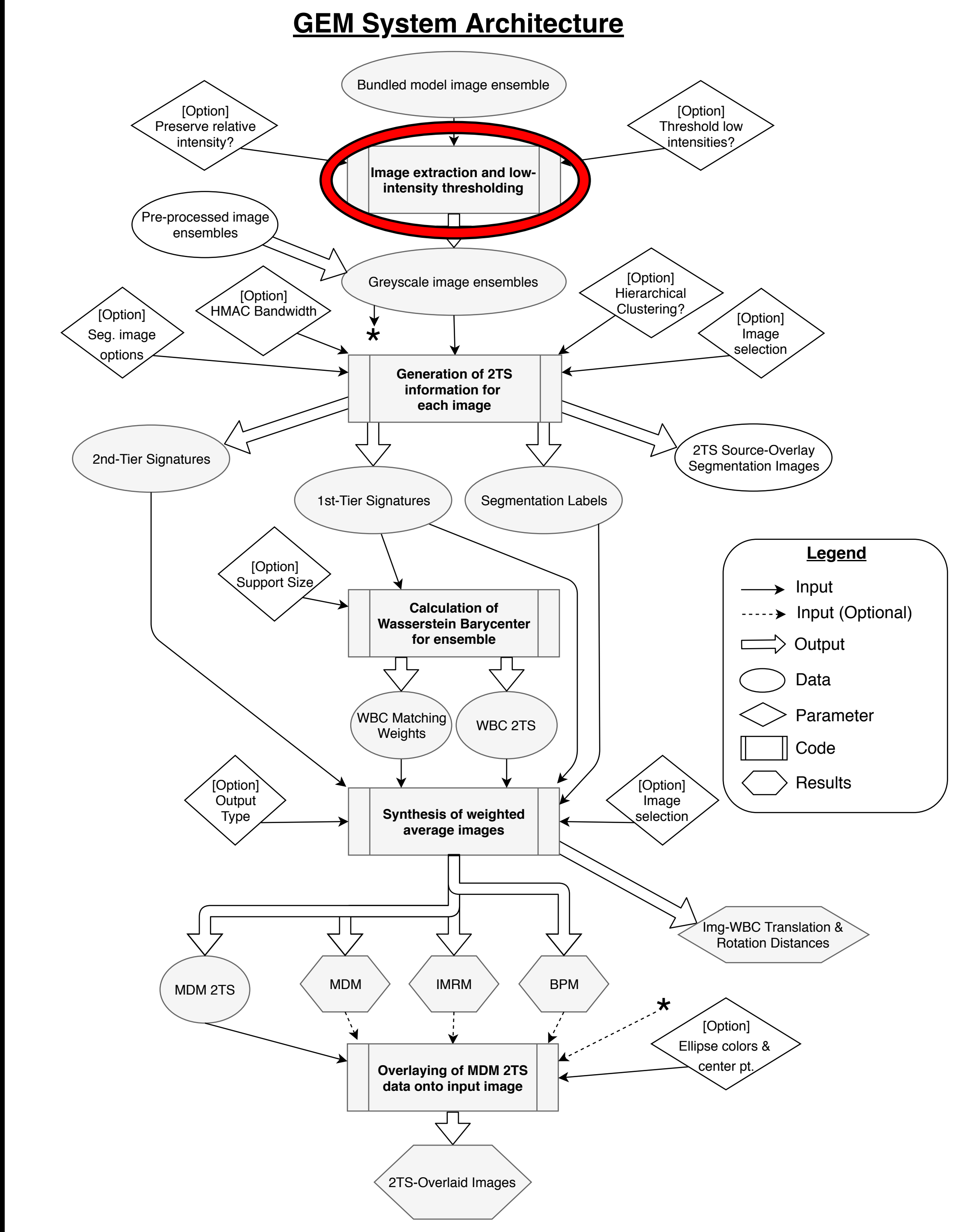
- **Bayesian Posterior Mean (BPM) Image-** Estimate ”true” image by treating forecasts as random samples for the Bayesian posterior mean. (“Probability-based”)

- **In-sample Mean under Rigid Motion (IMRM) Image-** In-sample mean with ensemble members translated and rotated to best alignment with the MDM.

