High Resolution Catastrophe Modeling using CUDA

Dag Lohmann, Stefan Eppert, Guy Morrow
KatRisk LLC, Berkeley, CA

http://www.katrisk.com

March 2014, Nvidia GTC Conference, San Jose

Acknowledgements

- This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

KatRisk

About us and this talk

- We are a company (started August 2012 in US and UK) working on global catastrophe models
- More details can be found on http://www.katrisk.com
- Open source example code in R “How to build a catastrophe model” is also on the website
- If you are interested in working together with us send me an email: Dag.Lohmann@KatRisk.com

KatRisk

Agenda

- What problem are we trying to solve?
- What are catastrophe models?
- Using CUDA for high resolution flood risk maps
- Next Steps: Event sets
- Closing thoughts
- Questions

Natural Disasters Examples

Insurance Needs Simplified

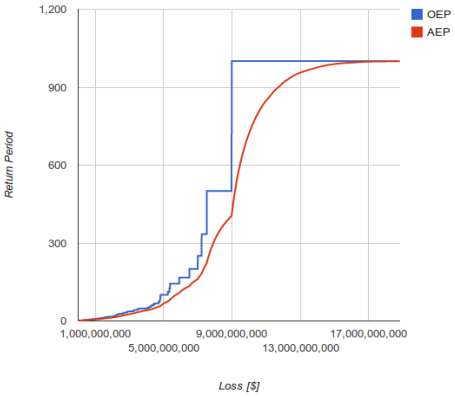
- Two basic business processes:
  - Underwriting
  - Portfolio Management
- Underwriting needs information about local risk at the location level (if that is available – otherwise needs good idea about area)
- Portfolio Management needs information about correlation between locations and aggregation of risks

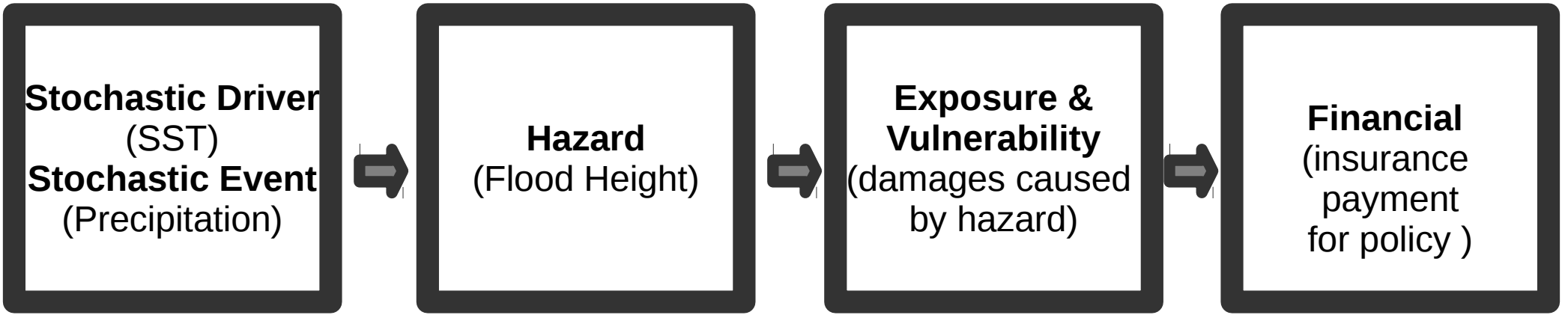
What is catastrophe modeling?

- Catastrophe: an event causing great and often sudden damage or suffering
- Why model catastrophes?
  - Because catastrophes are rare events that cause great harm
  - Build resilient societies where risk adverse behavior is rewarded
  - Actuarial pricing is based on loss experience (which is too short) or simplified models of nature
    - Fit frequency and severity distribution to past claims
    - Calculate loss distribution
    - Use for pricing, portfolio management, reinsurance

Structure of a Catastrophe Model

- Catastrophe models provide
  - Catalog of synthetic events with occurrence probability
  - Damages and insured losses to portfolio
  - Actuarial statistics (AAL, EP-curves, ...)
  - Where do historic events fall on the EP curve





Flood Model Example

Information in a Catastrophe Model

- Cat Models extend a company's loss experience with a synthetic event set
  - Fixed set of unobserved but realistic events
  - Calculate hazard intensity for all exposed locations
    - Locations are described by building characteristics
  - Calculate resulting damage (P&C)
  - Apply financial structures to model output
    - e.g. home owner insurance contract or reinsurance treaty
  - Output typically is an Event Loss Table
    - From which we can then derive all analytics, e.g. loss by postcode, loss by line of business (LOB), EP, etc.

