Modelling river flood risk for the whole continental US

Paul Bates, University of Bristol

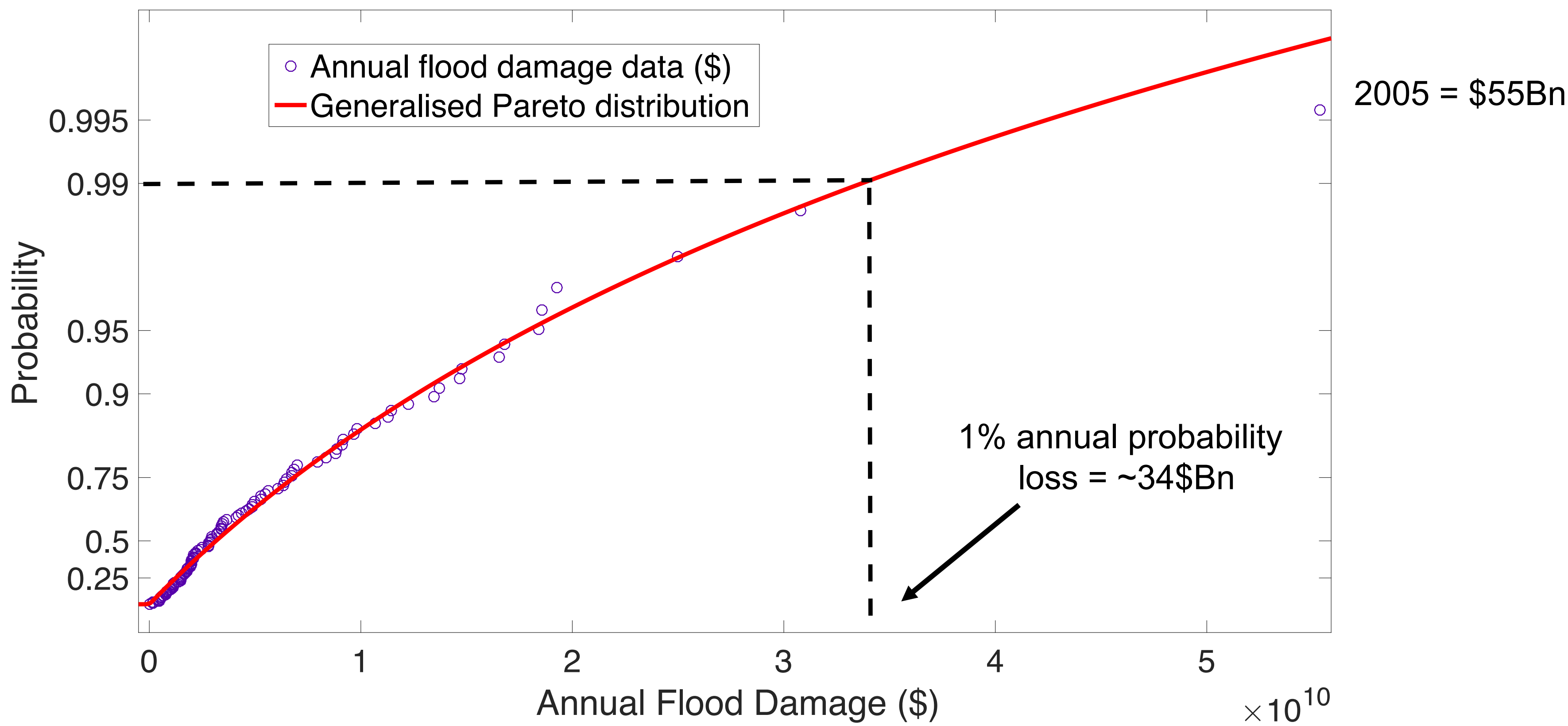
(With lots and lots of help from Niall Quinn, Chris Sampson, Andy Smith, Ollie Wing and Jeff Neal)

www.fathom.global

The problem – US flood river risk

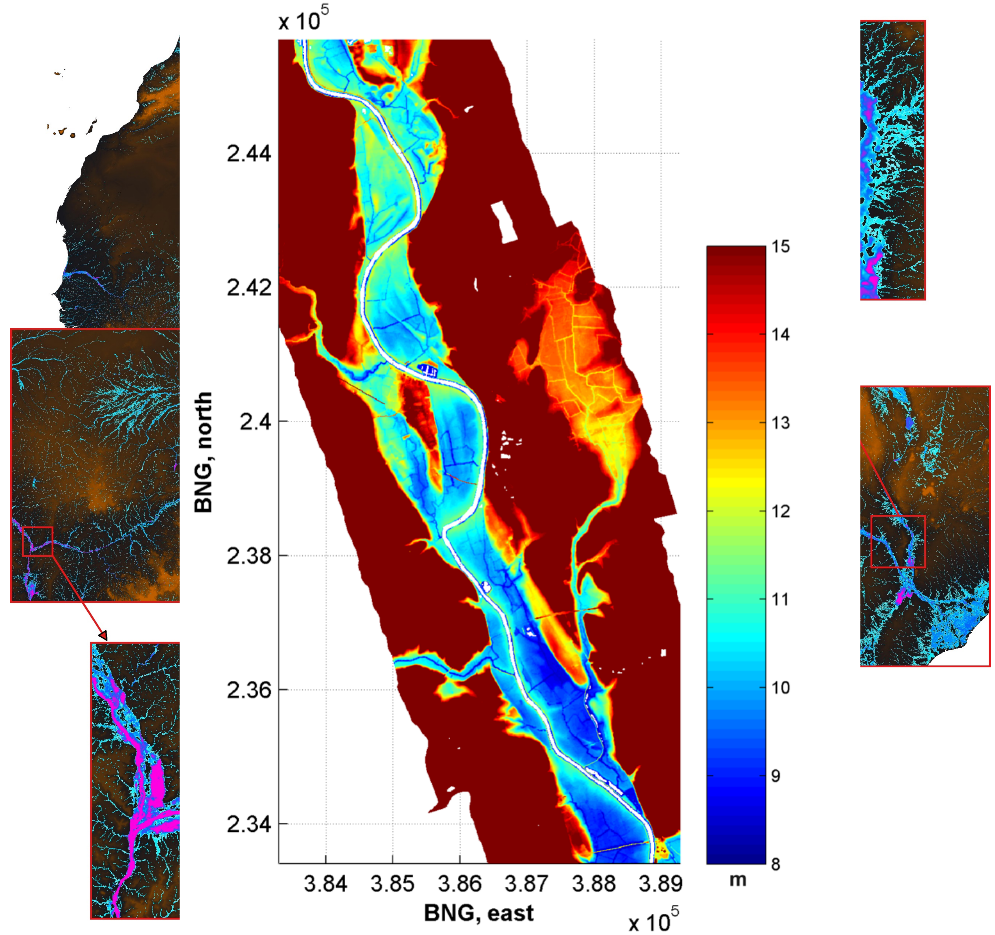
- **Substantial losses each year**
  - Average annual damage 1903-2014 = ~$5Bn
    - (NWS data for river flood, not including coastal events, actual losses at 2014 prices)
  - ~100 fatalities a year on average, with no trend over the same time
  - Significant social disruption
- **National Flood Insurance Programme costs $190M p.a. and is $25Bn in debt**
- **Recent events have raised public awareness of the issue**

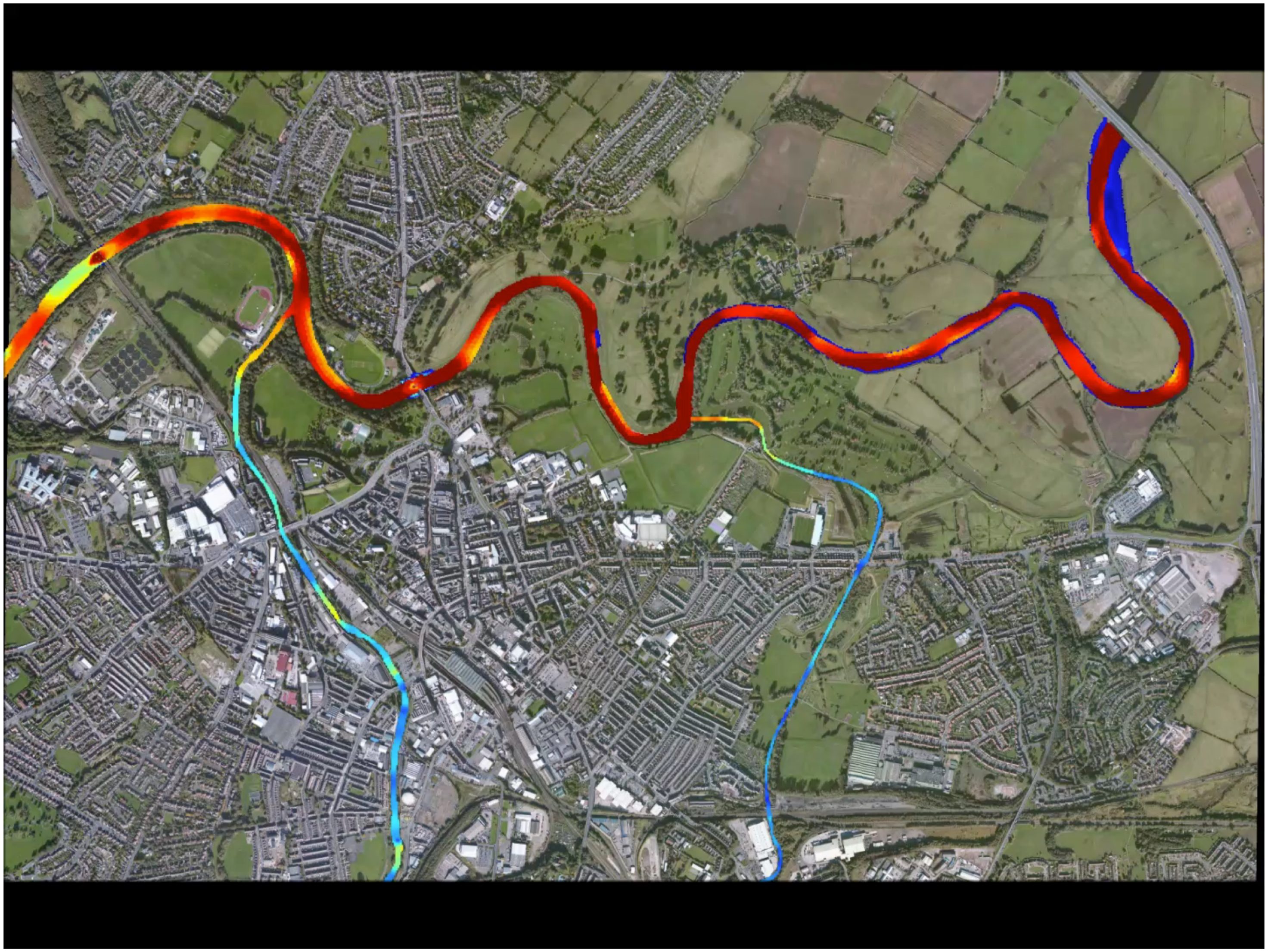
US Flood losses – National Weather Service data