Methodology

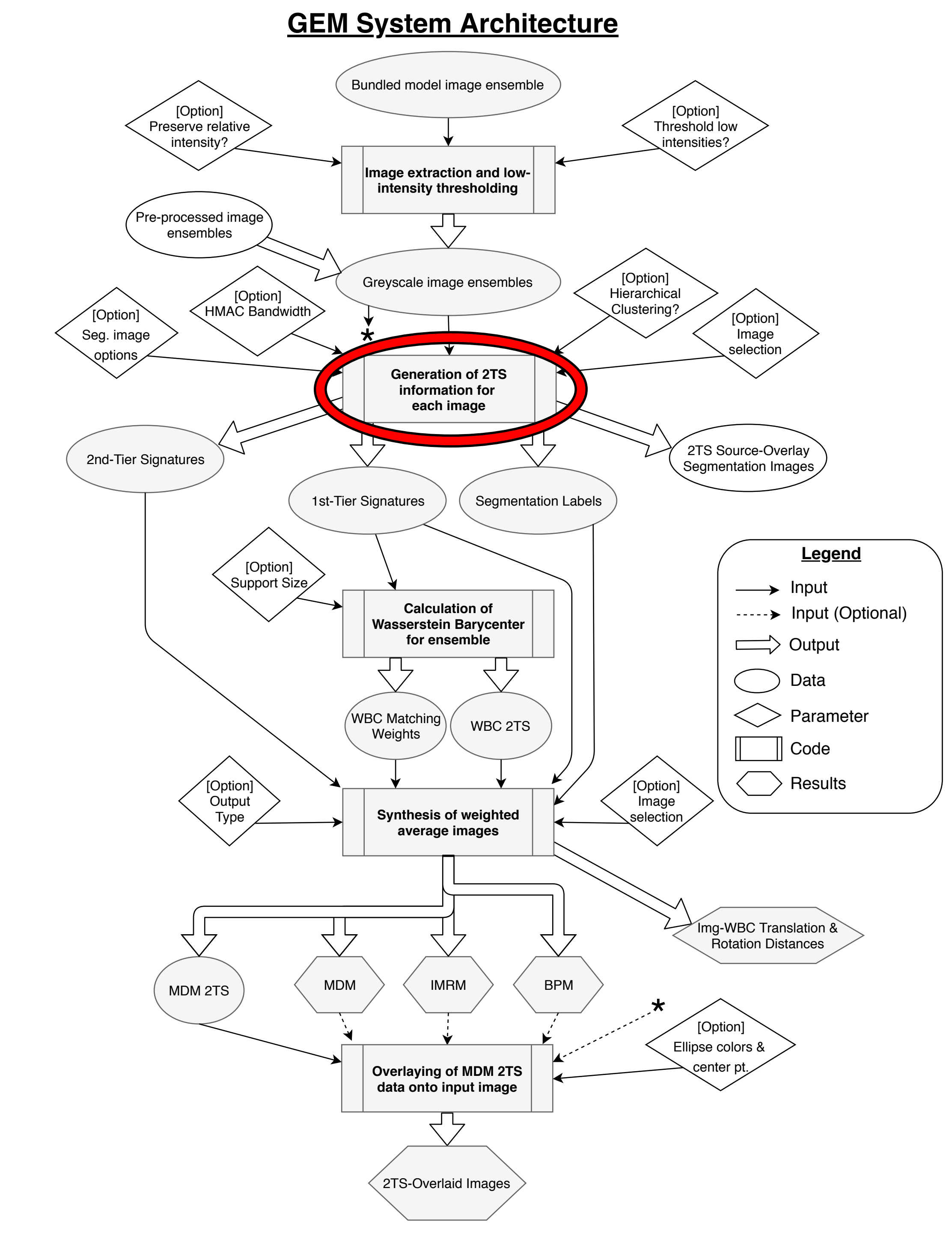


[All processing done in the C and Matlab programming languages.]