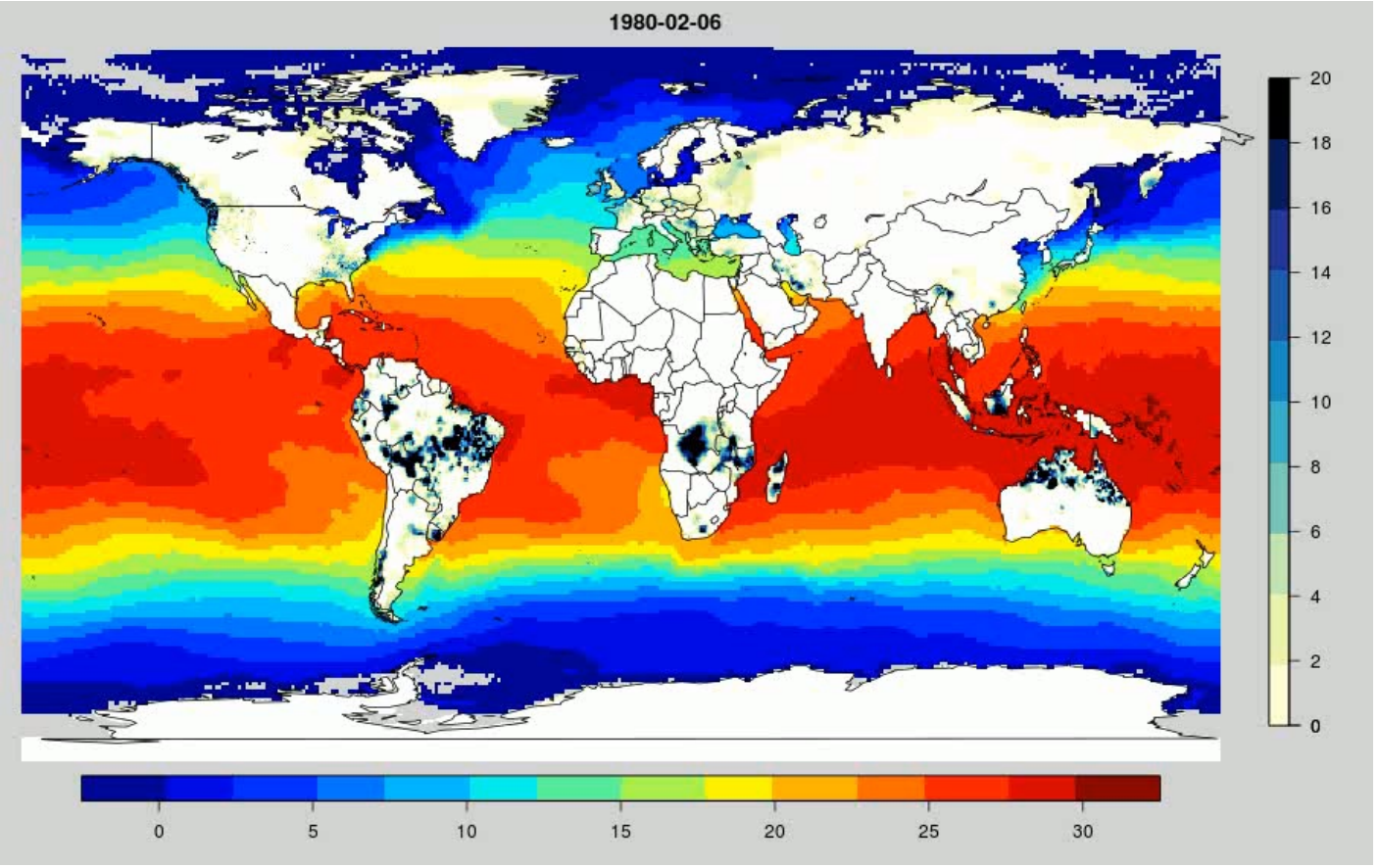
KatRisk Water Modeling chain

Precipitation conditional On SST

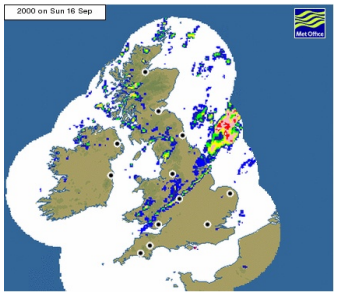
Rainfall-Runoff-Model

River Routing

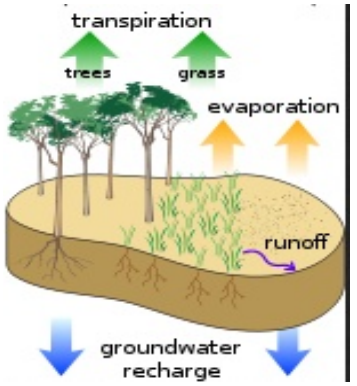
Inundation



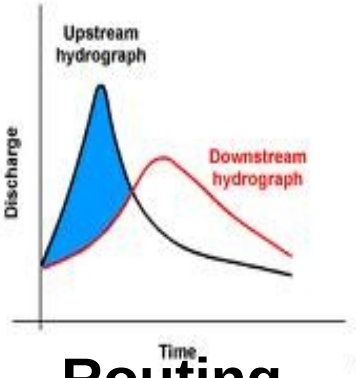
Hazard Model Components



Precipitation +Meteorology



Runoff



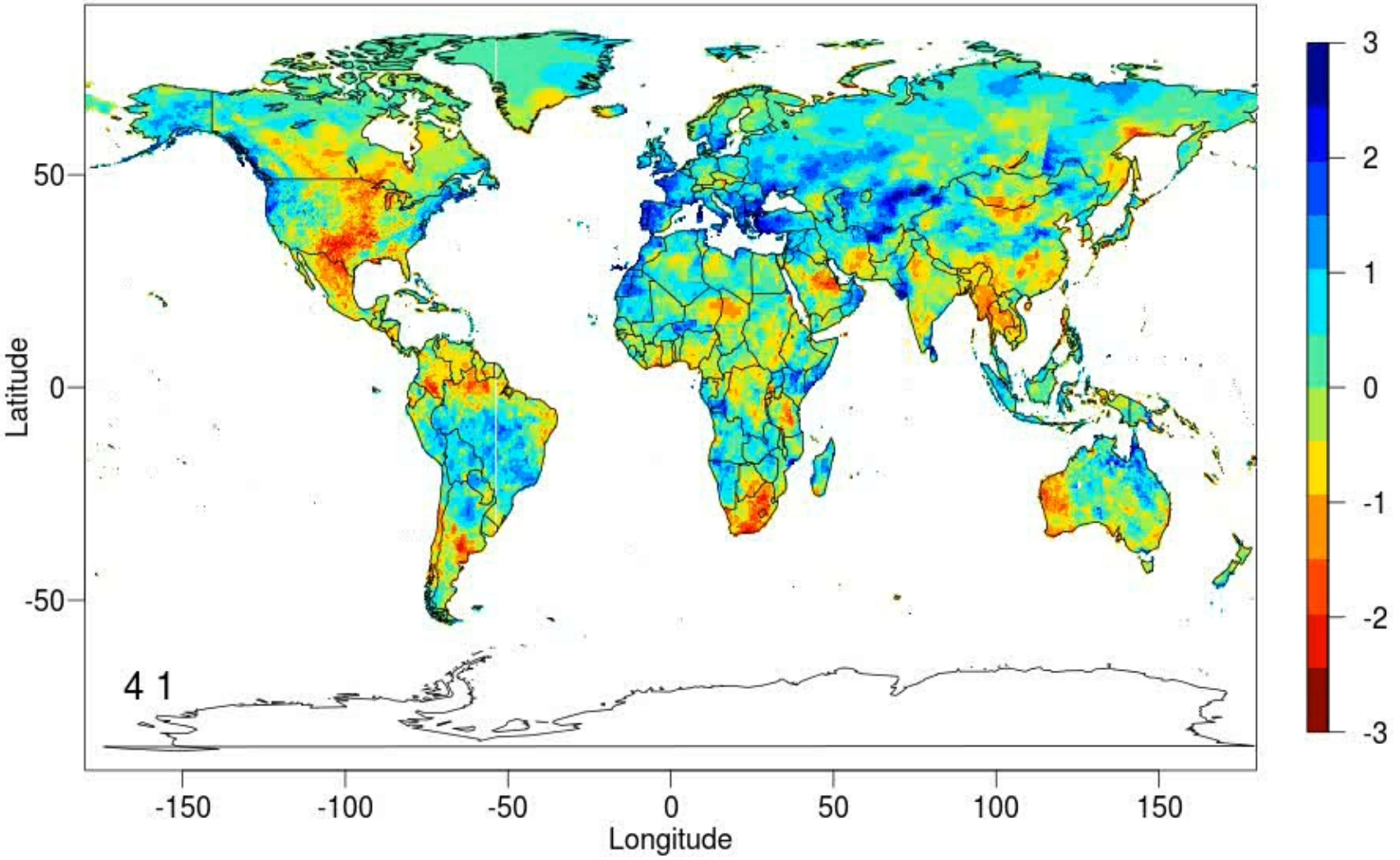
Routing

Inundation

Component	Output	SE Asia data volume (15,000,000 km^2)
Stochastic Precipitation	monthly precipitation anomalies on 0.25 deg grid	27,000 cells
Downscaling	3-hourly precipitation per catchment	40,000 catchments
Rainfall-Runoff Model	3 h runoff, saturated areas, groundwater	
Routing	10 min river discharge at outlets	
Inundation	Flood-depth at each cell	17b cells

Global Flood & Drought Model

- Normalized Surface Water (monthly time step shown)



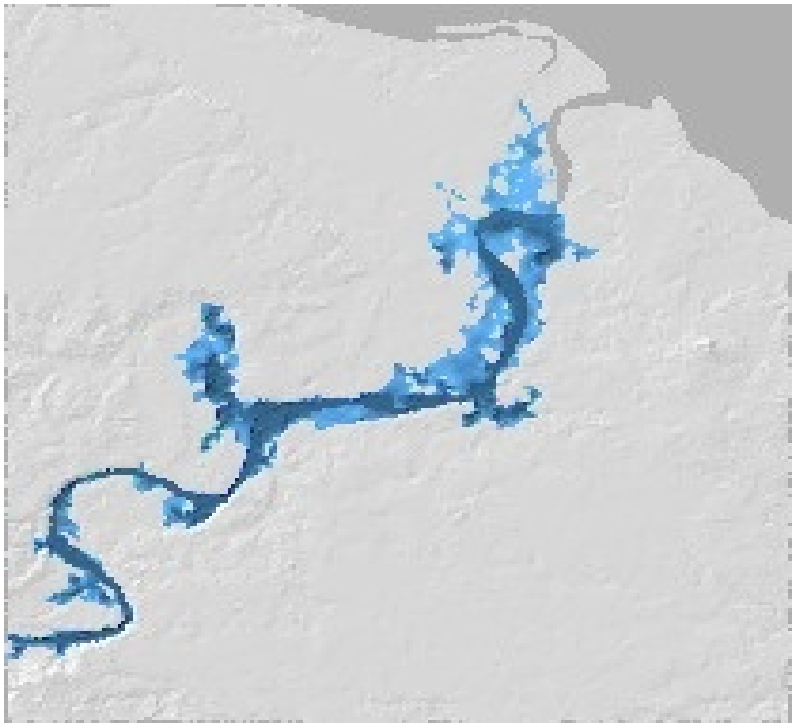
Inundation Model

- 2D Shallow Water Equations
- Riverine Flooding, dam break, Tsunami

$$h_t + (hu)_x + (hu)_y = 0,$$

$$(hu)_t + \left(hu^2 + \frac{1}{2} gh^2\right)_x + (huv)_y = -ghB_x,$$

$$(hv)_t + \left(hv^2 + \frac{1}{2} gh^2\right)_y + (huv)_x = -ghB_y.$$



- High resolution, small time steps → CUDA

SWE on CUDA

- Explicit upwind scheme, regular grid
- State variables *h*, *hu* and *hv*
- Kernels:
  - Set boundary conditions
  - Calculate flux
  - Get maximum time step (depending on celerity)
  - Integrate over time step
- Overlapping computational domain per block with ghost cells on edges

MPI scheduling

- Each GPU calculates flooding along a river stretch (“catchment”)
- Use MPI to distribute work to GPUs
- US: ca. 15000 catchments, 15m 10m cells per catchment, single run takes 1 to 24 GPU hours
- World: ca. 500K catchments, 90m resolution, single run takes 30s to 10min per GPU

TITAN supercomputer (ORNL)
#2 TOP 500
17,590 Tflop LINPACK
18,688 NVIDIA K20 GPUs
18,688 16-core Opteron CPUs