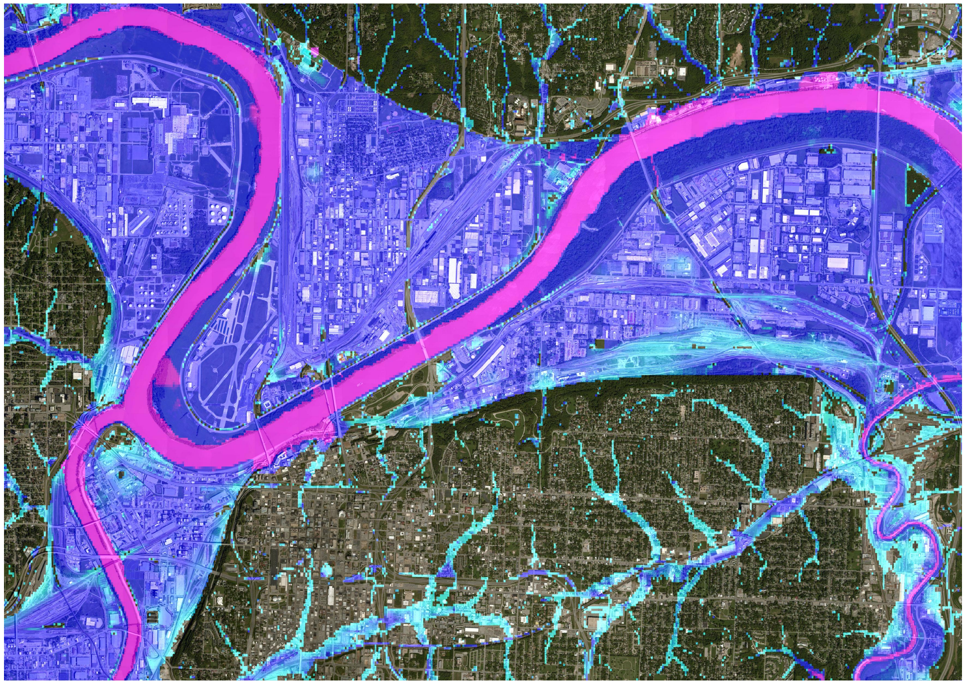
Flood modelling

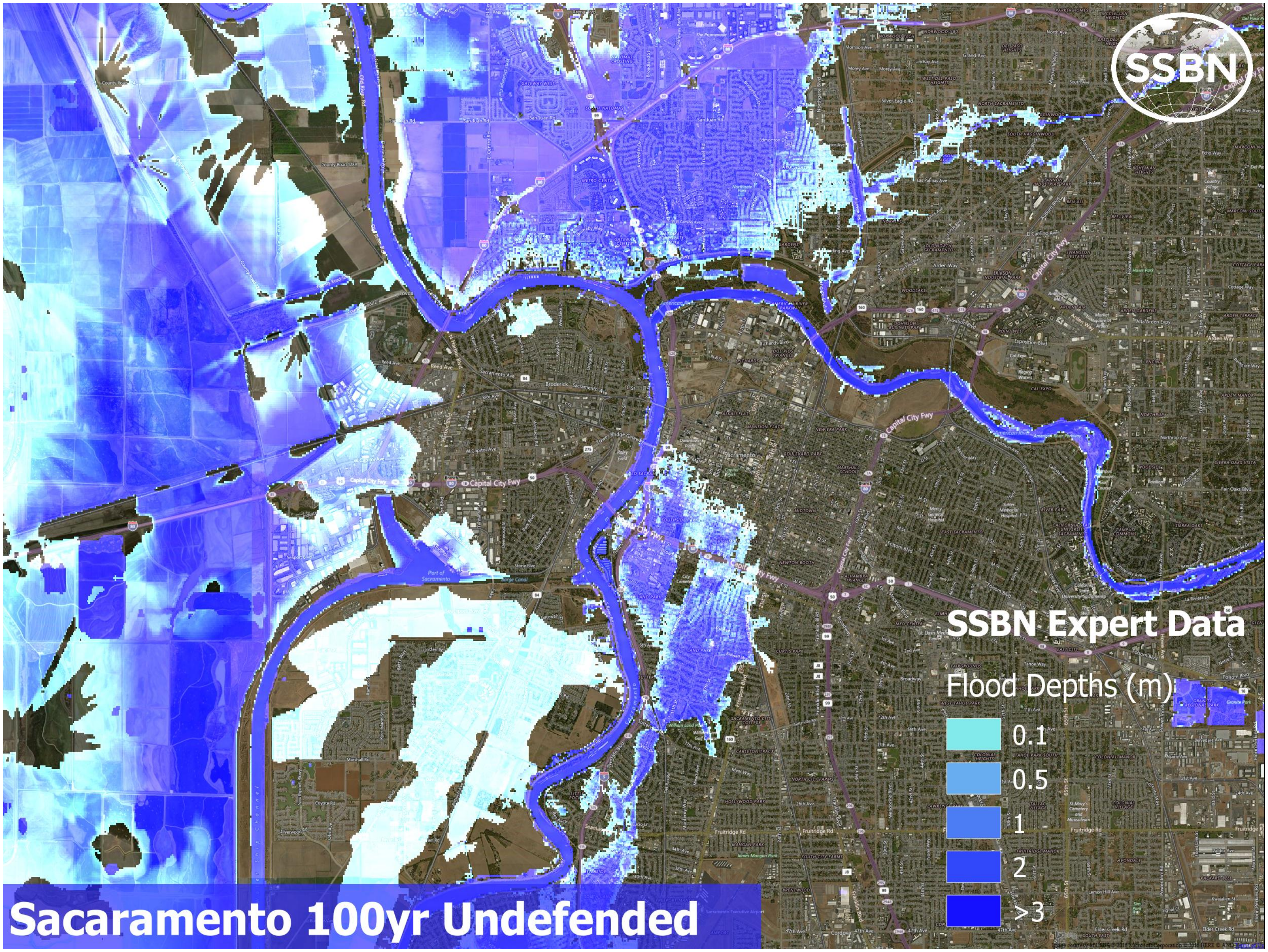
- Over the last 20 years remote sensing has revolutionized flood modelling
  - From early work at reach scales with LiDAR data ….
  - to global data sets enabling continental and global models at <100m resolution

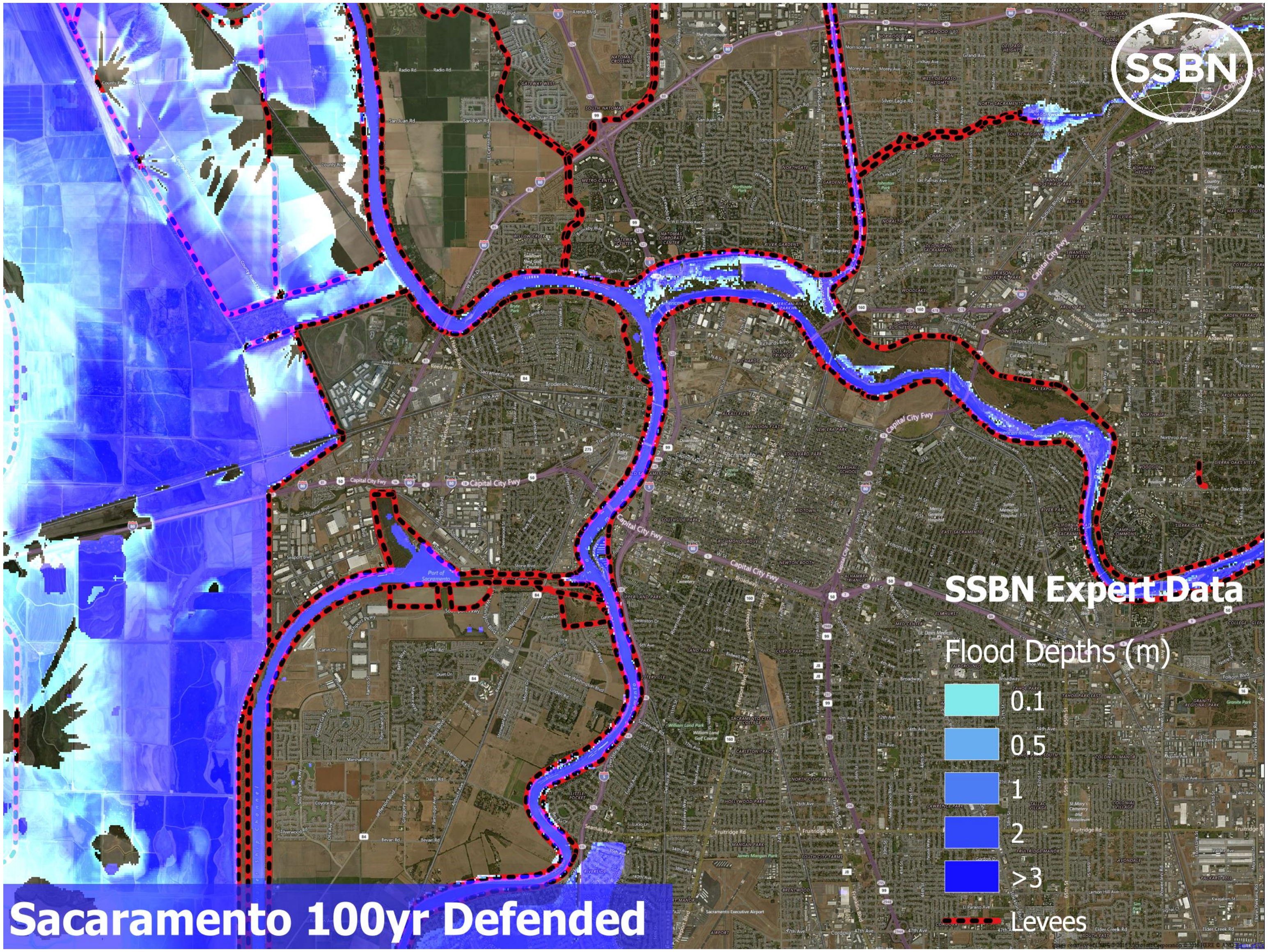




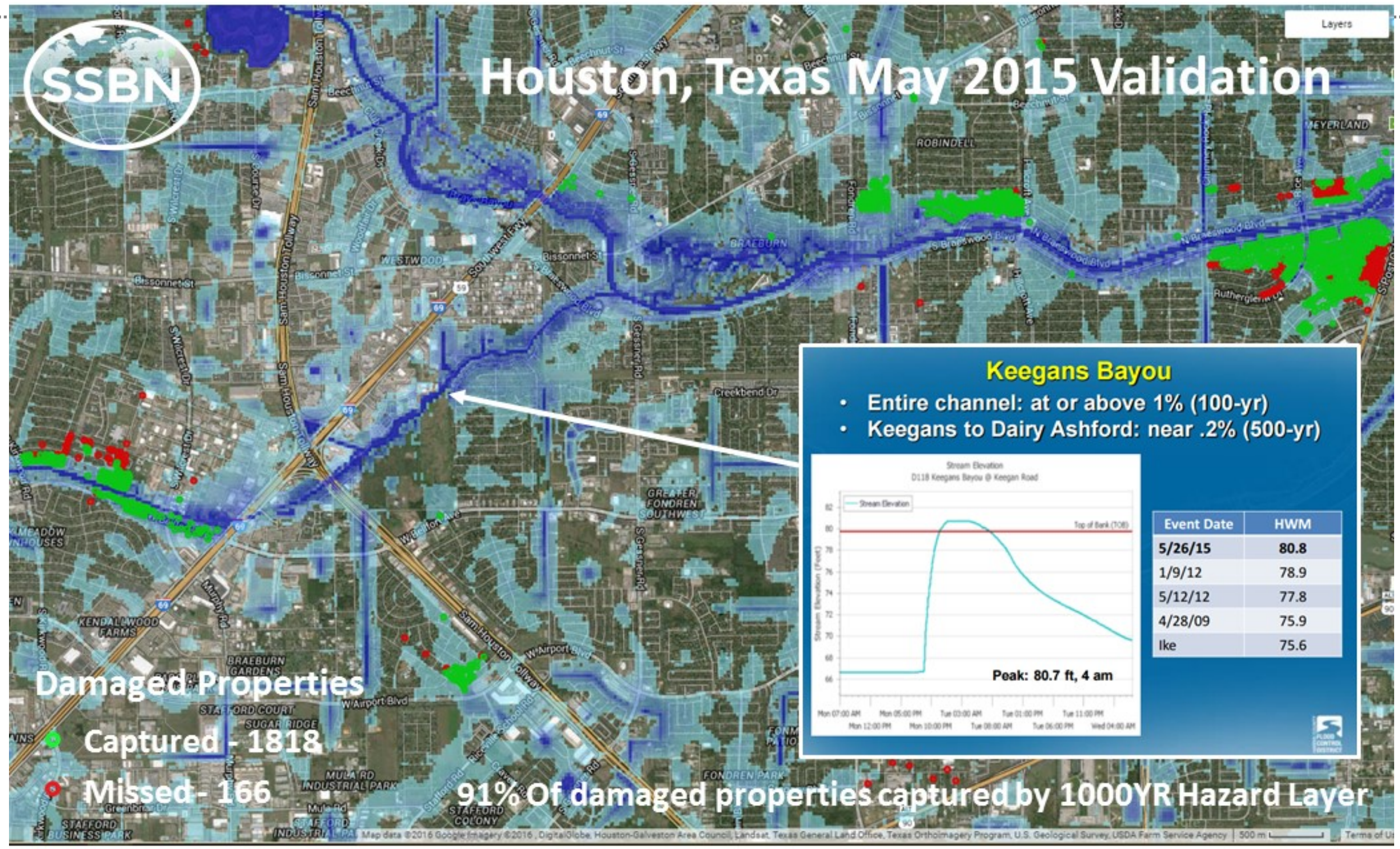
- Whole US model
  - DEM from US National Elevation Dataset (NED)
  - 30m resolution
  - All river channels explicitly represented
  - Boundary conditions from regional flood frequency analysis of rainfall and flow
  - Includes US Army National Levee Dataset
  - 10 return periods from 5 to 1000 years




SSBN
SSBN Expert Data
Flood Depths (m)
0.1
0.5
1
2
>3
Sacaramento 100yr Undefended
Capital City Fwy


SSBN
SSBN Expert Data
Flood Depths (m)
0.1
0.5
1
2
>3
Levees
Sacaramento 100yr Defended

Validation – Memorial day, Houston 2015



- 1984 properties defined as being inundated with flood waters.
- Event magnitude estimates range from 100-500 year event depending on exact location.
- Hit rate of 70% and 91% for the 100 year and 1000 year hazard layers respectively.