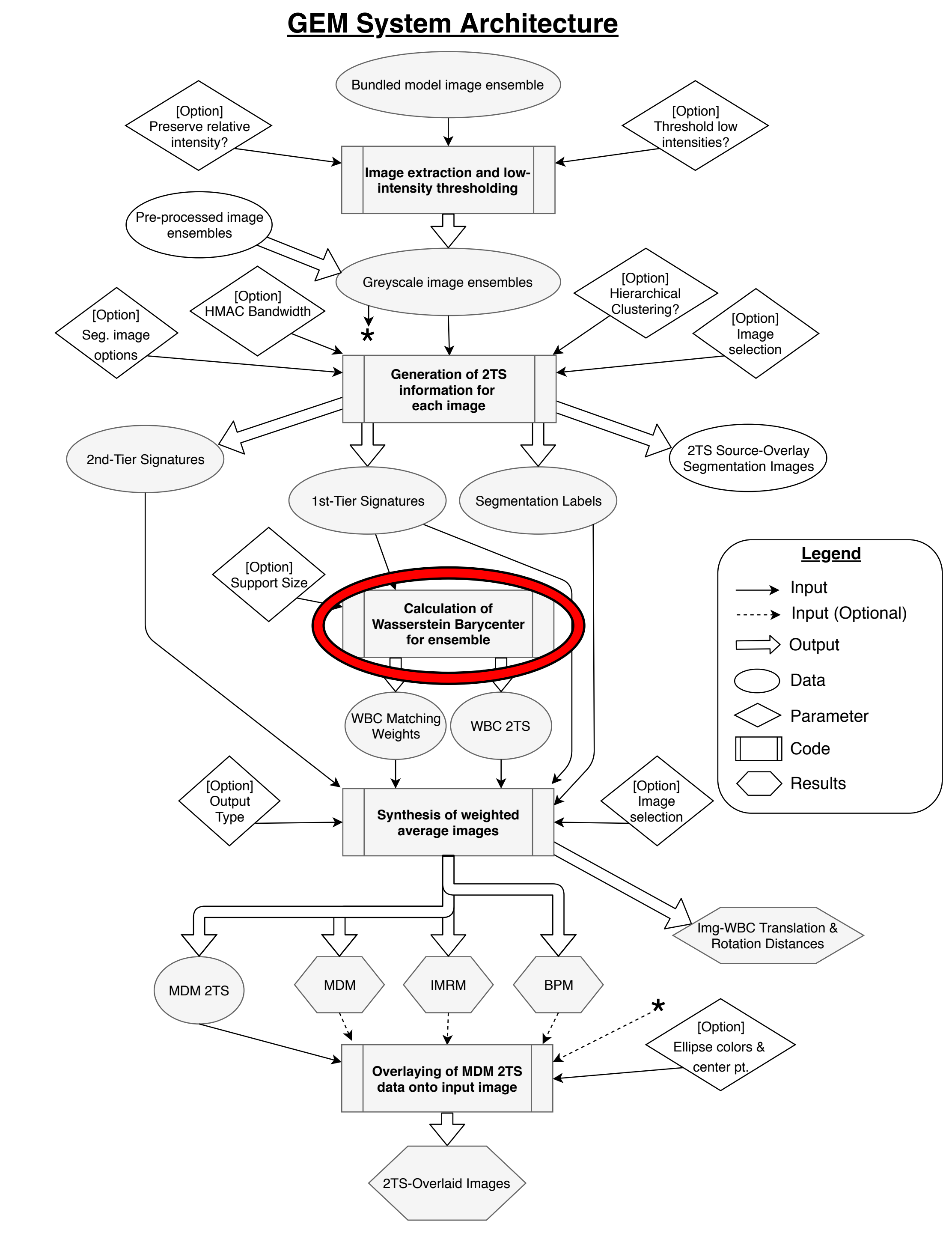
Methodology

1. Generate thresholded versions of each ensemble member.



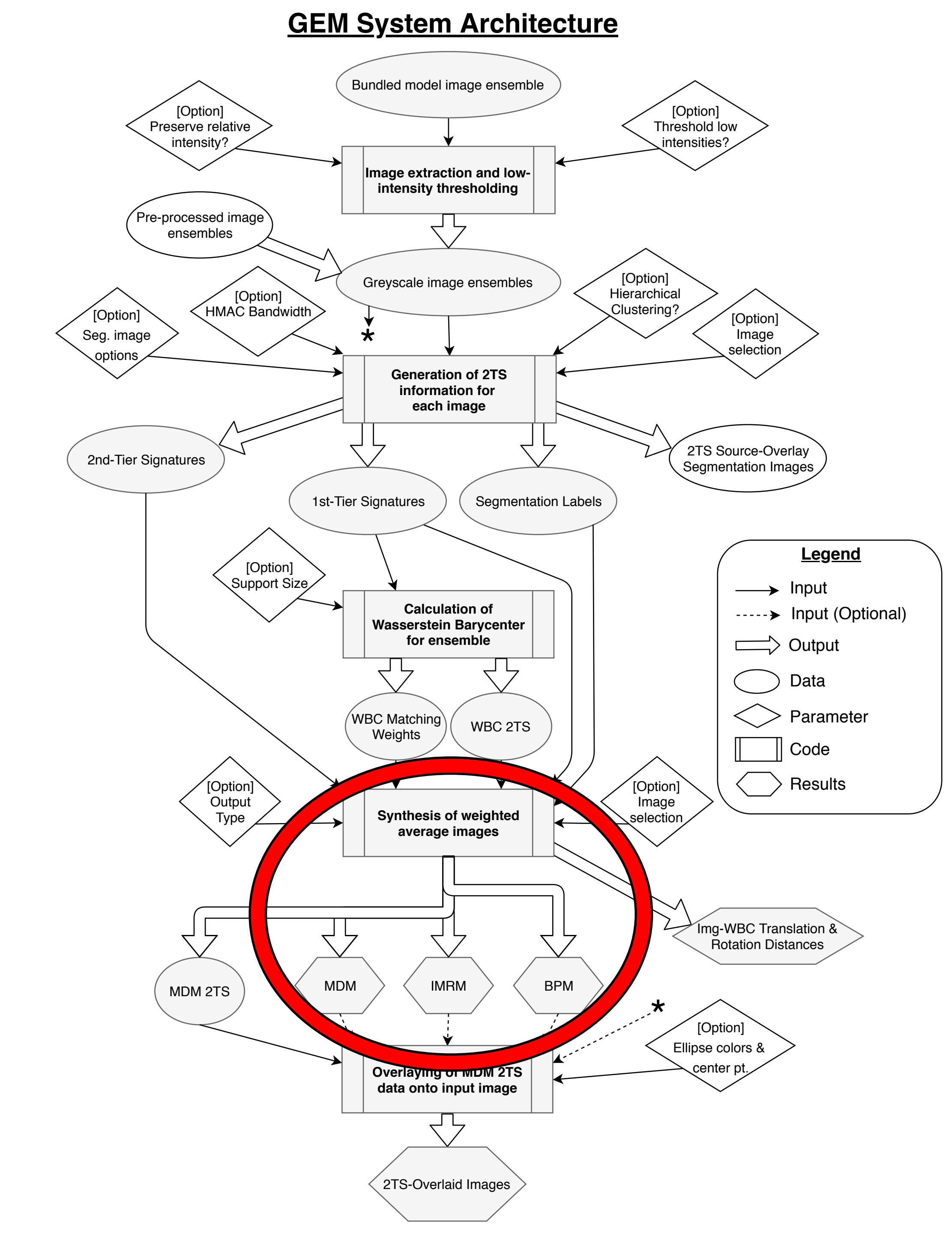
Methodology

2. Calculate 2TS for each ensemble member at 4 distinct bandwidths