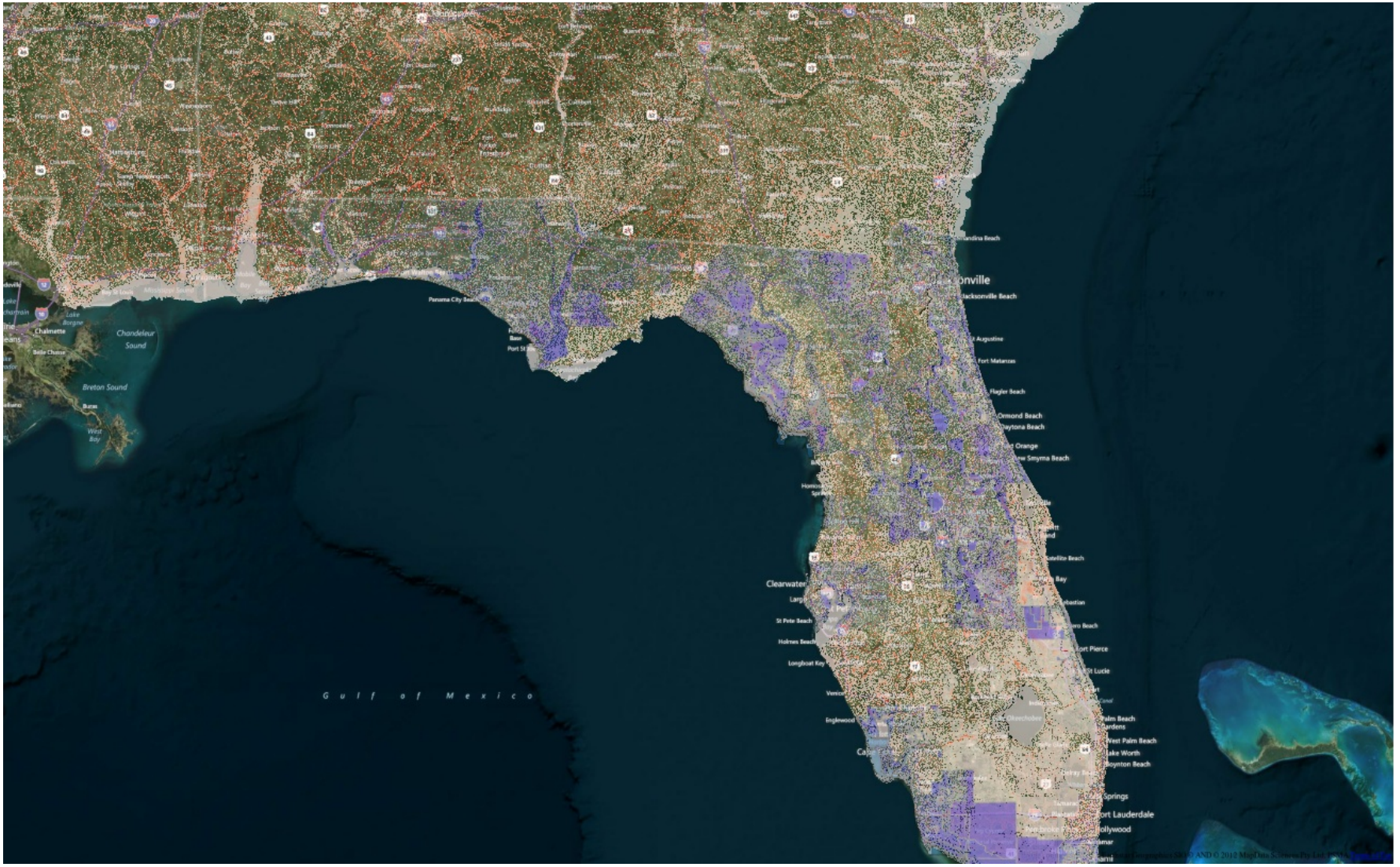
Benefits of GPU code

- GPU 50 – 200x faster than single-core CPU (depending on catchment)
- ORNL TITAN compute capability allows us to calculate global 90m and US 10m resolution flood maps with accurate 2D flooding for first time
- Better spatial differentiation of risk to minimize damages from future catastrophic events