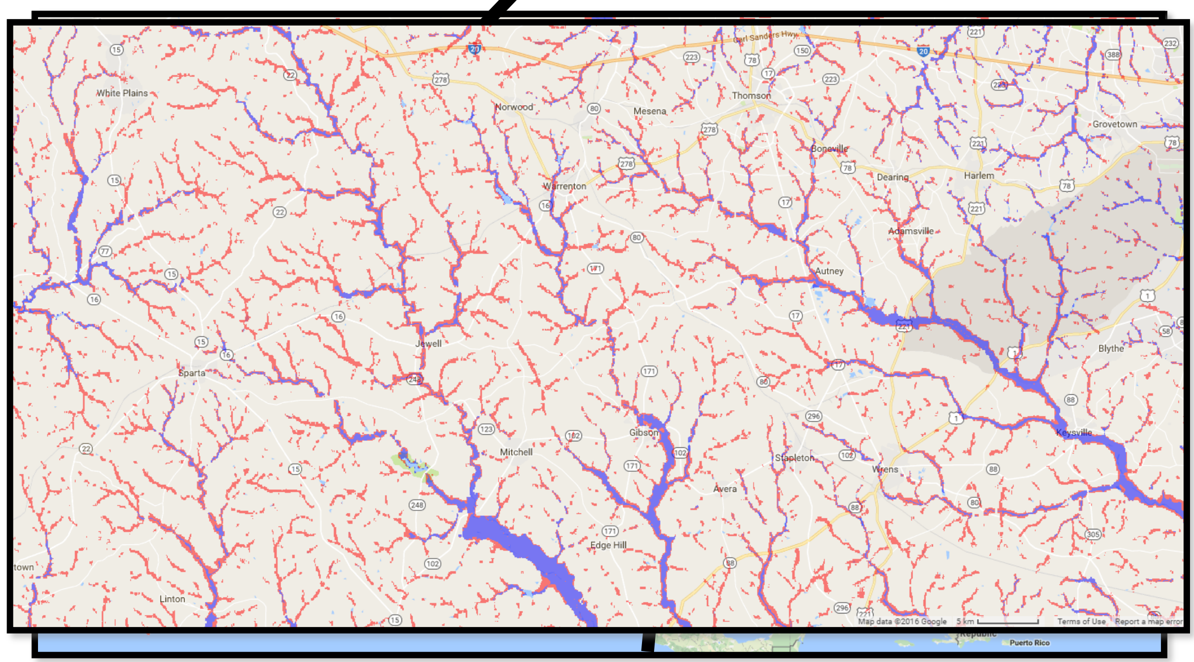
Layer	Captured	Missed	Hit Rate %
20YR	683	1301	34
100YR	1391	593	70
1000YR	1818	166	91

100-500YR observed event magnitude

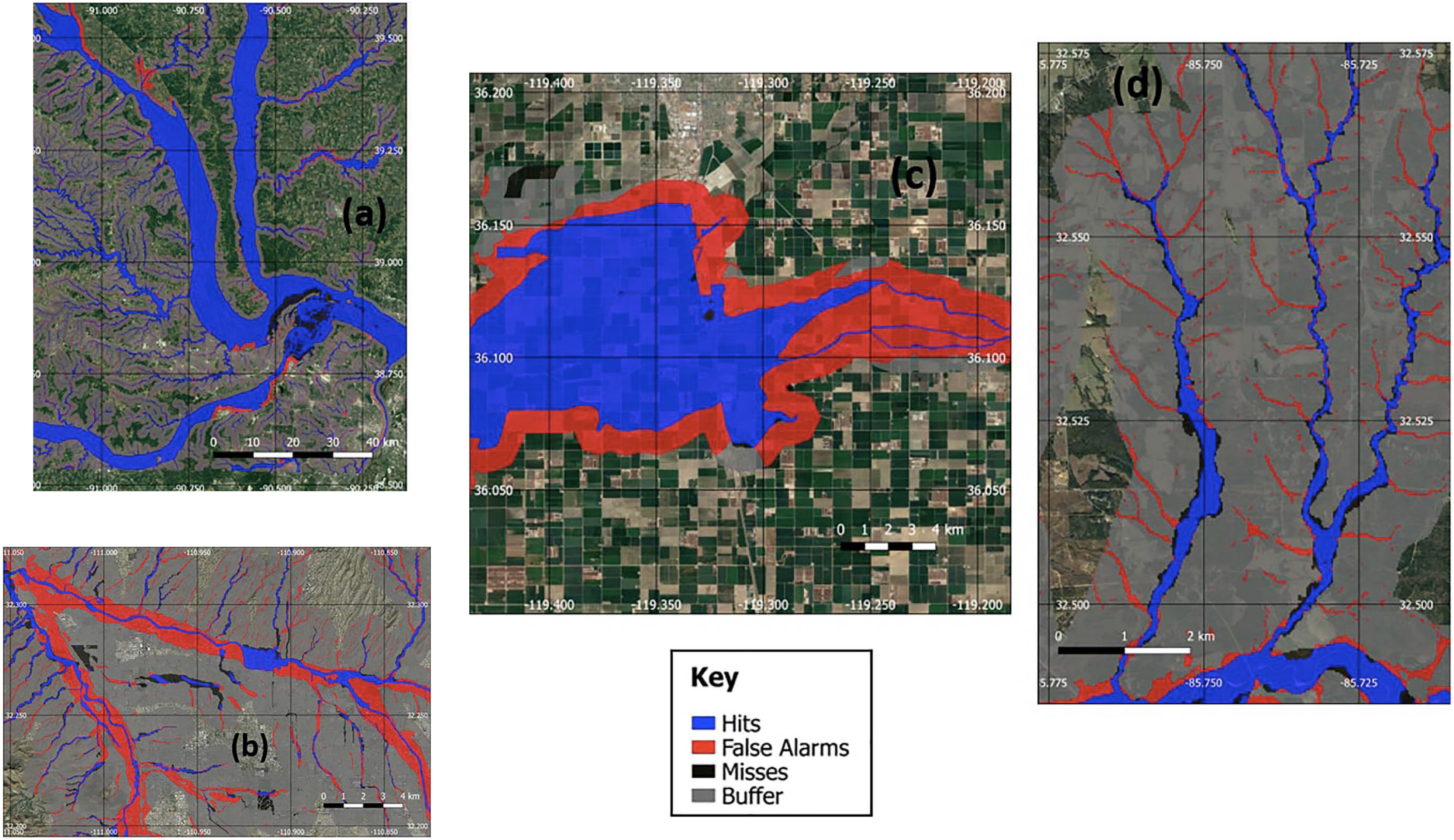
Global model validation: FEMA

- An amalgamation of local studies carried out by FEMA to determine the 1 in 100-year flood extent
- Re-sampled to 90m resolution
- Unexamined areas: both declared and undeclared
- Lack of headwater area coverage problematic
- Analysis performed in Google Earth Engine

No data ?



Over prediction ?


(a)
0 10 20 30 40 km
(b)
0 1 2 3 4 km
(c)
-119.400 -119.350 -119.300 -119.250 -119.200
36.200
36.150
36.100
36.050
0 1 2 3 4 km
(d)
-85.750 -85.725
32.575
32.550
32.525
32.500
0 1 2 km
Key
Hits
False Alarms
Misses
Buffer

Global model validation: FEMA

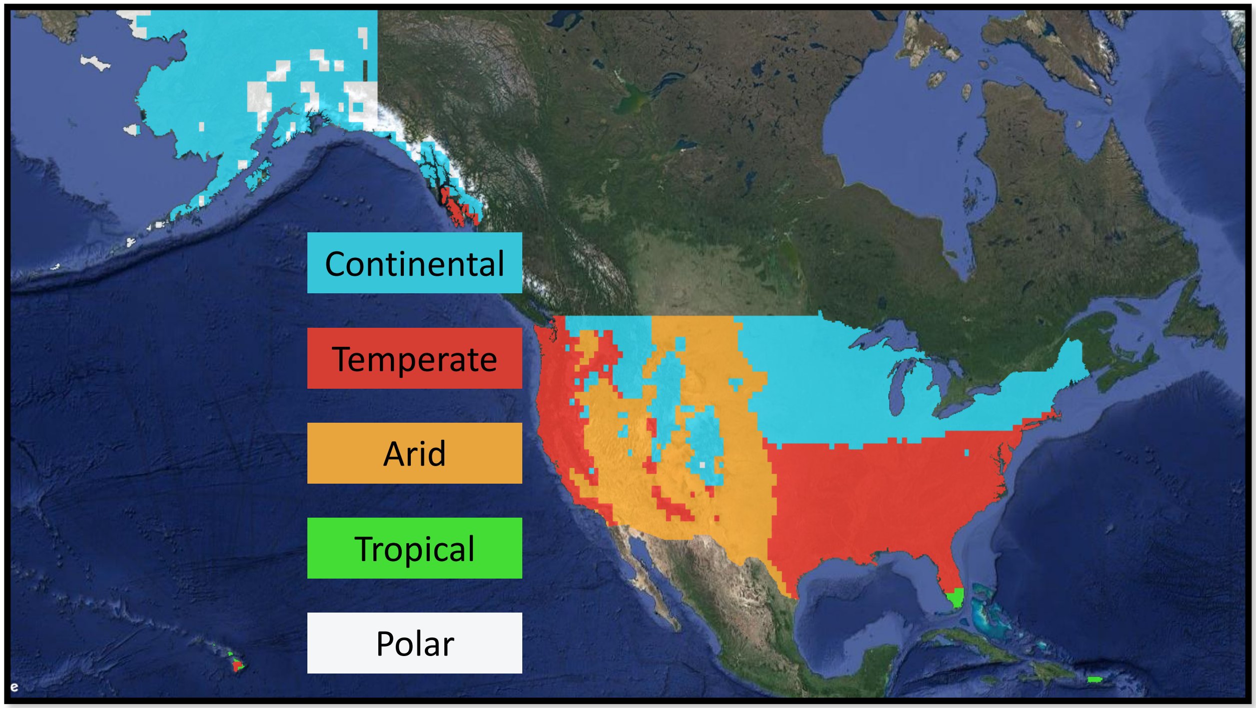
Hite Rate = 81%

- 84% for catchments >80 km^2
- 86% for 'high quality' FEMA data at all scales

CSI = 59%

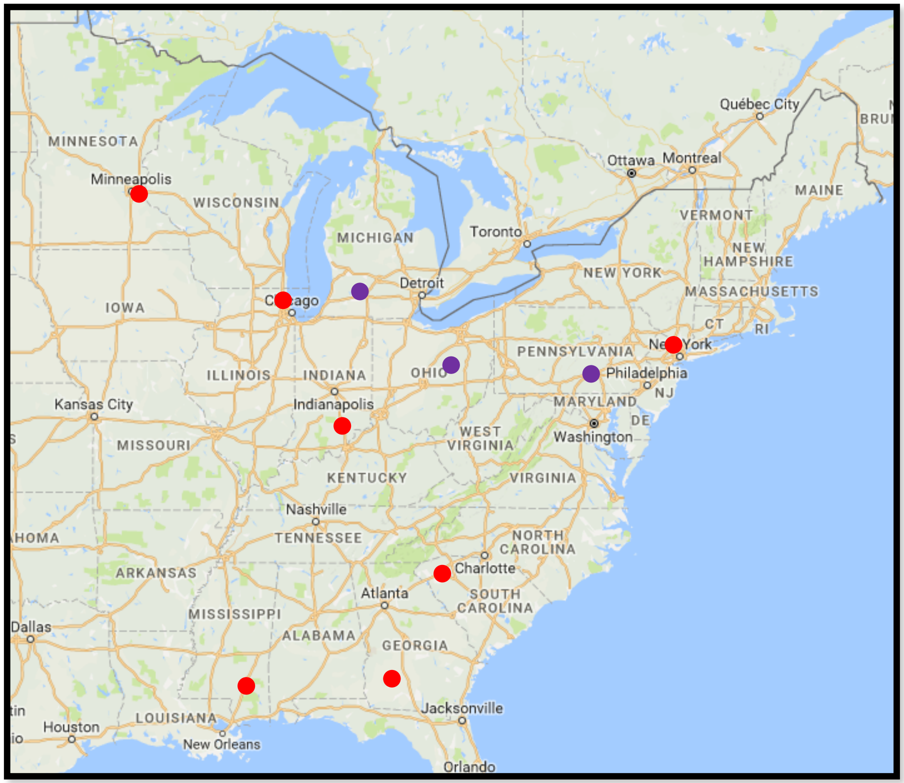
Köppen-Geiger zones

- Temperate = 84%
- Arid = 73%
- Continental = 78%

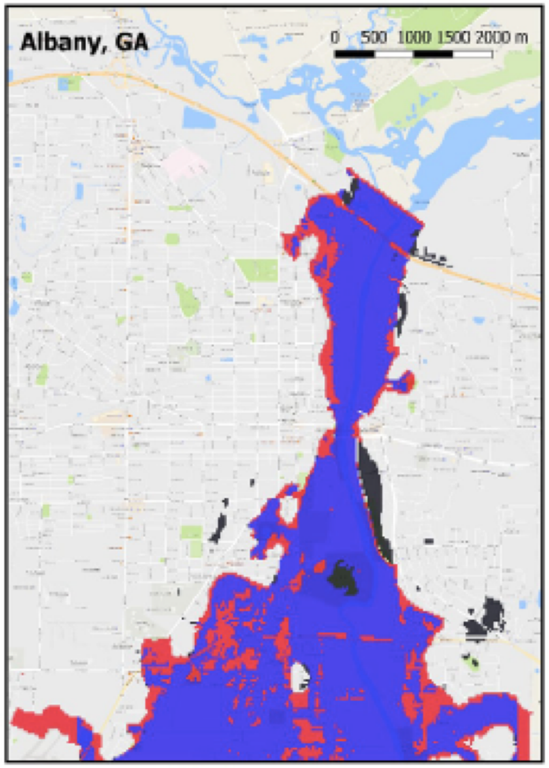


Catchment-scale validation: USGS 1D models

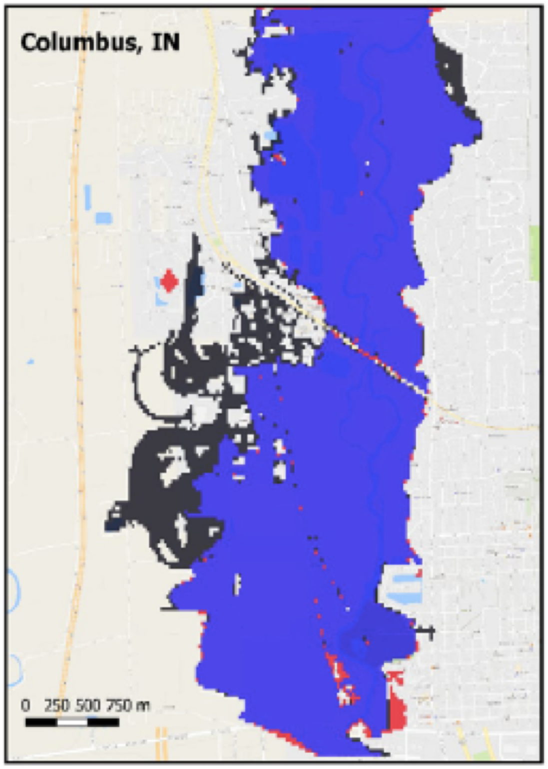
- Isolated local modelling studies – usually a few kilometres of a single stream
- 10 sites with 100-year simulations
- 3 further sites with events of varying magnitude