   - Generate cloud patch overlays



Methodology

3. Calculate WB for each bandwidth at 4 distinct support sizes
   - Total of 16 granularity levels



Methodology

4. Synthesize BPM, MDM, IMRM, and pixel-wise average (PWA) ensemble mean images

Parameter Examples

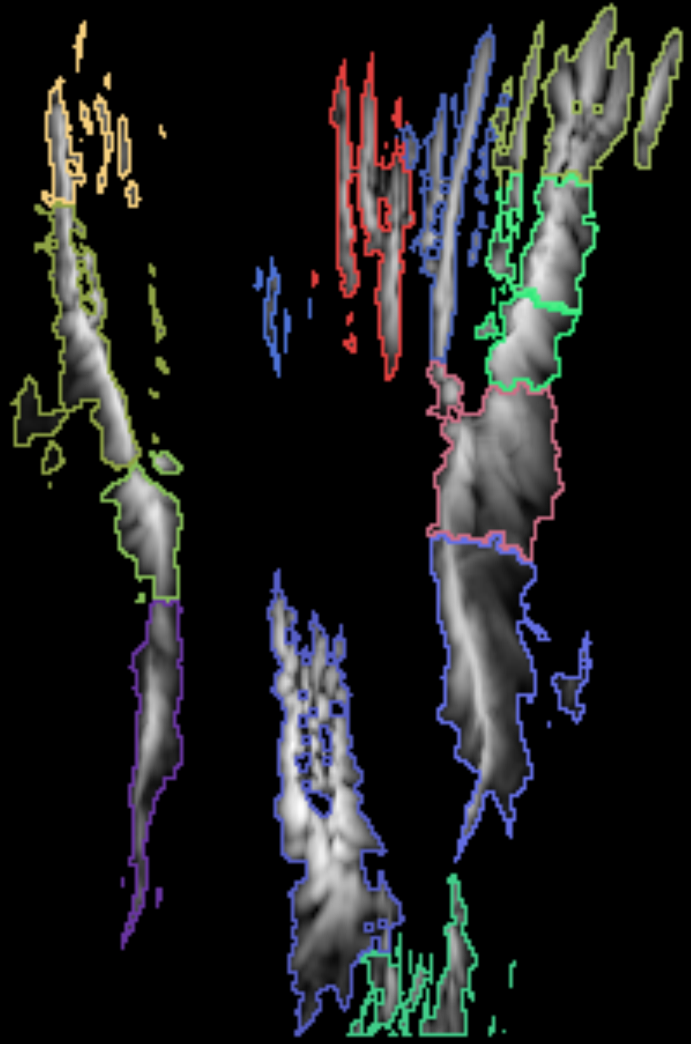
Bandwidth

Ensemble member #10: Thresholded

Bandwidth A = (2, 3.5)

Bandwidth values ↑

Fewer patches,
each are larger



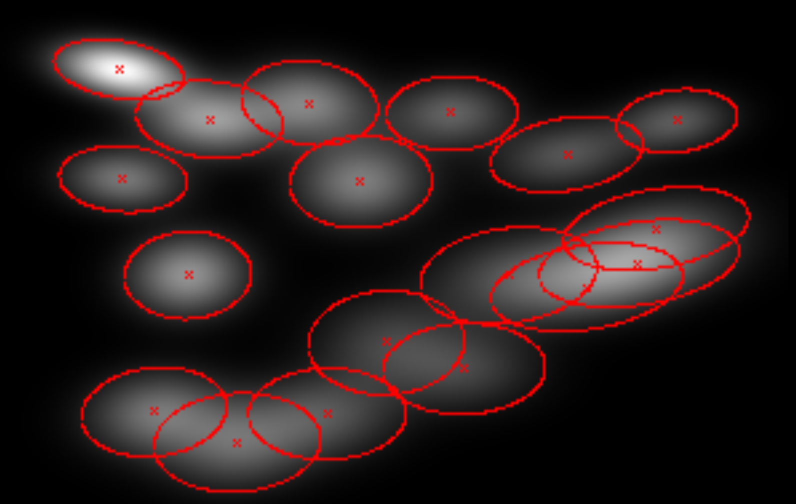
Bandwidth C = (5, 8)