USA Flood Maps

USA Flood Map on 10m resolution

Zoom into Florida



KatRisk and FEMA data shown. KatRisk red, FEMA blue

Something about color scales

Red = KatRisk only, no FEMA

White = Katrisk only, low flood depth

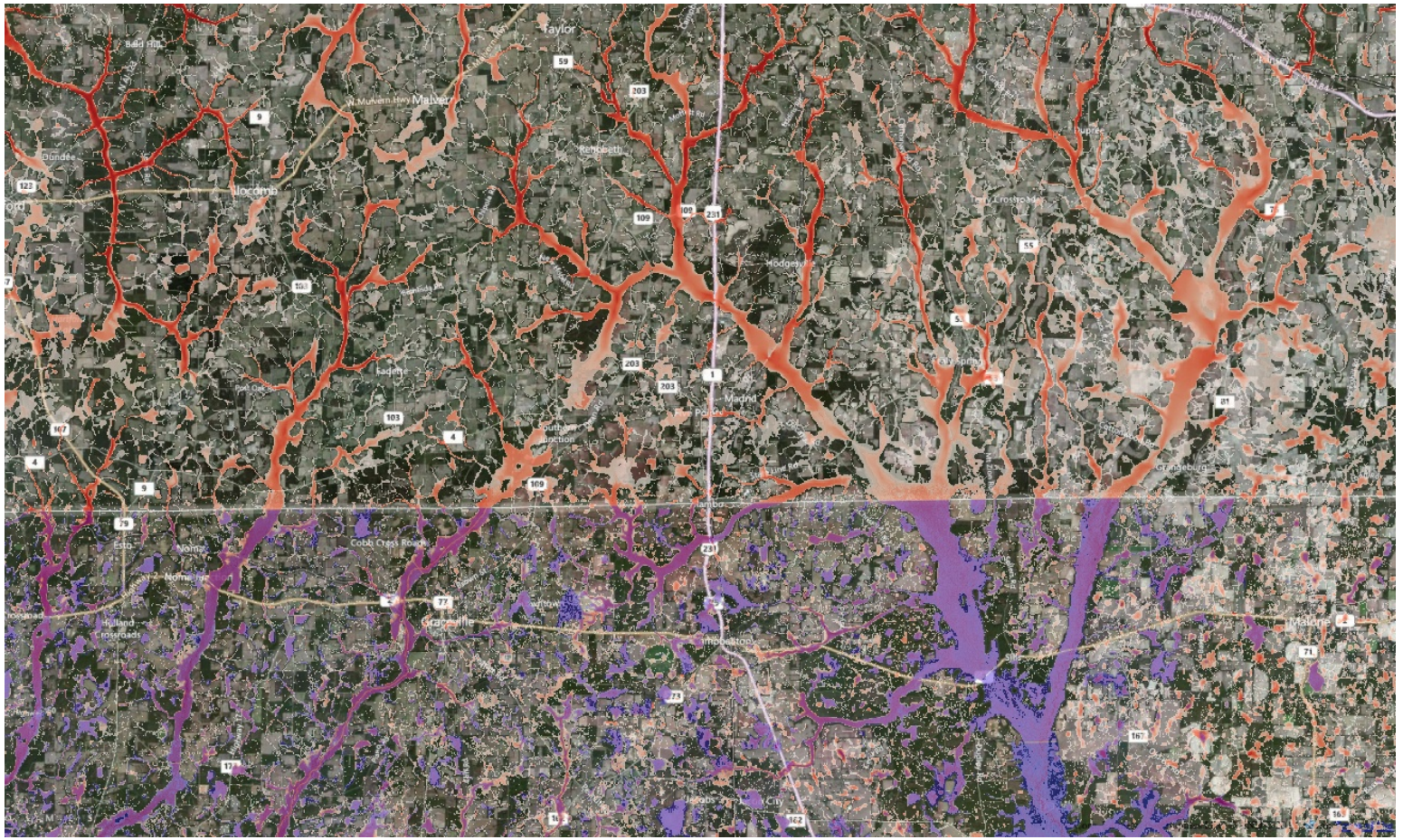
Purple = FEMA and KatRisk overlap

Blue = FEMA > KatRisk

Flood Maps Zoom

Red = KatRisk only, no FEMA

White = Katrisk only, low flood depth



Purple = FEMA and KatRisk overlap

Blue = FEMA > KatRisk

Flood Maps Zoom

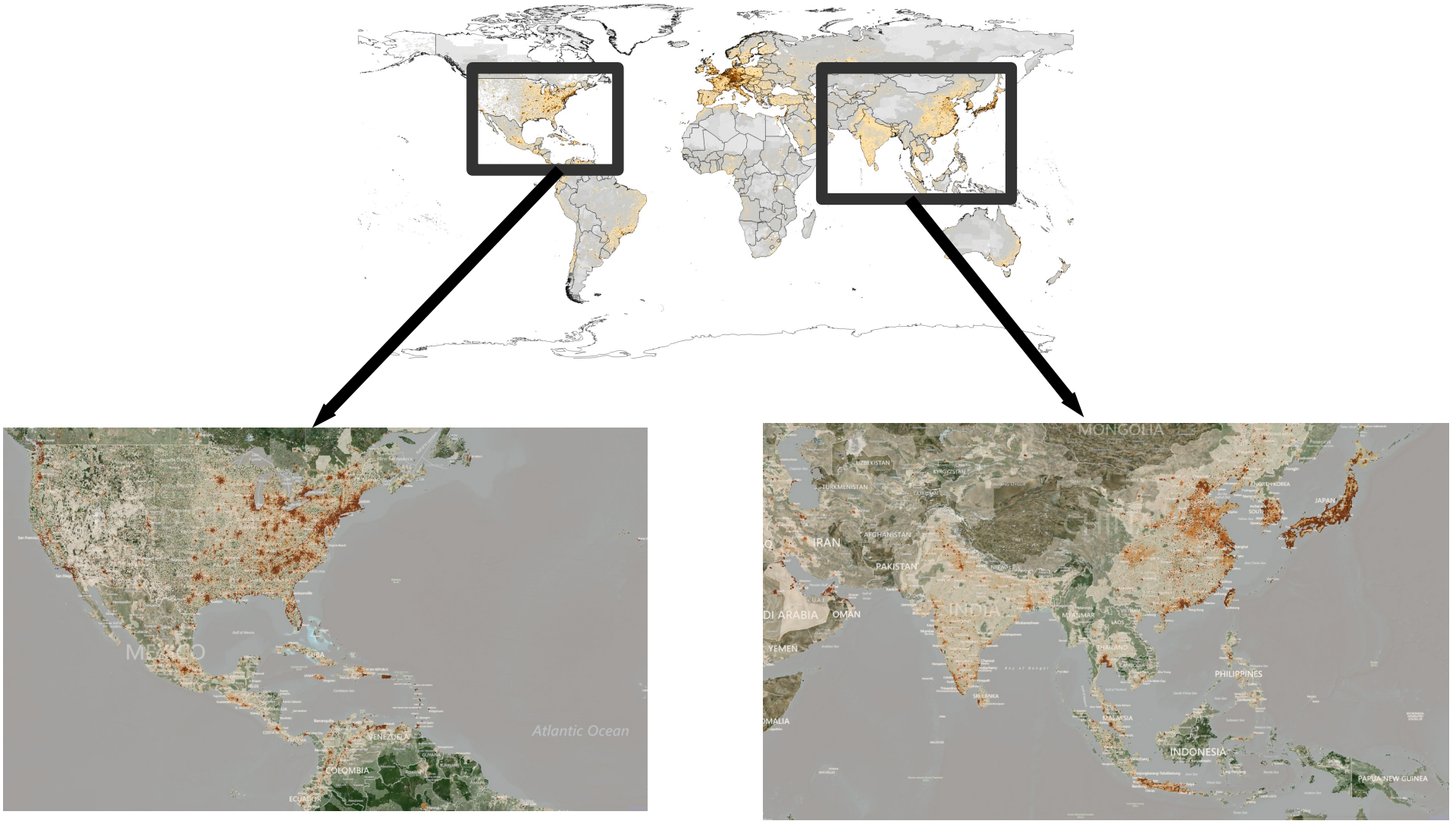
Red = KatRisk only, no FEMA

White = Katrisk only, low flood depth

Purple = FEMA and KatRisk overlap

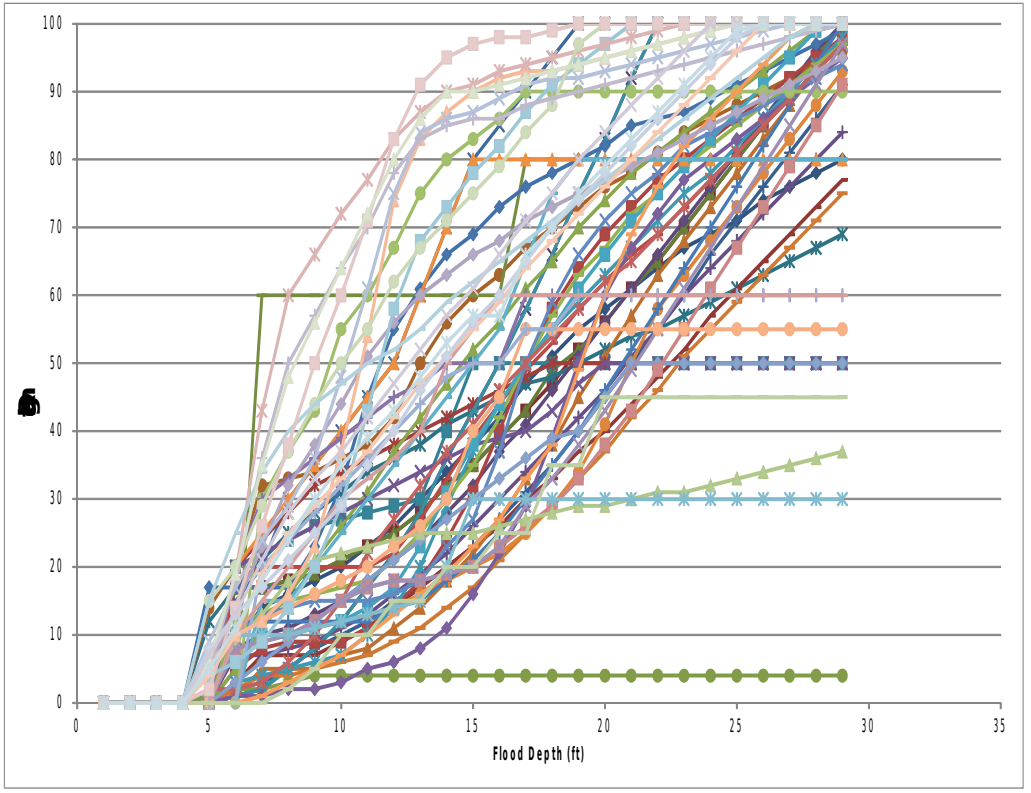
Blue = FEMA > KatRisk

Global Exposure – Risk Modeling

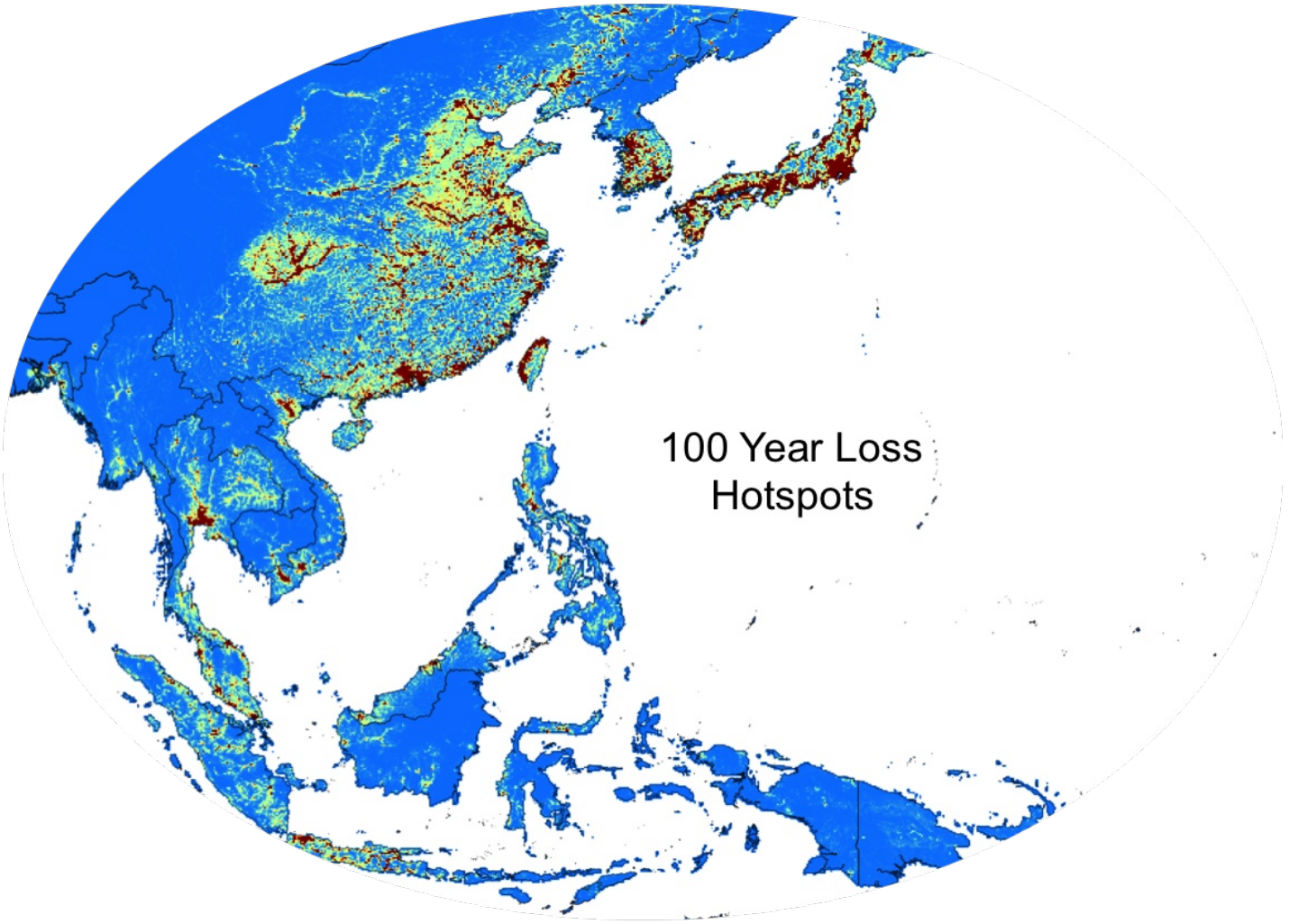


Global Vulnerability – What's my Loss?

- Published Curves
  - US Army Corps of Engineers
  - Flood Hazard Research Centre (Middlesex)
  - Research papers
- Loss data analysis
- Building surveys
- Component-based approach