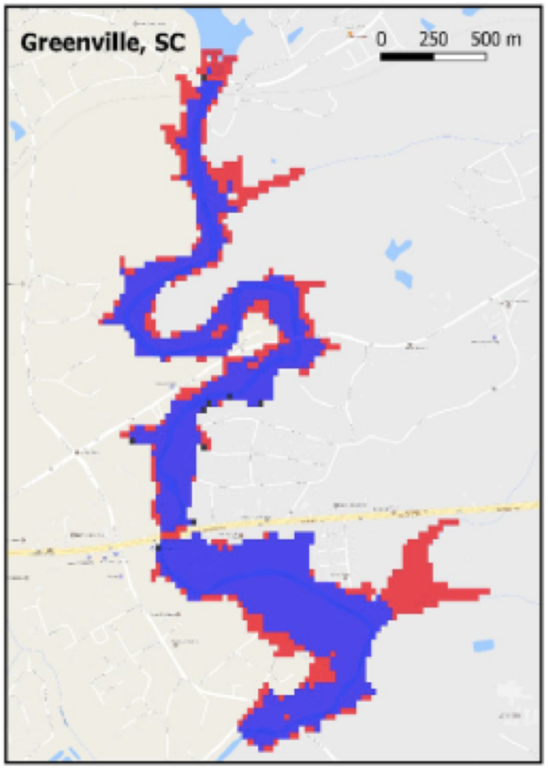
LOCATION	HIT RATE (%)
Albany, GA	**93.8**
Columbus, IN	**83.3**
Greenville, SC	**99.7**
Hattiesburg, MS	**93.7**
Lincolnshire, IL	**81.8**
Minneapolis, MN	**91.0**
Ridgewood, NJ	**88.6**



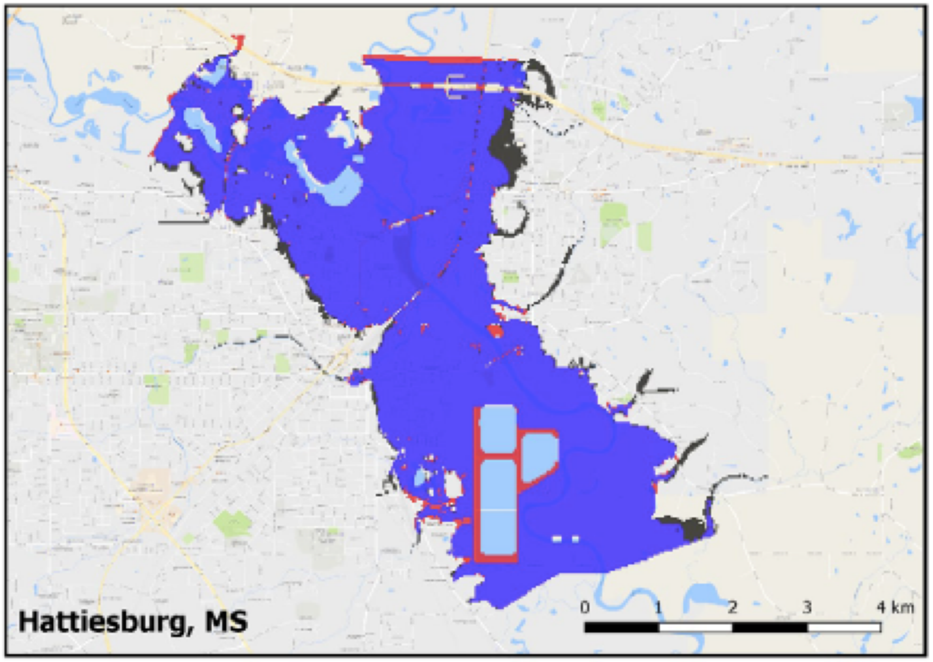
LOCATION	HIT RATE (%)
Albany, GA	**93.8**
Columbus, IN	**83.3**
Greenville, SC	**99.7**
Hattiesburg, MS	**93.7**
Lincolnshire, IL	**81.8**
Minneapolis, MN	**91.0**
Ridgewood, NJ	**88.6**



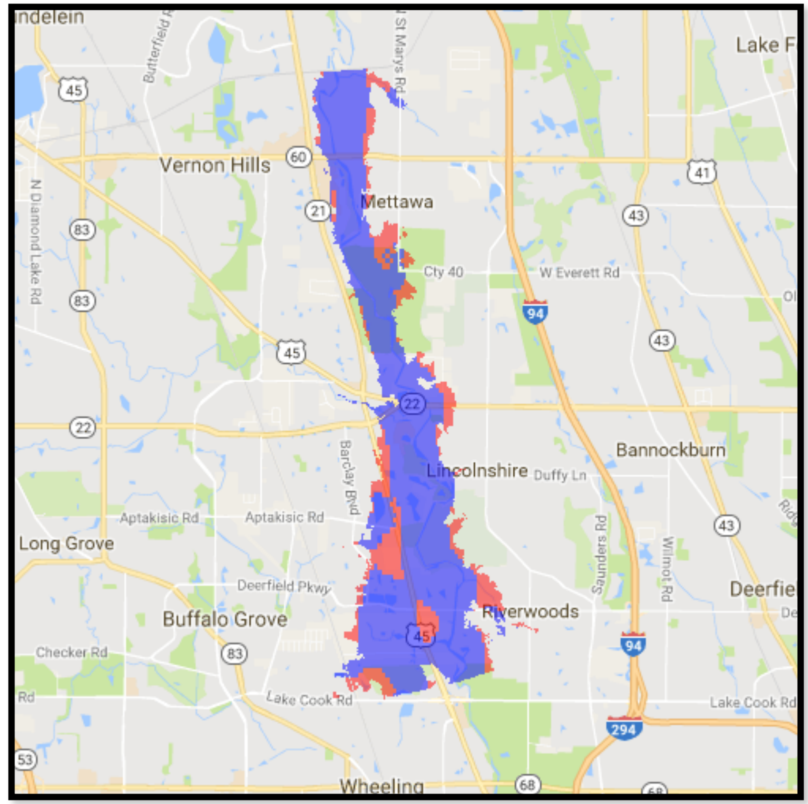
LOCATION	HIT RATE (%)
Albany, GA	**93.8**
Columbus, IN	**83.3**
Greenville, SC	**99.7**
Hattiesburg, MS	**93.7**
Lincolnshire, IL	**81.8**
Minneapolis, MN	**91.0**
Ridgewood, NJ	**88.6**



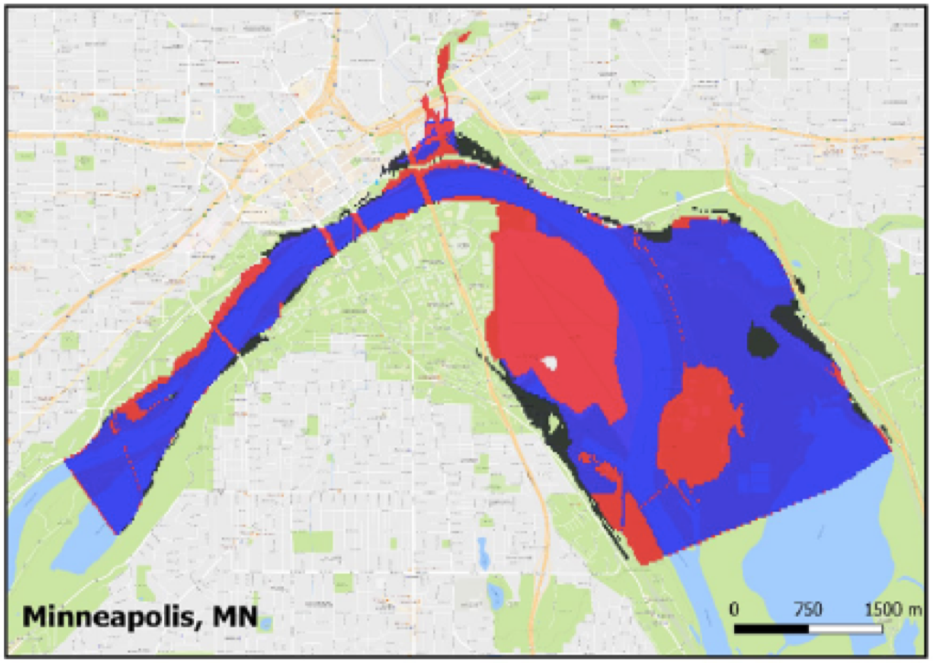
LOCATION	HIT RATE (%)
Albany, GA	**93.8**
Columbus, IN	**83.3**
Greenville, SC	**99.7**
Hattiesburg, MS	**93.7**
Lincolnshire, IL	**81.8**
Minneapolis, MN	**91.0**
Ridgewood, NJ	**88.6**



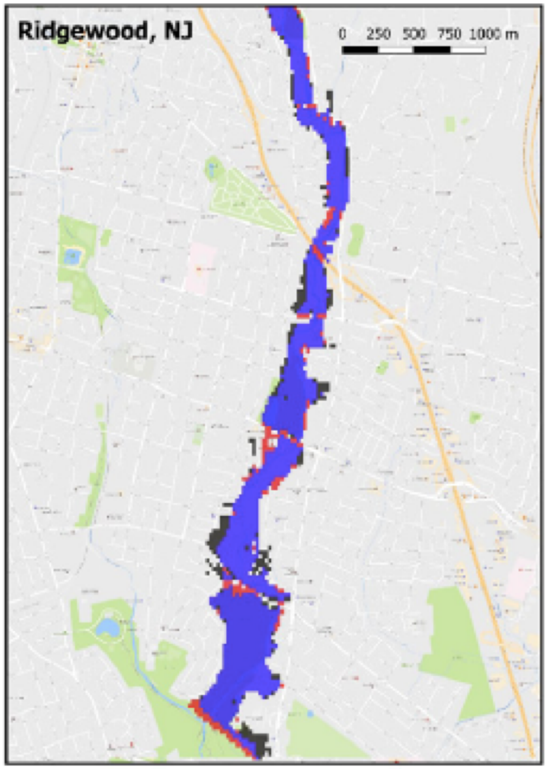
LOCATION	HIT RATE (%)
Albany, GA	**93.8**
Columbus, IN	**83.3**
Greenville, SC	**99.7**
Hattiesburg, MS	**93.7**
Lincolnshire, IL	**81.8**
Minneapolis, MN	**91.0**
Ridgewood, NJ	**88.6**



LOCATION	HIT RATE (%)
Albany, GA	**93.8**
Columbus, IN	**83.3**
Greenville, SC	**99.7**
Hattiesburg, MS	**93.7**
Lincolnshire, IL	**81.8**
Minneapolis, MN	**91.0**
Ridgewood, NJ	**88.6**