Support Size

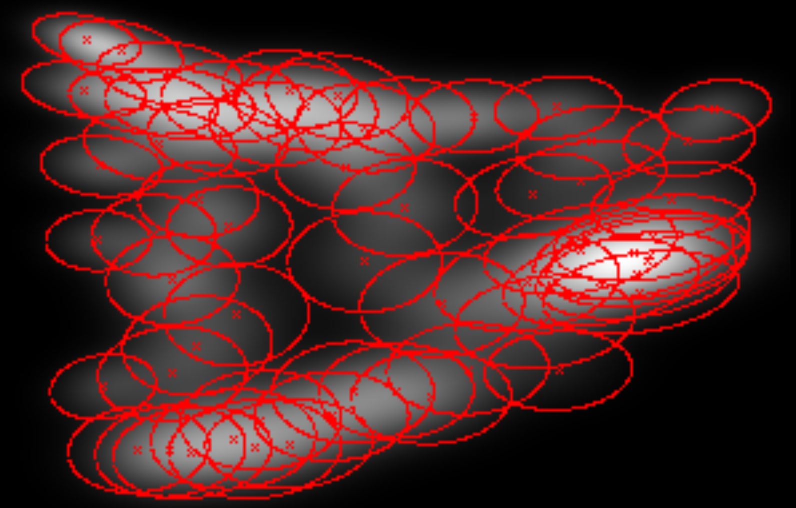
MDM:
Bandwidth B (3,5),
Thresholded

Support size ↑

More reference points,
higher granularity



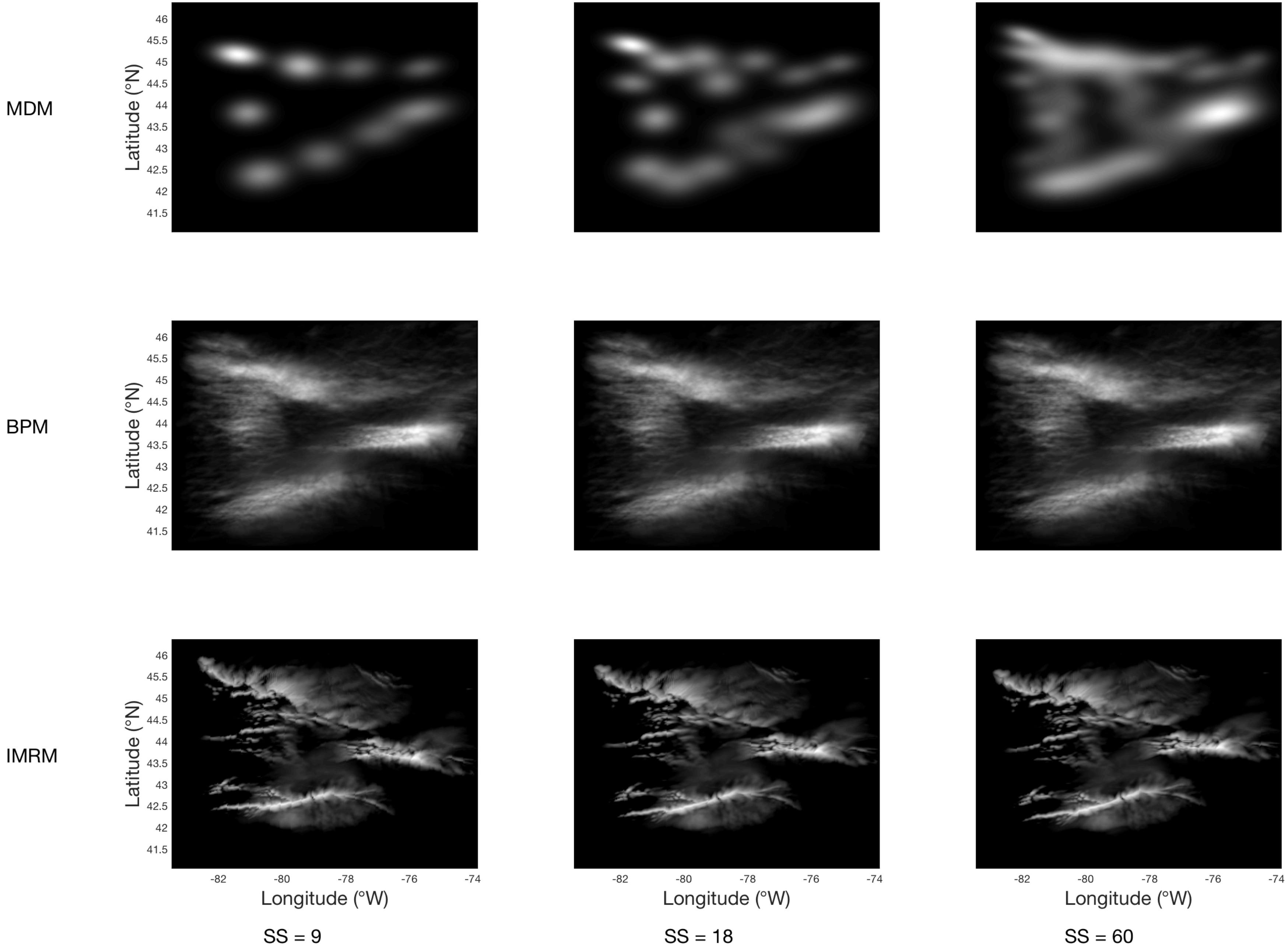
Support Size 18



Support Size 60

Results

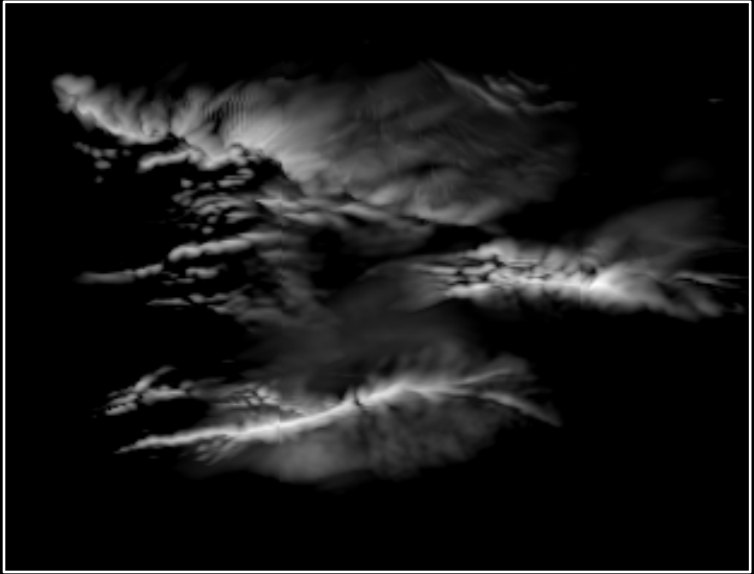
WB-Ensemble-Avg Images: Bandwidth B, Thresholded



GEM Outputs Vs. Pixelwise Average

Bandwidth B (3,5)
Support Size 18

PWA



IMRM

MDM

BPM

Preliminary Conclusions & Future Work

- Preliminary Conclusions:
  - All mean image variants retain more information than the PWA, and will be useful for different applications
  - The MDM shows the greatest sensitivity to granularity and thresholding
    - Provides a useful visual representation of the 2TS, and might indicate locations of interest ("intensity kernels") within the snowbands
  - The BPM shows a more reasonable spread of possibilities
  - The IMRM provides a realistic best member
  - Thresholded results are more representative of the primary snowband structures

- Future work:
  - Larger ensemble, more cases
  - Determine best method of system calibration (granularity limits) for a given case
  - Develop more direct meteorological applications and demonstrative examples