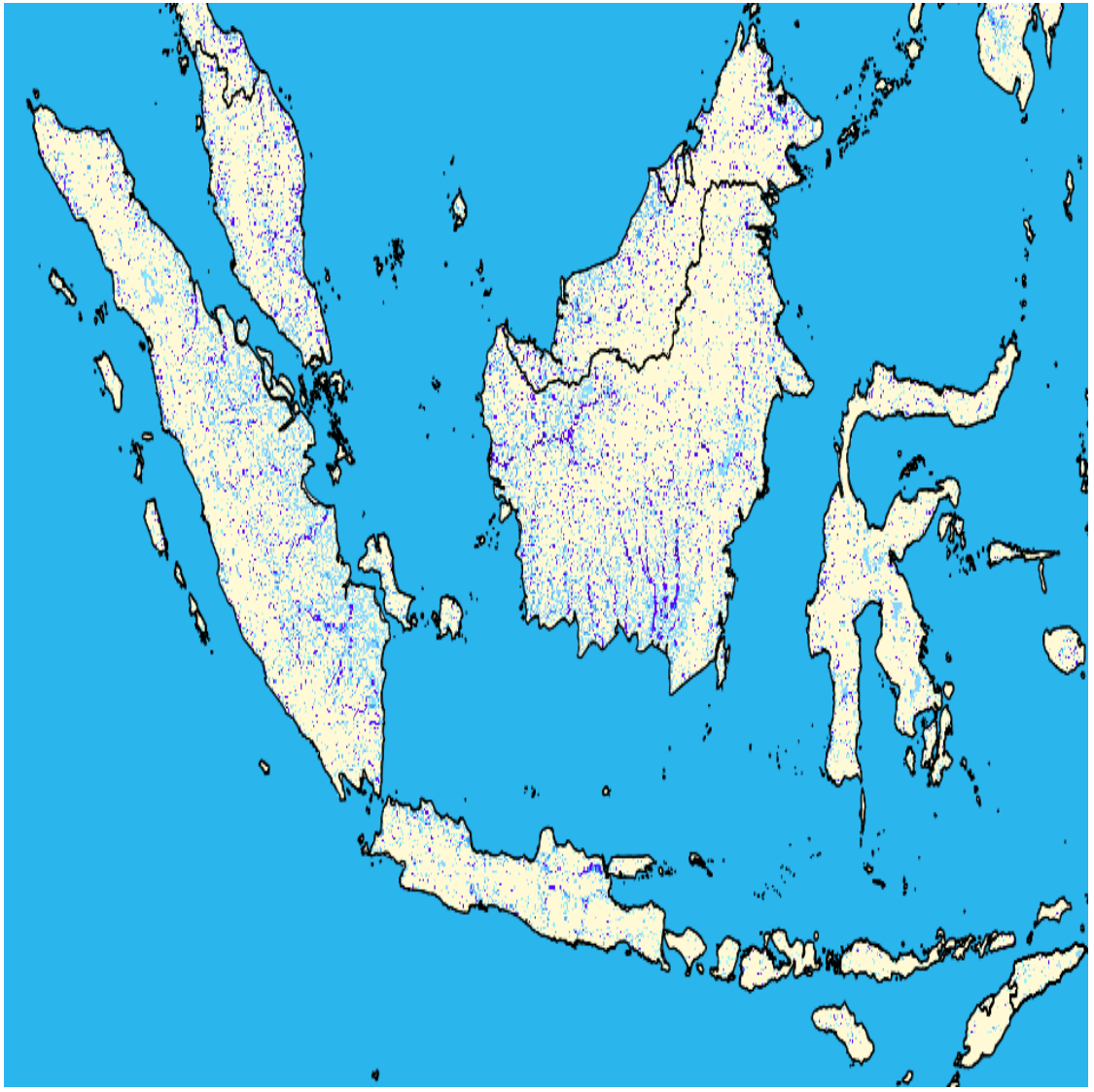


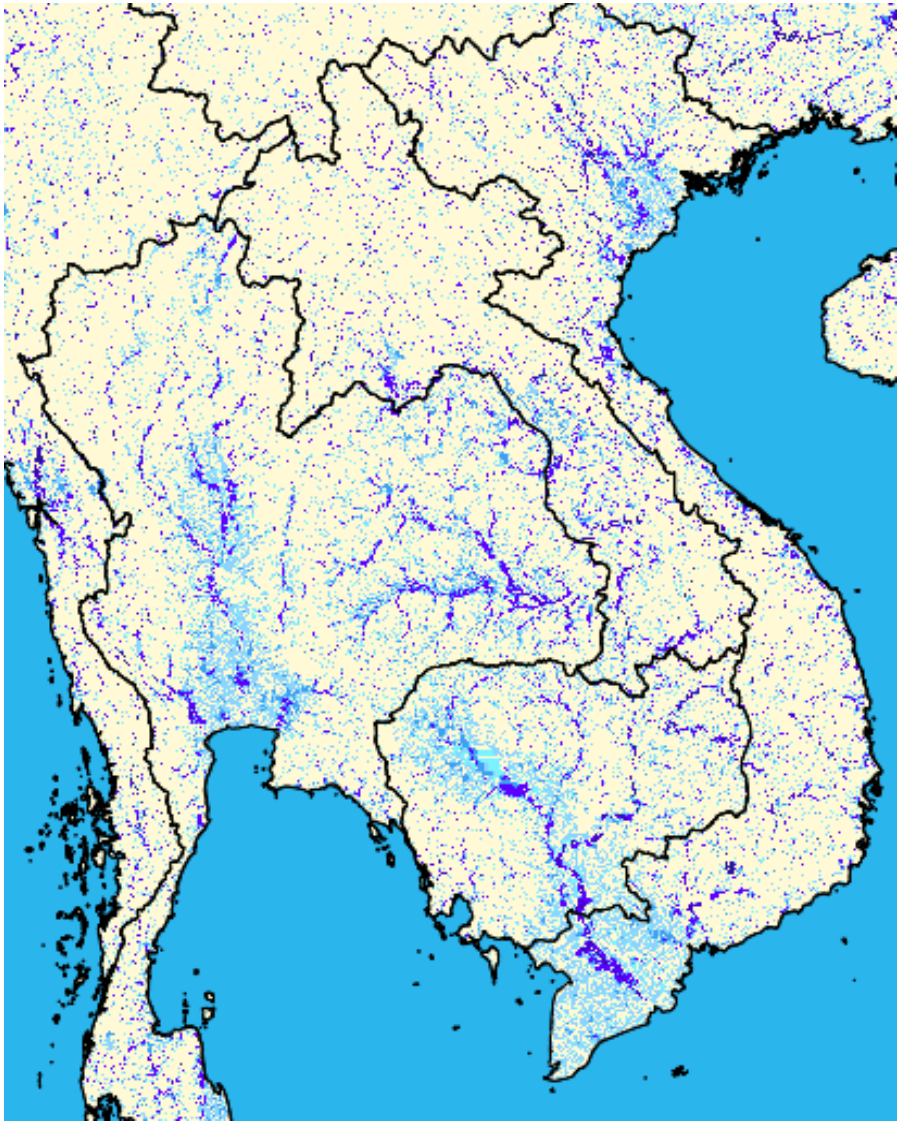
100 yr Loss Hotspots Asia



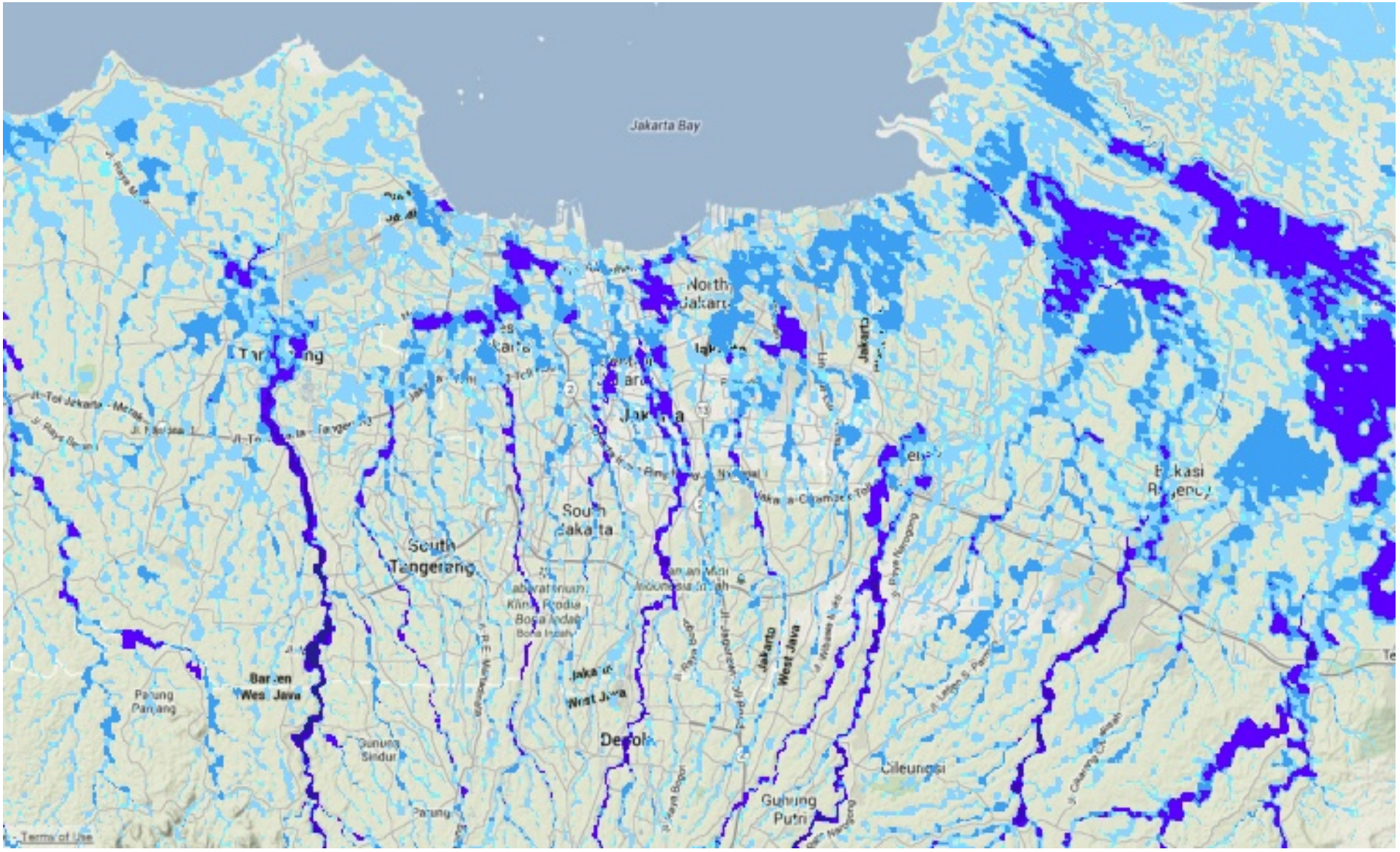
Combine global exposure and vulnerability to develop flood scores and hot spots