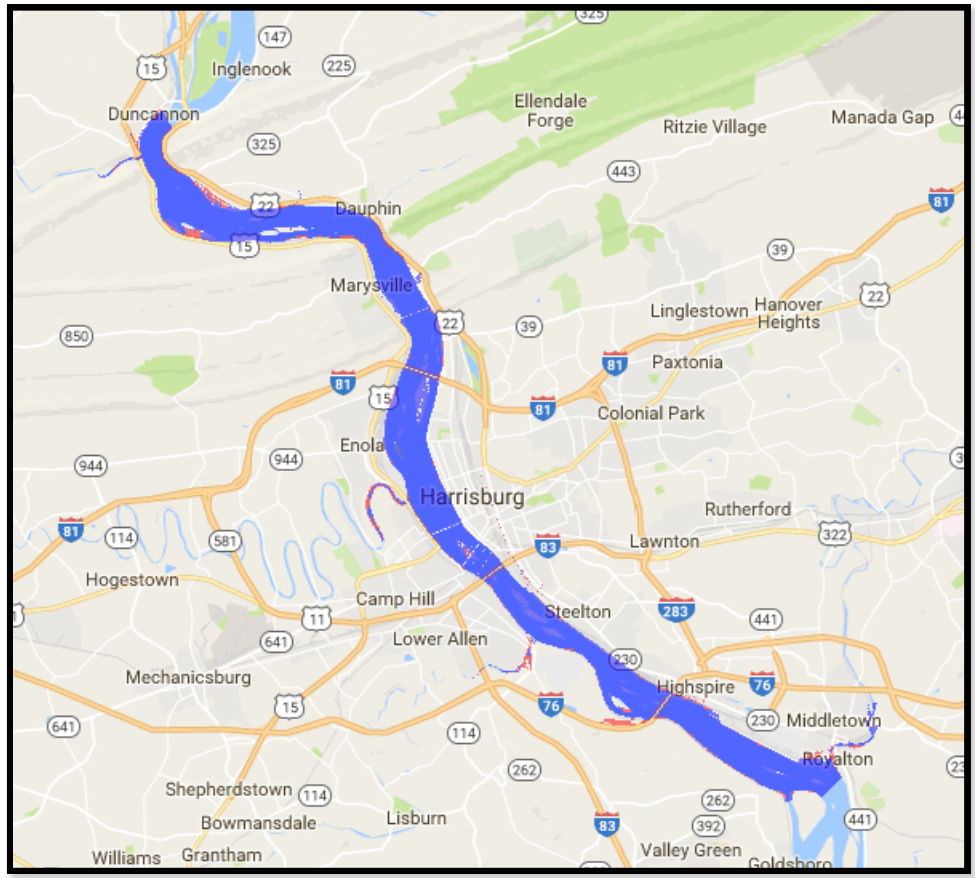


LOCATION	HIT RATE (%)
Albany, GA	**93.8**
Columbus, IN	**83.3**
Greenville, SC	**99.7**
Hattiesburg, MS	**93.7**
Lincolnshire, IL	**81.8**
Minneapolis, MN	**91.0**
Ridgewood, NJ	**88.6**

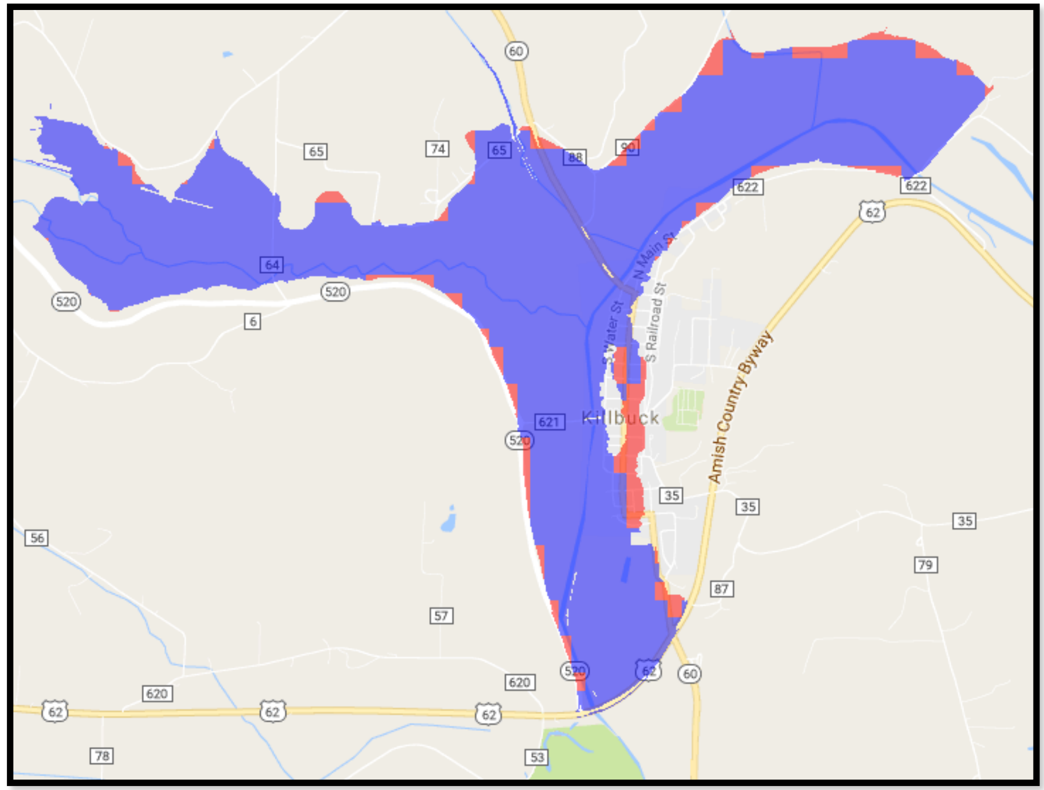


LOCATION	HIT RATE (%) 1 in 500
Battle Creek, MI	**90.7**

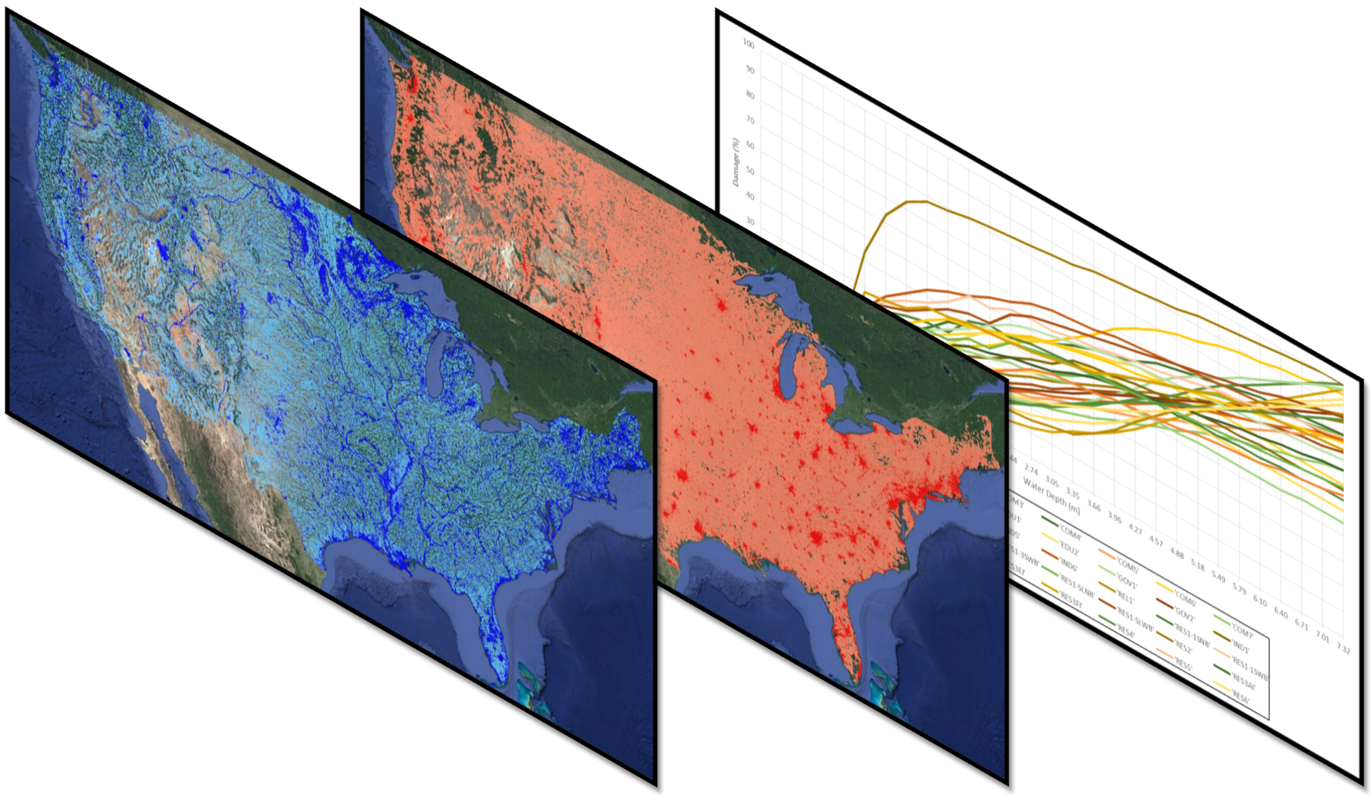
LOCATION	HIT RATE (%) 1 in 10
Harrisburg, PA	**95.1**



LOCATION	HIT RATE (%) 1 in 500
Killbuck, OH	**94.3**

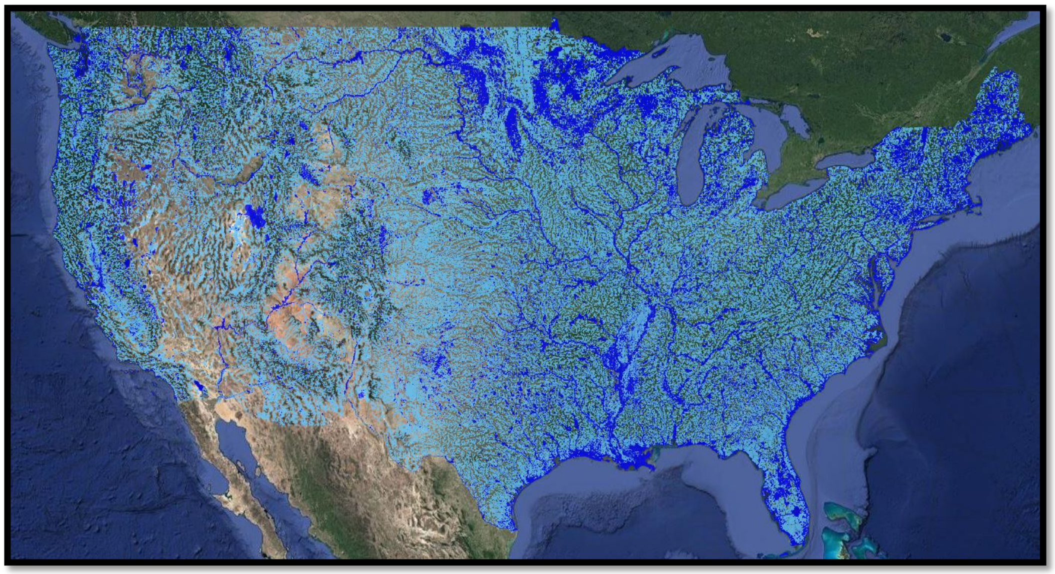


Basics of risk calculation



- Hazard
    - Flooded area
- Exposure
    - Value of buildings within floodplain
    - Number of people within floodplain
- Vulnerability
    - Potential damages

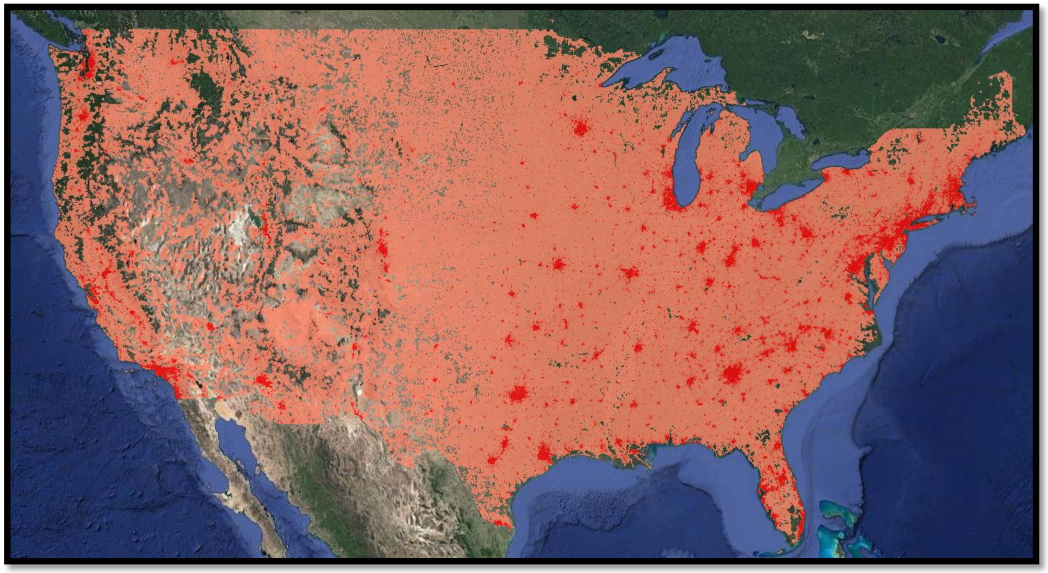
Hazard data



- ~30m resolution flood hazard model of CONUS (Wing *et al.*, 2017)