References

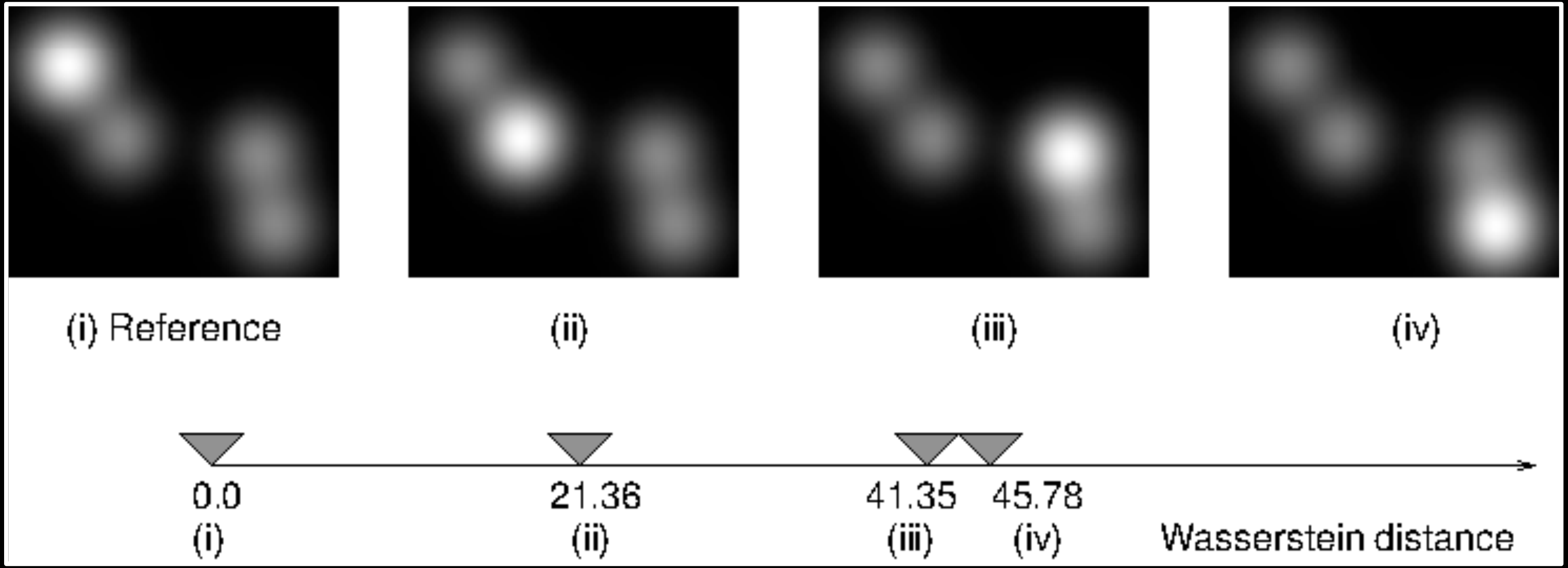
1. Li, Jia and Fuqing Zhang. "Geometry-Sensitive Ensemble Mean based on Wasserstein Barycenters: Proof-of-Concept on Cloud Simulations", *Journal of Computational and Graphical Statistics*, 2018. DOI: 10.1080/10618600.2018.1448831
2. Zhuomin Zhang. "Geometry-Sensitive Ensemble Mean based on Wasserstein Barycenters: Proof-of-Concept on model-simulated weather radar reflectivity", 2018.
3. Ng, Sam. "Satellite image of Lake-Effect Multiple Snow Bands event", *Weather 5280*, 25 Nov. 2014, https://www.weather5280.com/blog/2014/11/25/lake-effect-snow-explained/

Questions?

Extra Slides

Wasserstein Distance Example

- The brightness of each support point region represents the intensity value assigned to it.
- All WD calculated with respect to the reference distribution, A.
- The WD for each differs significantly, as they consider both the intensity distributions and the Euclidean distance between the support points.



Li & Zhang 2018