SE Asia Region Flood Maps

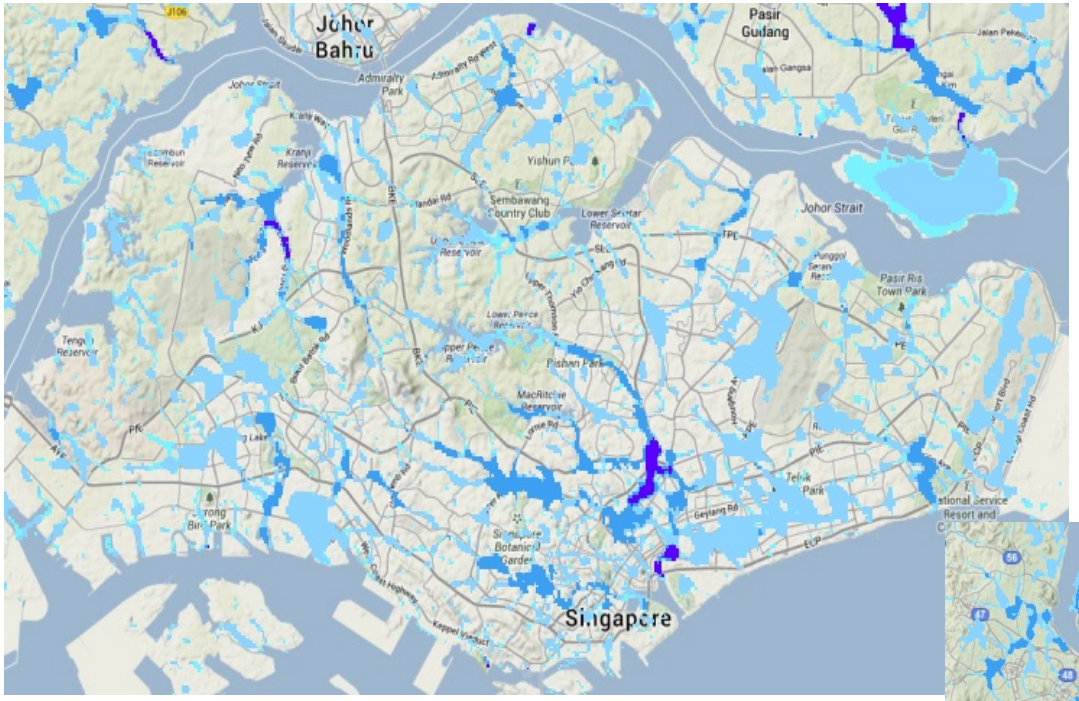


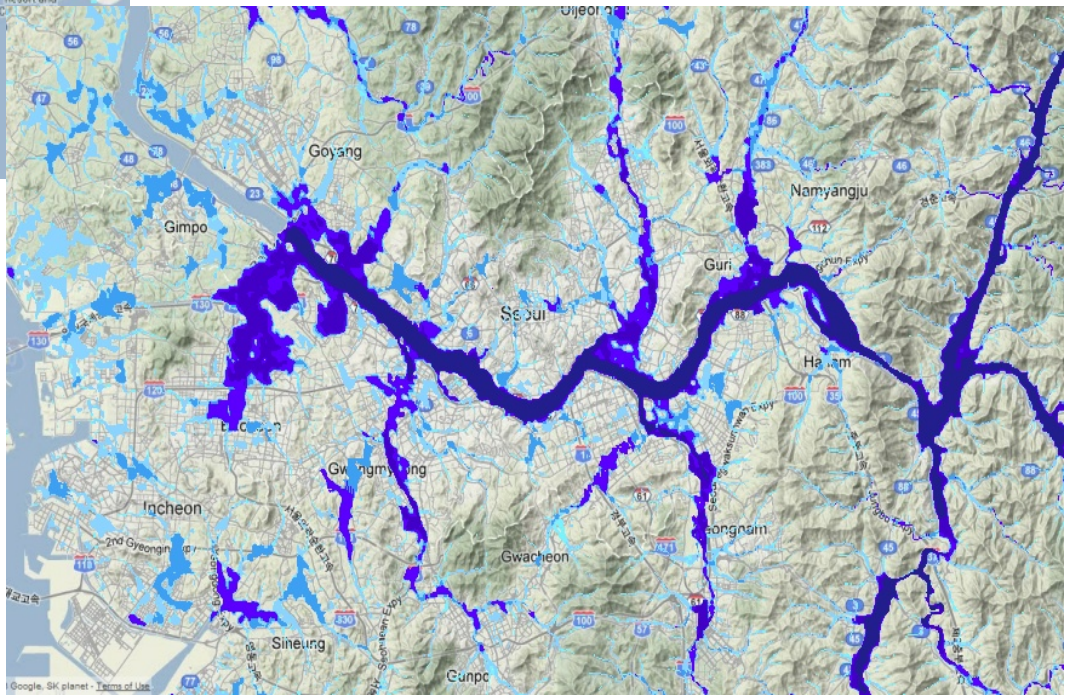


Asia Flood Maps (Jakarta)

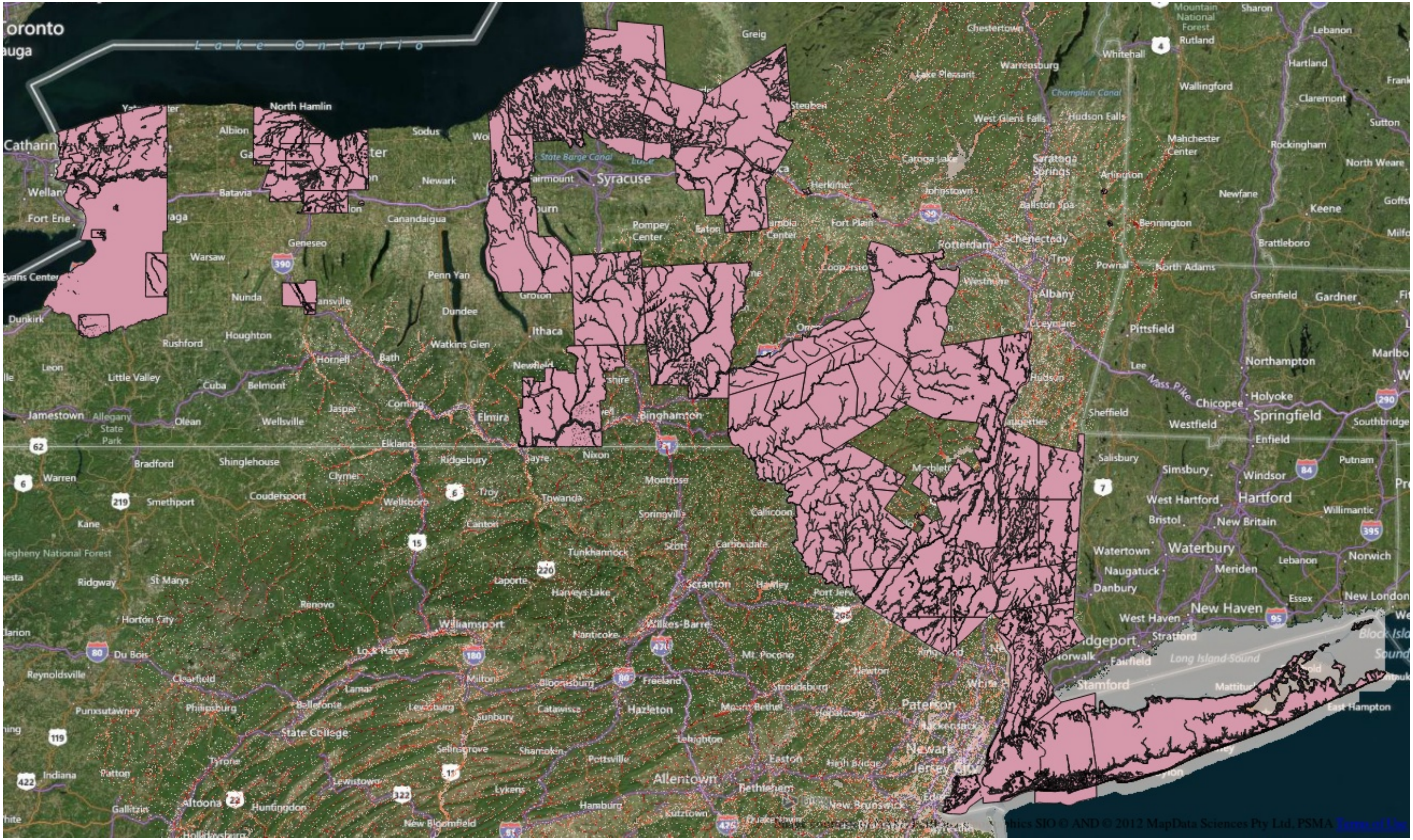


Asia Flood Maps (Singapore, Seoul)





More USA Flood Maps (NY)