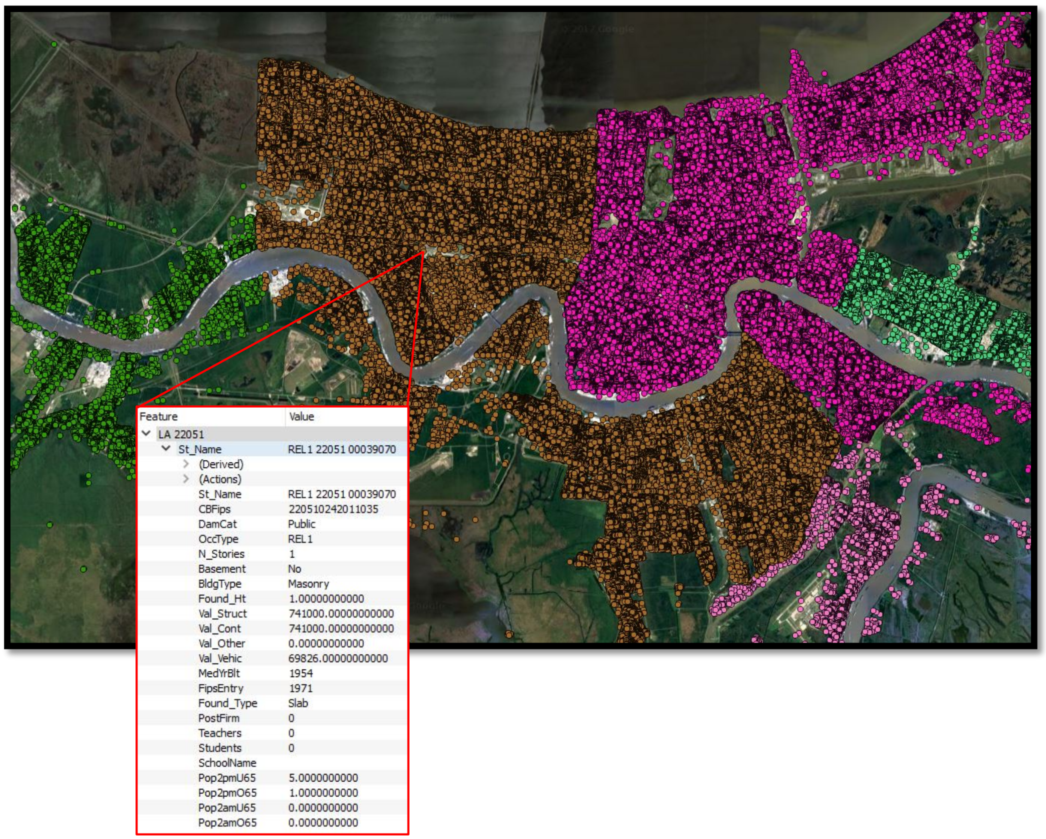
Wing, O. E. J. *et al.* (2017), Validation of a 30 m resolution flood hazard model of the conterminous United States, *Water Resour. Res.*, 53, 7968–7986, doi:10.1002/2017WR020917.

Socio-economic data



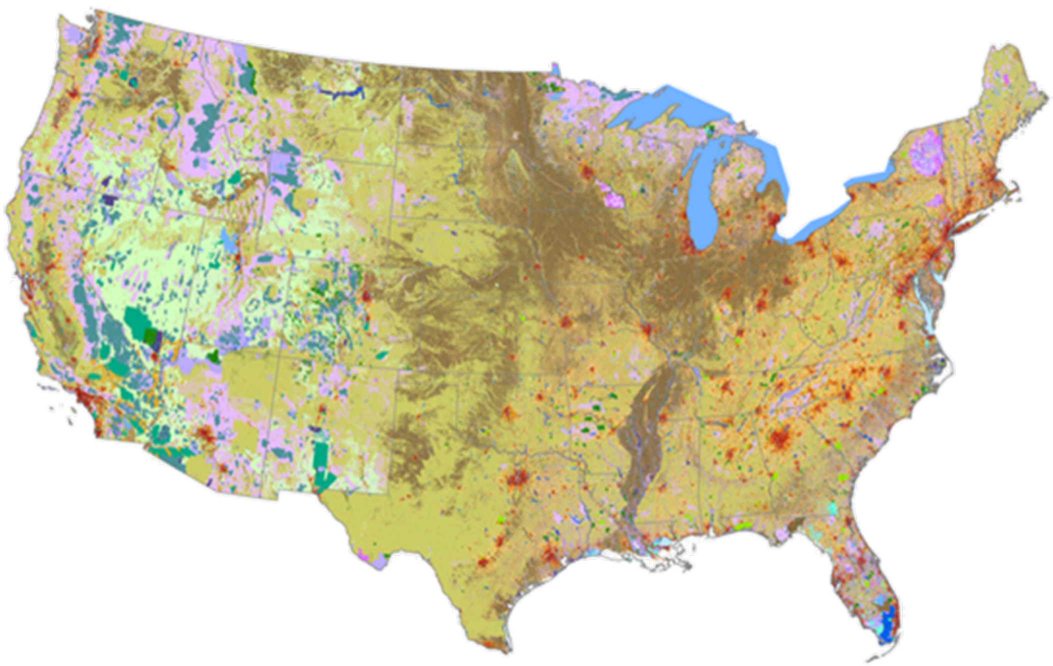
- EPA EnviroAtlas dasymetric population map
- Assigns 2010 census populations to 30m pixels based on land-use and slope

Socio-economic data



- FEMA National Structure Inventory
- Information on over 100M buildings in the CONUS:
  - Location
  - Value
  - Type
  - Number of storeys
  - Presence of basement

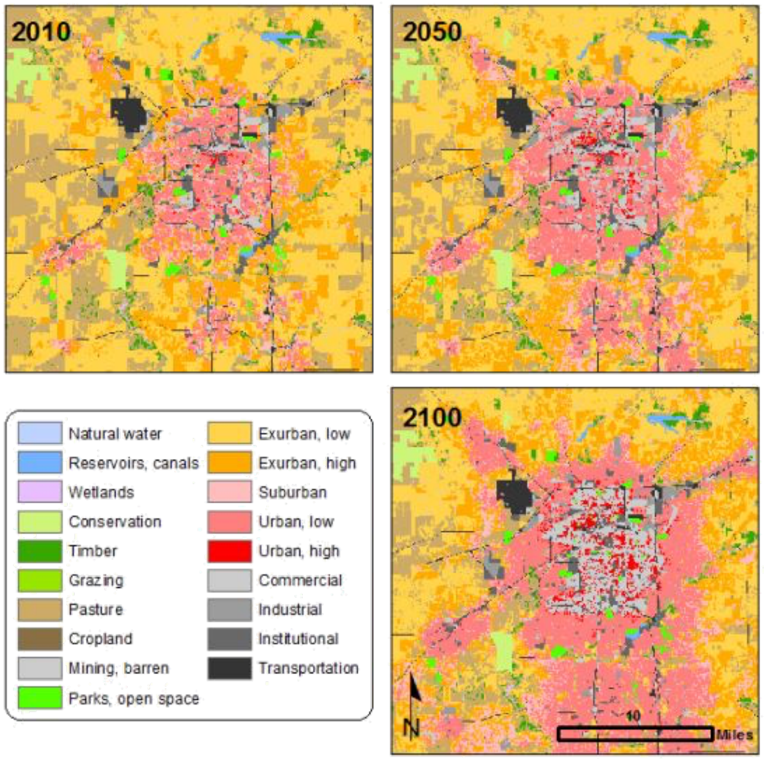
Socio-economic data



- National Land Use Database (Theobold, 2014)
- Indicates developed areas across the CONUS

Theobold, D. M. (2014), Development and Applications of a Comprehensive Land Use Classification and Map for the US, *PLOS ONE*, 9(4), e94628, doi:10.1371/journal.pone.0094628.

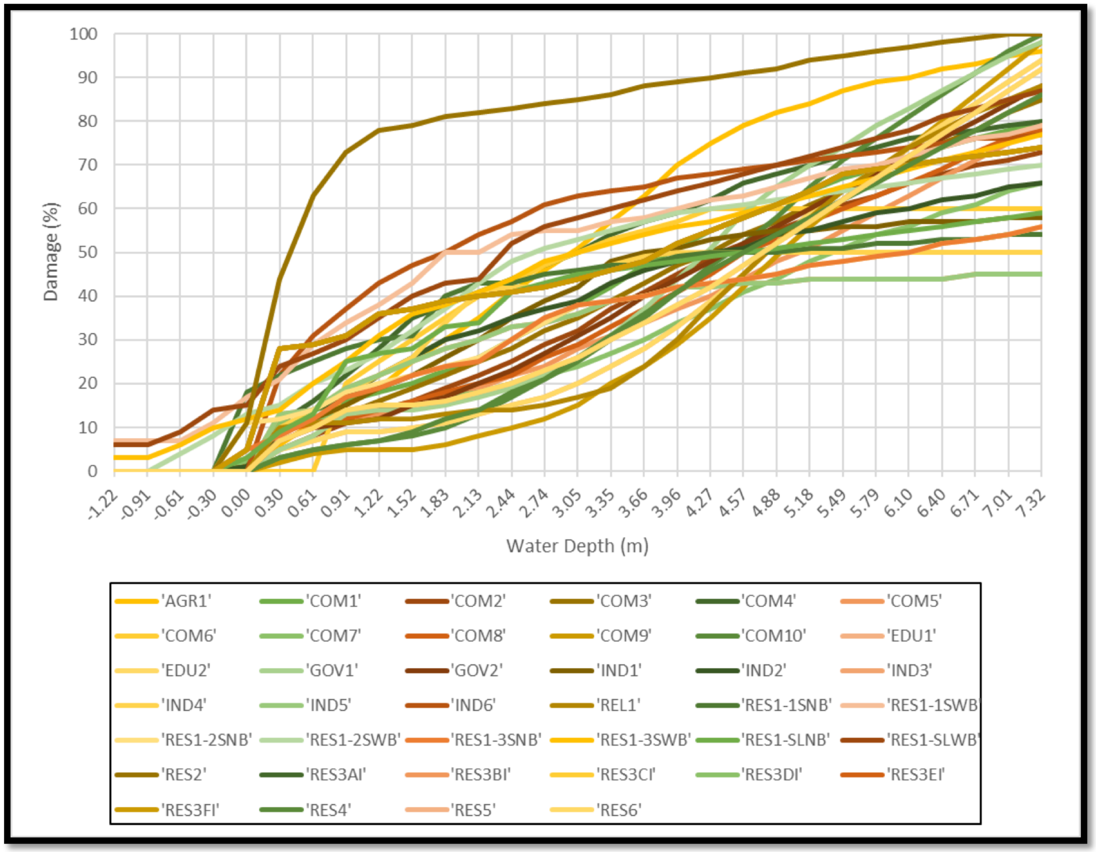
Future projections



EPA (2016), Updates to the Demographic and Spatial Allocation Models to Produce Integrated Climate and Land Use Scenarios (ICLUS) Version 2. EPA/600/R-16/366F, National Center for Environmental Assessment, Washington, DC.

- EPA Integrated Climate and Land Use Scenarios (ICLUS) project
- Projects population and land use change up to 2100:
  - SSP2 = tracks US census projection
  - SSP5 = high growth case

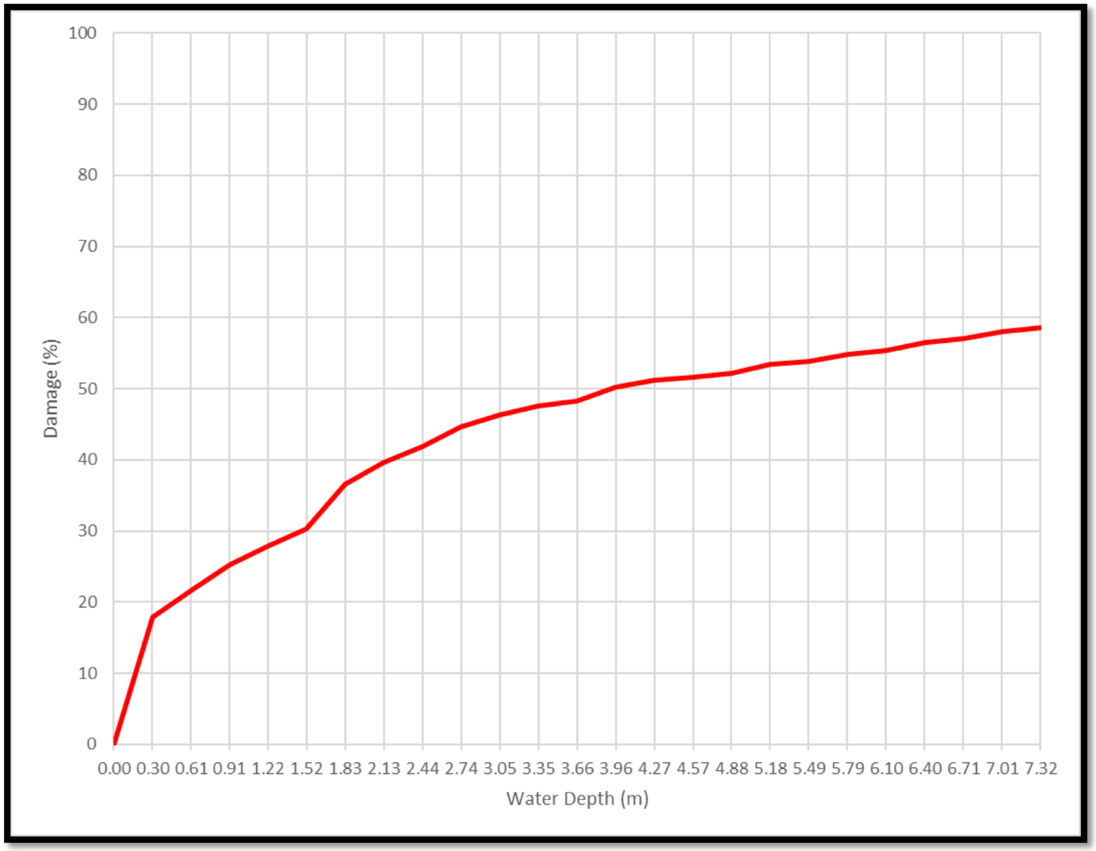
Vulnerability functions



- Simple relationship between water depth and % damage to an asset obtained from USACE

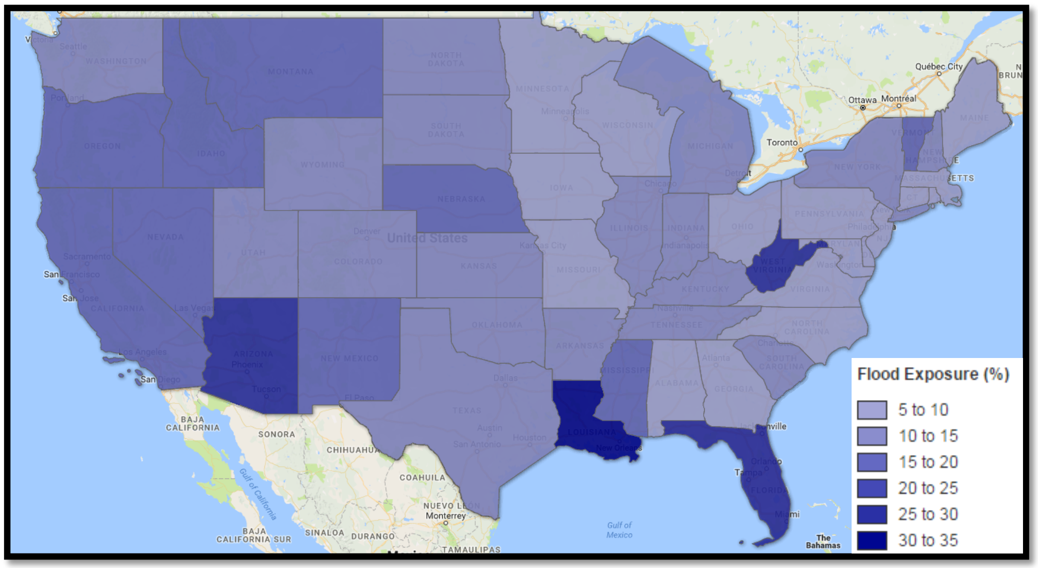
- Specific function for each building type when used with FEMA NSI

Future vulnerability functions



- Simple relationship between water depth and % damage to an asset obtained from USACE
- Generalised curve used for ICLUS projections

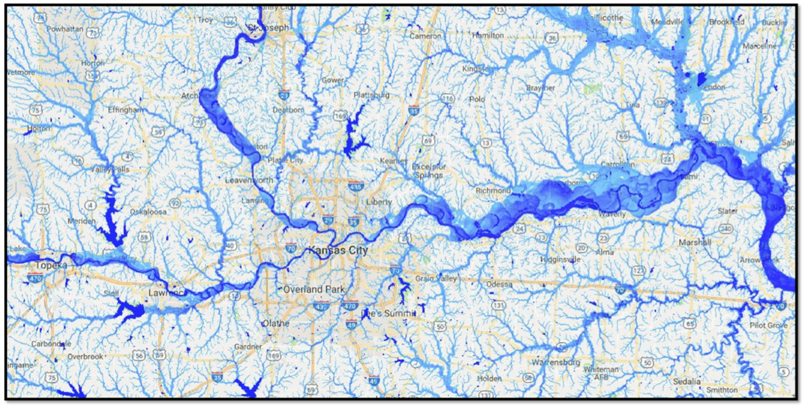
Population exposure

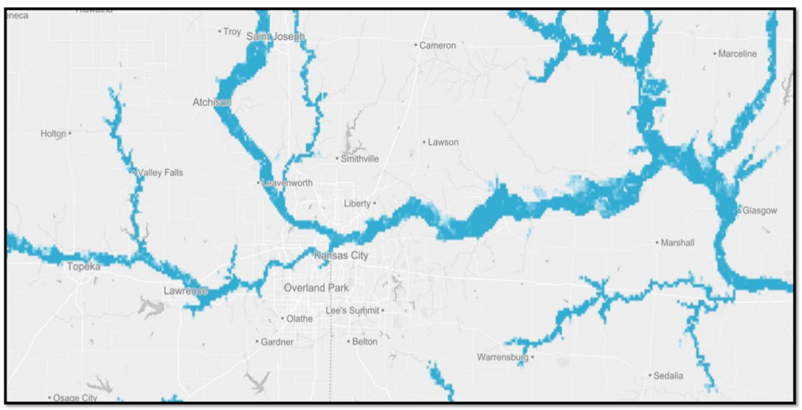


- Present-day

Return Period	Exposure (millions)	Exposure (%)
1 in 50	33.5	11.0
1 in 100	**40.8**	**13.3**
1 in 500	61.4	20.0
FEMA (1 in 100)	13.0	4.2
Aqueduct (1 in 100)	15.7	5.1

Population exposure





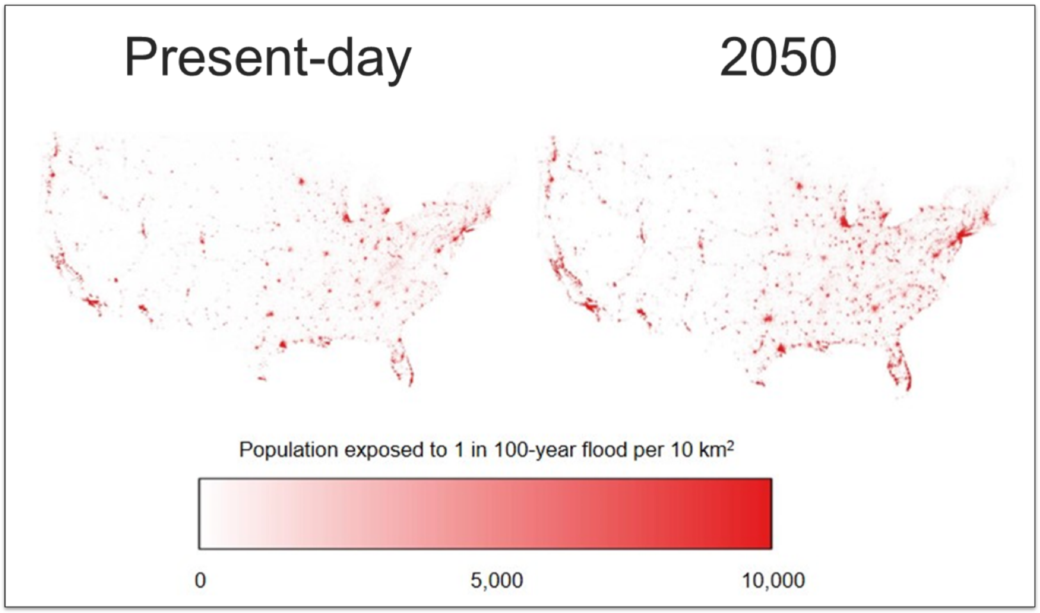
Return Period	Exposure (millions)	Exposure (%)
1 in 100	40.8	13.3
FEMA (1 in 100)	13.0	4.2
Aqueduct (1 in 100)	15.7	5.1

www.floods.wri.org

Winsemius, H. C. *et al.* (2013), A framework for global river flood risk assessments, *Hydrol. Earth Syst. Sci.*, 17, 1871–1892, doi:10.5194/hess-17-1871-2013.

Ward, P. J. *et al.* (2013), Assessing flood risk at the global scale: model setup, results, and sensitivity, *Environ. Res. Lett.*, 8, doi:10.1088/1748-9326/8/4/044019.

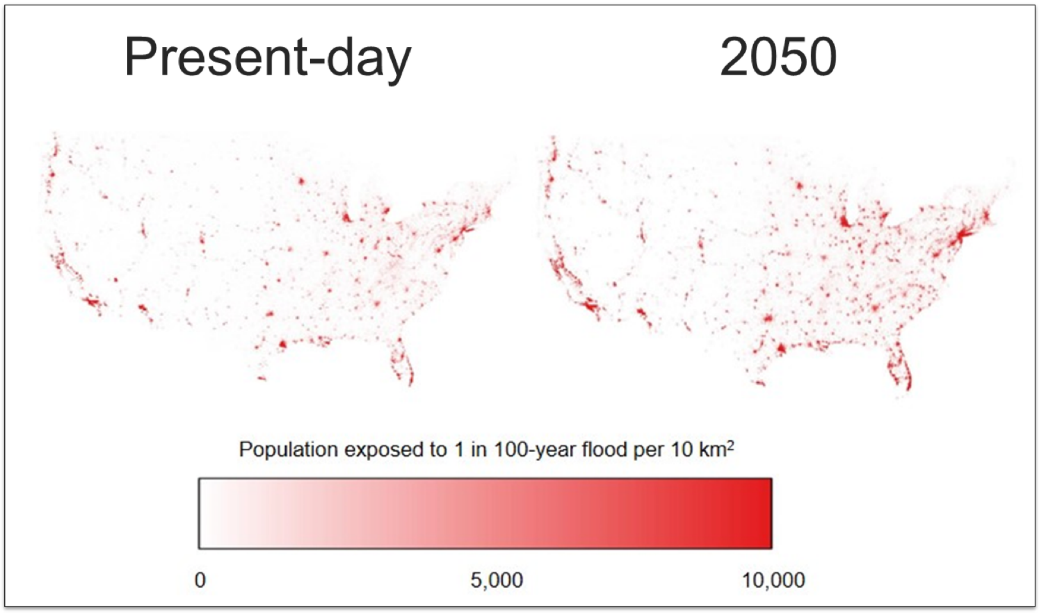
Population exposure



- SSP2 2050

Return Period	Exposure (millions)	Exposure (%)
1 in 50	51.3	13.1
1 in 100	**61.2**	**15.6**
1 in 500	86.8	22.2

Population exposure

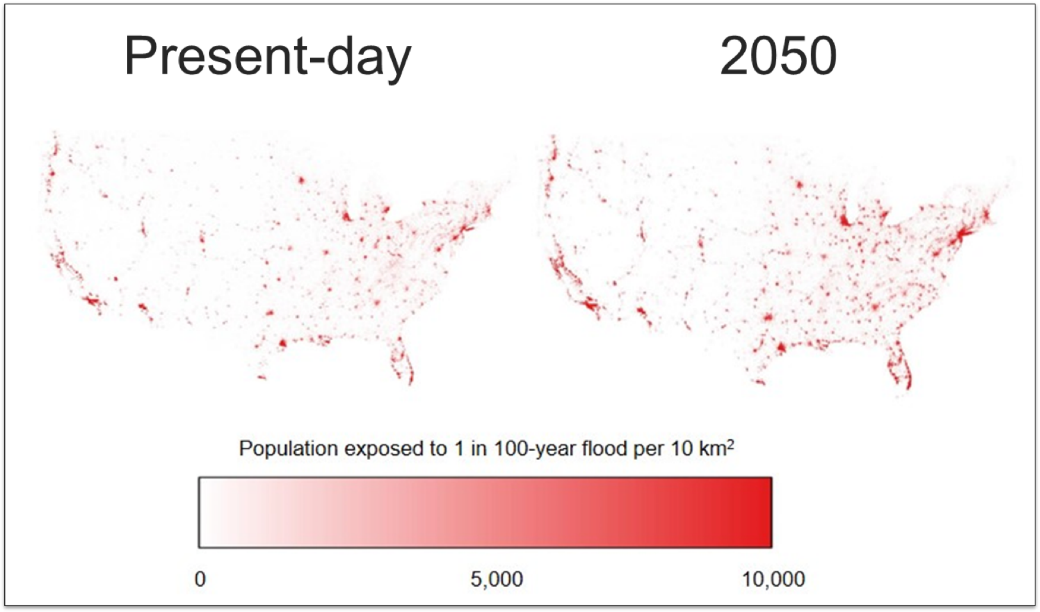


- SSP2 2050

Return Period	Exposure (millions)	Exposure (%)
1 in 50	51.3	13.1
1 in 100	**61.2**	**15.6**
1 in 500	86.8	22.2

- Changes from present-day (100-year floodplain)
  - 20.4M (↑ 50%)