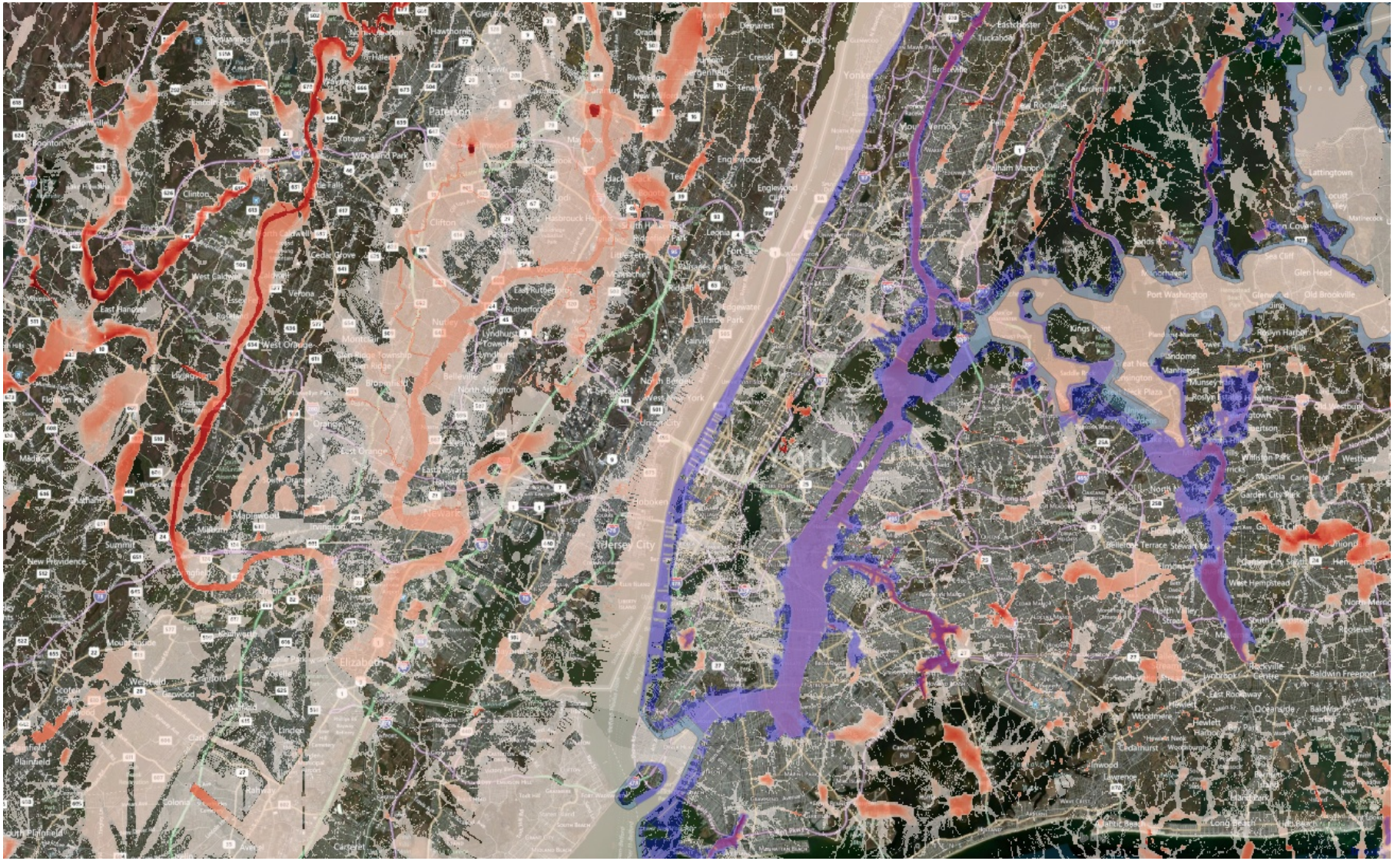


NY State FEMA and KatRisk coverage

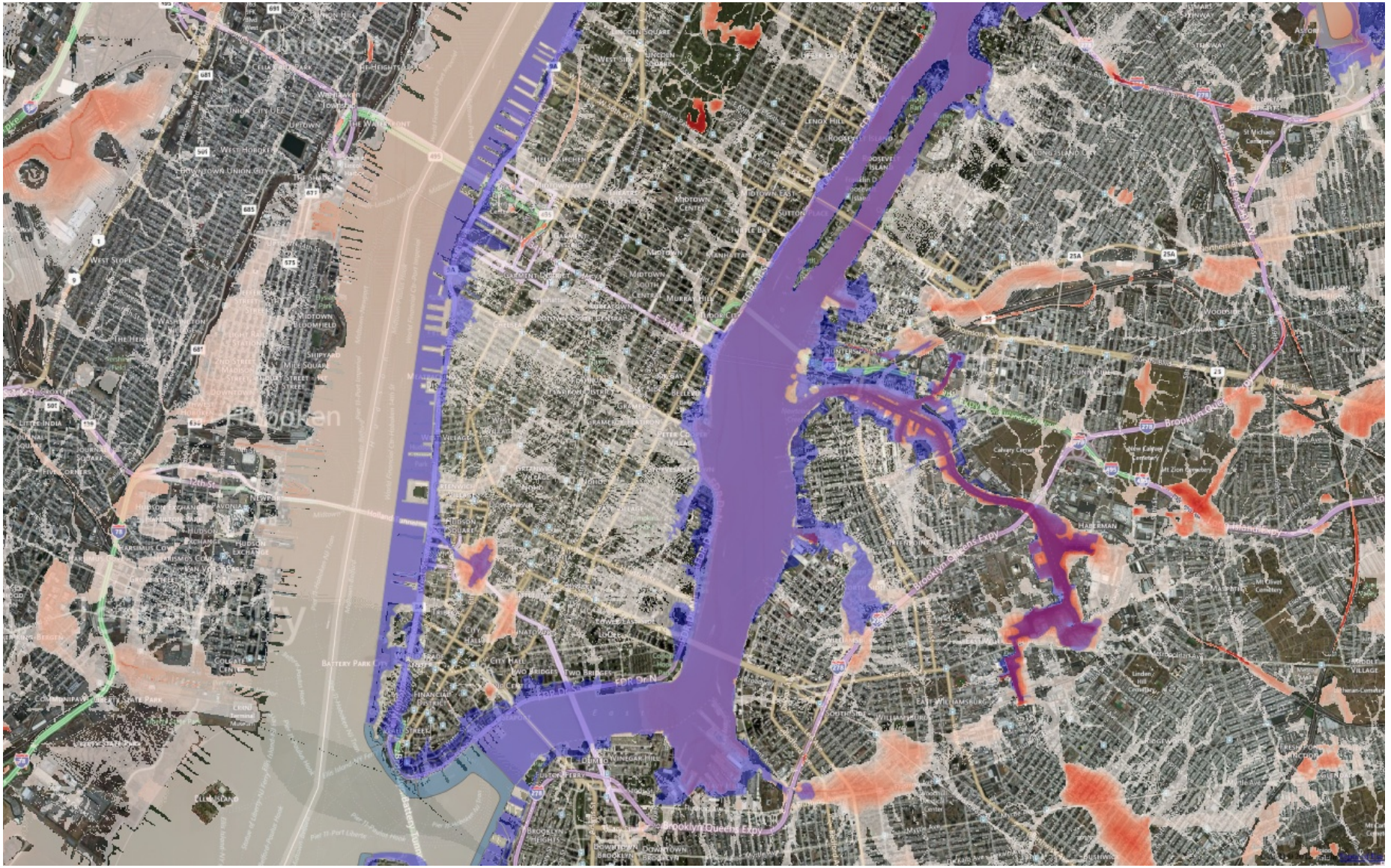
KatRisk

Zoom into NY Flood Maps

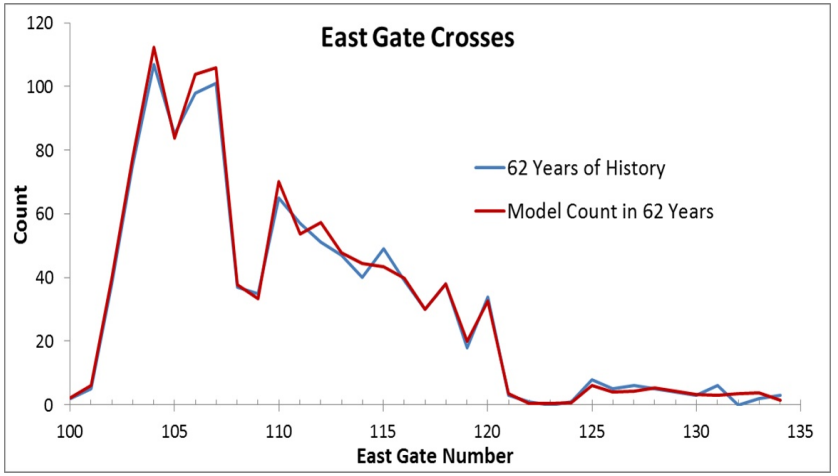
NYC Flood Map

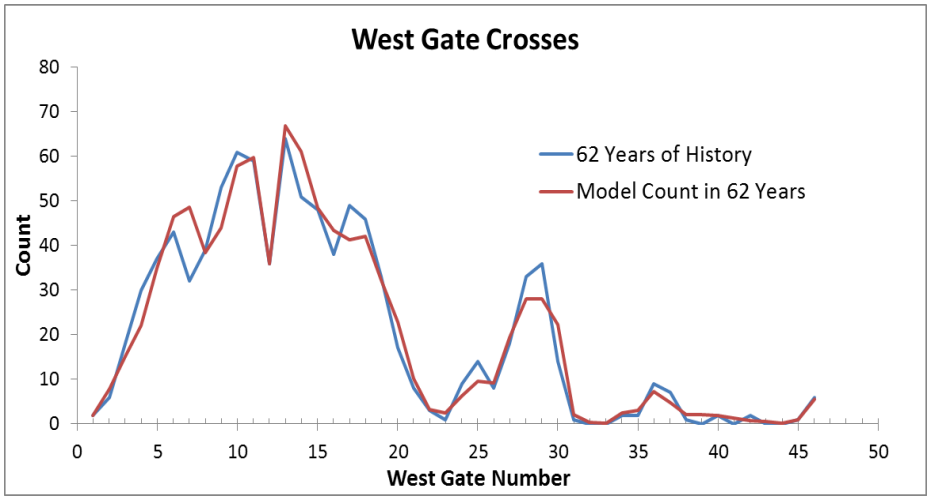


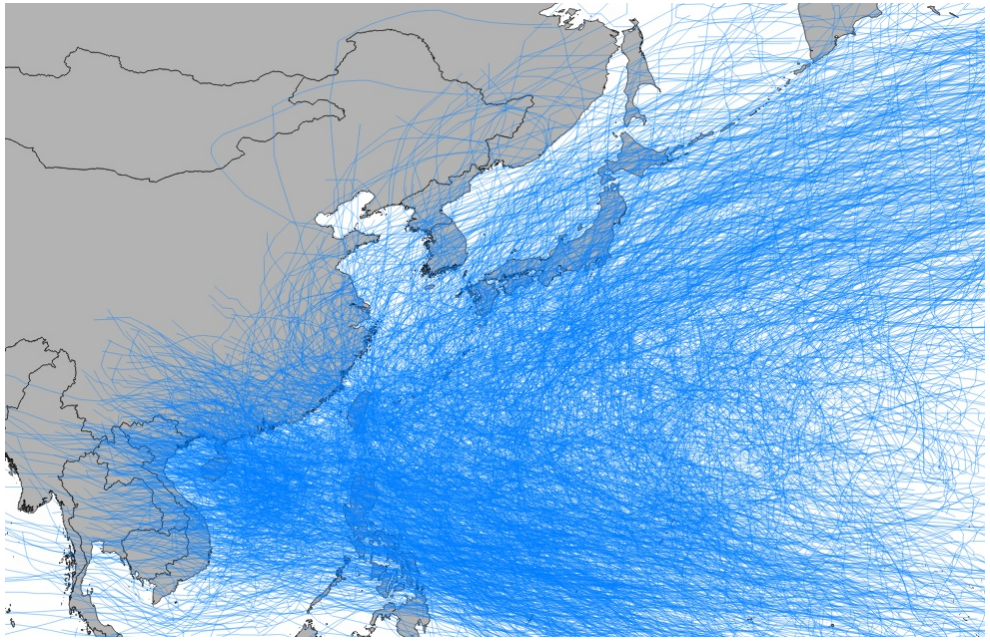
NYC Flood Map

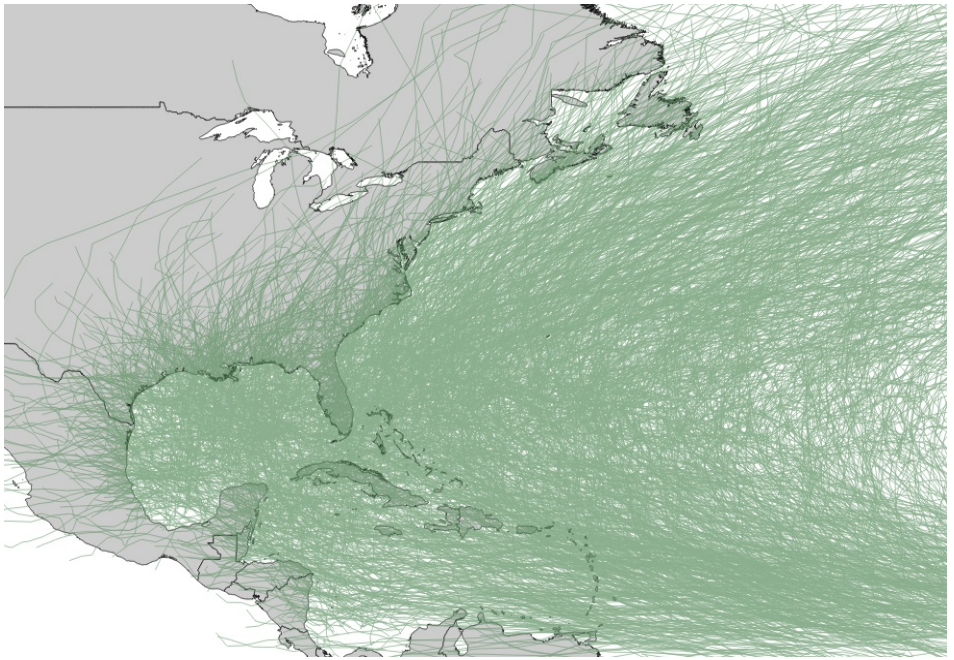


Next Steps: USA and Asia Track Models



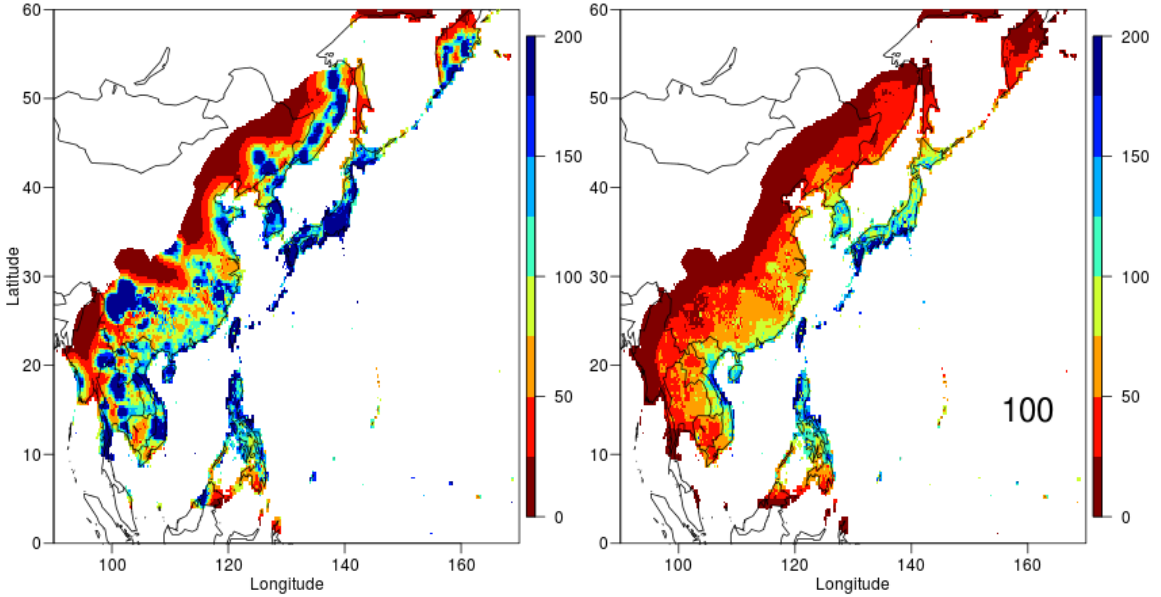


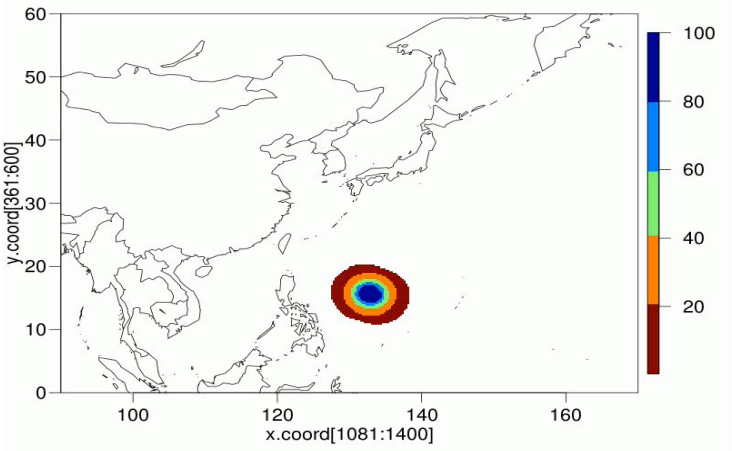




TC Precipitation Modeling

- TC precipitation can be modeled with conceptual models that reflect the space-time correlation structure of TC events
- Left figure shows a possible extreme 10 year maximum TC rain realization, the center figure shows the 10 year return period estimate from TC rain after 10k years of bias corrected simulations.





Summary

- Modeling for insurance = impact modeling
- Modeling needs to be on the scale of large hazard gradients, in flood that is often 10m to 100m
- We can model fluvial and pluvial flooding worldwide thanks to accelerator cards (Nvidia enabled Titan, ORNL)
- Next steps are building a probabilistic event set for portfolio management

Appendix: Risk Vocabulary

- **Hazard**: severity, rate, return period
- **Vulnerability**: inventory, construction classes
- **Exposure**: line of business, TIV and TSI
- **Statistics and analytics**: frequency, PDF and CDF, exceedance probability, OEP, AEP, Bayesian statistics, loss cost, PML, EP uncertainty
- **Frequently used distributions**: Poisson, Beta, Negative Binomial, Gamma