Population exposure

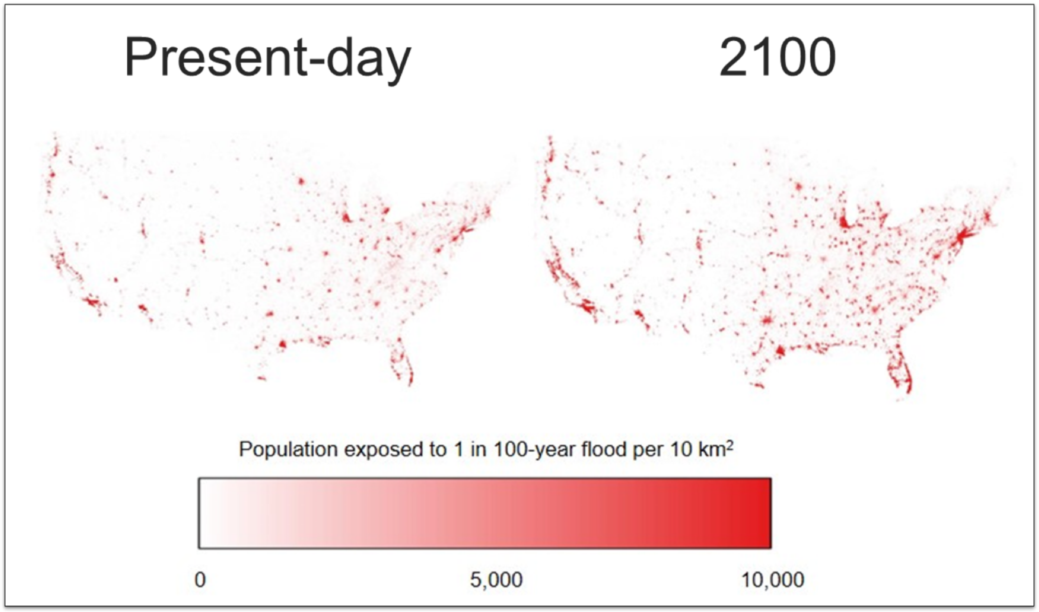


- SSP2 2050

Return Period	Exposure (millions)	Exposure (%)
1 in 50	51.3	13.1
1 in 100	**61.2**	**15.6**
1 in 500	86.8	22.2

- Changes from present-day (100-year floodplain)
  - 2.3pp ↑ in proportion of total population

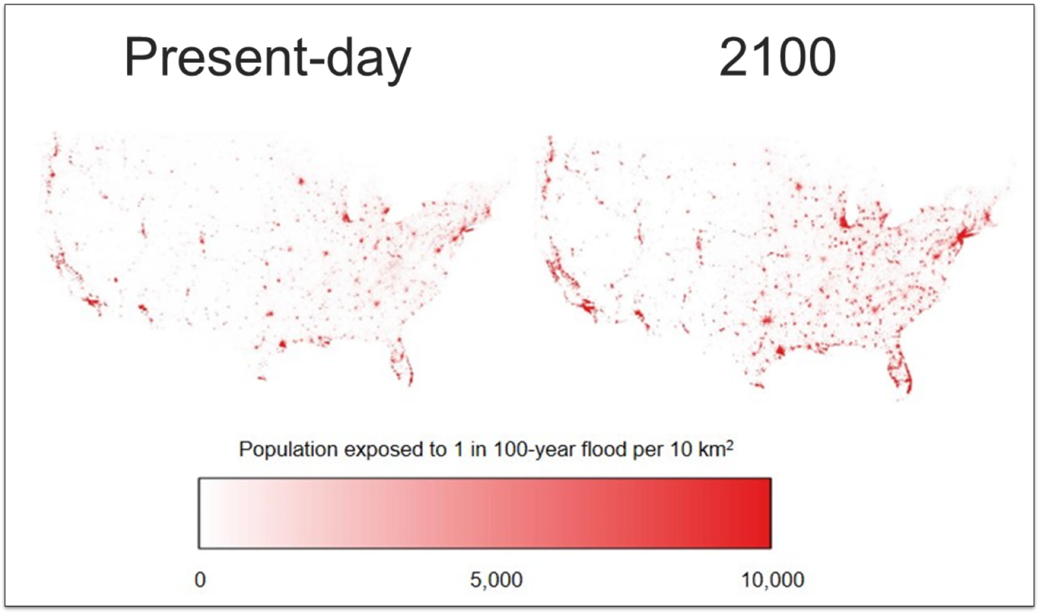
Population exposure



- SSP2 2100

Return Period	Exposure (millions)	Exposure (%)
1 in 50	63.1	13.9
1 in 100	**74.8**	**16.4**
1 in 500	104.5	23.0

Population exposure

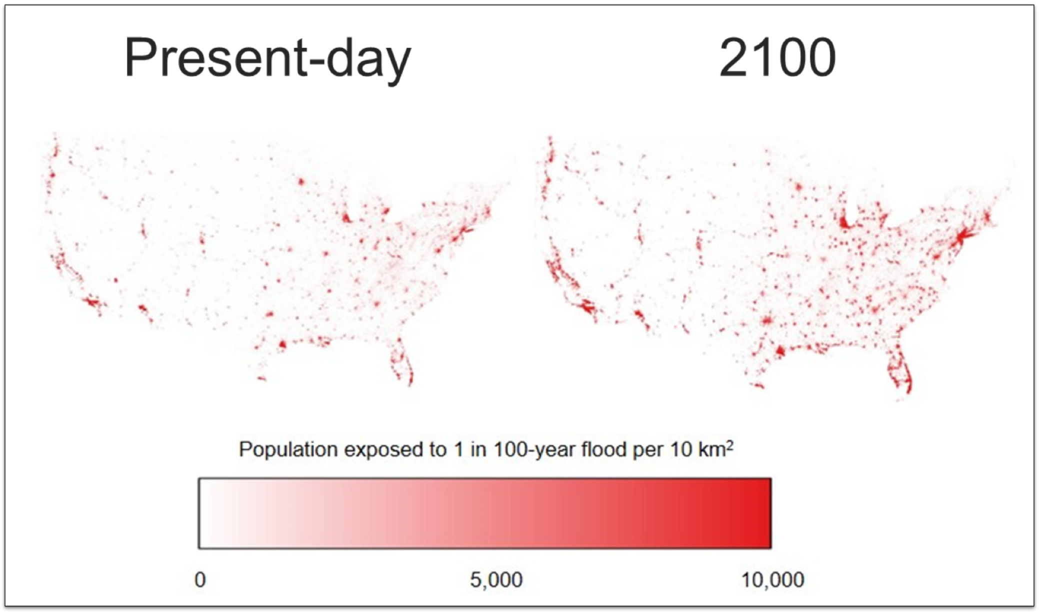


- SSP2 2100

Return Period	Exposure (millions)	Exposure (%)
1 in 50	63.1	13.9
1 in 100	**74.8**	**16.4**
1 in 500	104.5	23.0

- Changes from present-day (100-year floodplain)
  - 34.0M (↑ 83%)

Population exposure

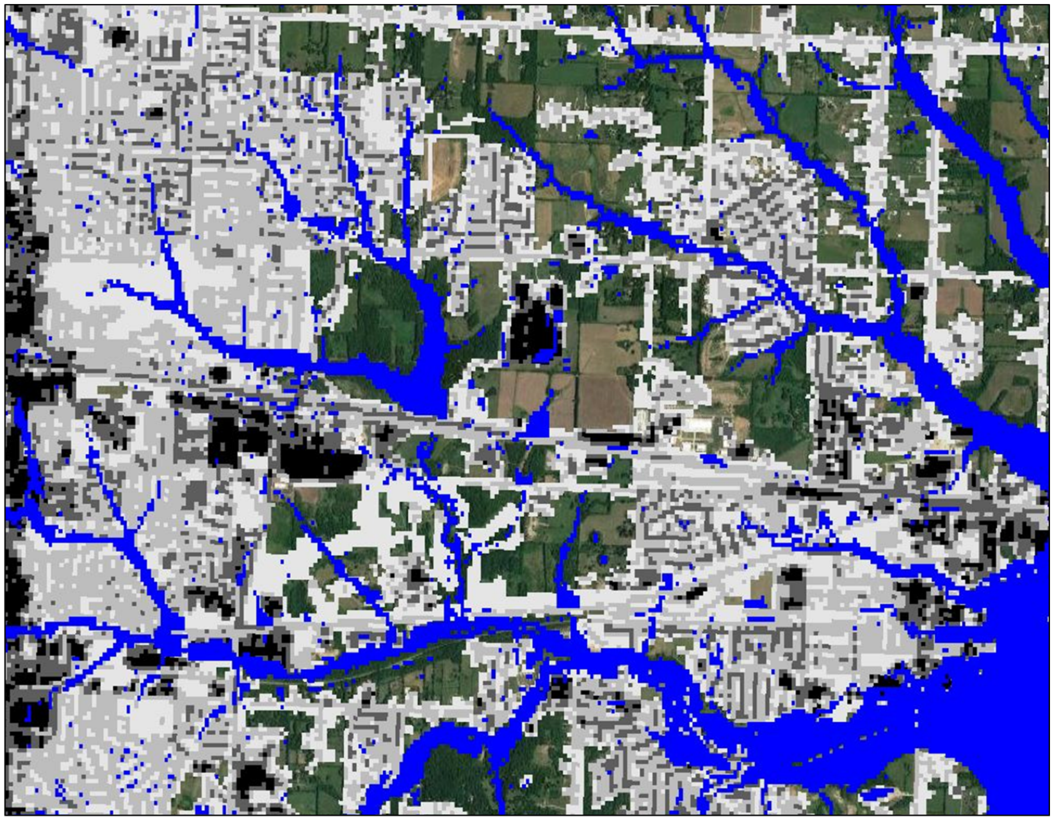


- SSP2 2100

Return Period	Exposure (millions)	Exposure (%)
1 in 50	63.1	13.9
1 in 100	**74.8**	**16.4**
1 in 500	104.5	23.0

- Changes from present-day (100-year floodplain)
  - 3.1pp ↑ in proportion of total population

Asset exposure and risk

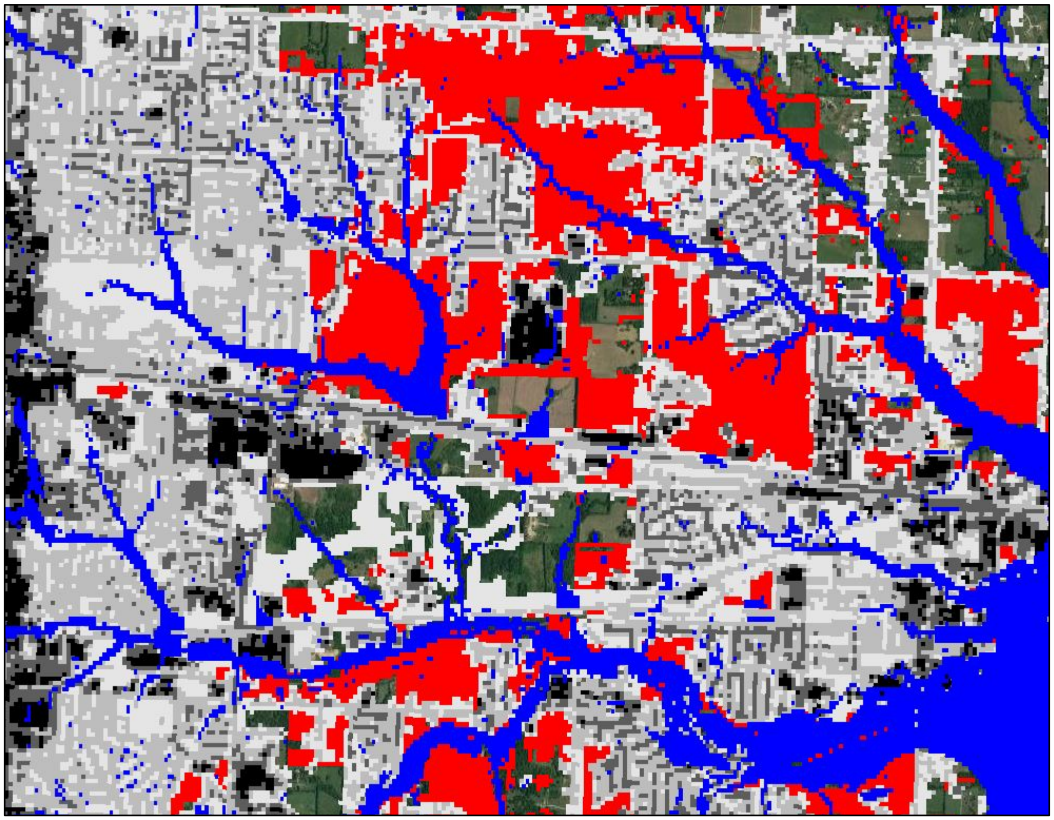


- Present-day

Return Period	Dev. Area (km^2)	Exposed assets ($tn)	Potential dmg ($tn)
1 in 50	140,657	4.6	0.9
1 in 100	**157,430**	**5.5**	**1.2**
1 in 500	203,775	8.2	1.9

- 100-year developed floodplain = approx. the land area of **Georgia**

Asset exposure and risk

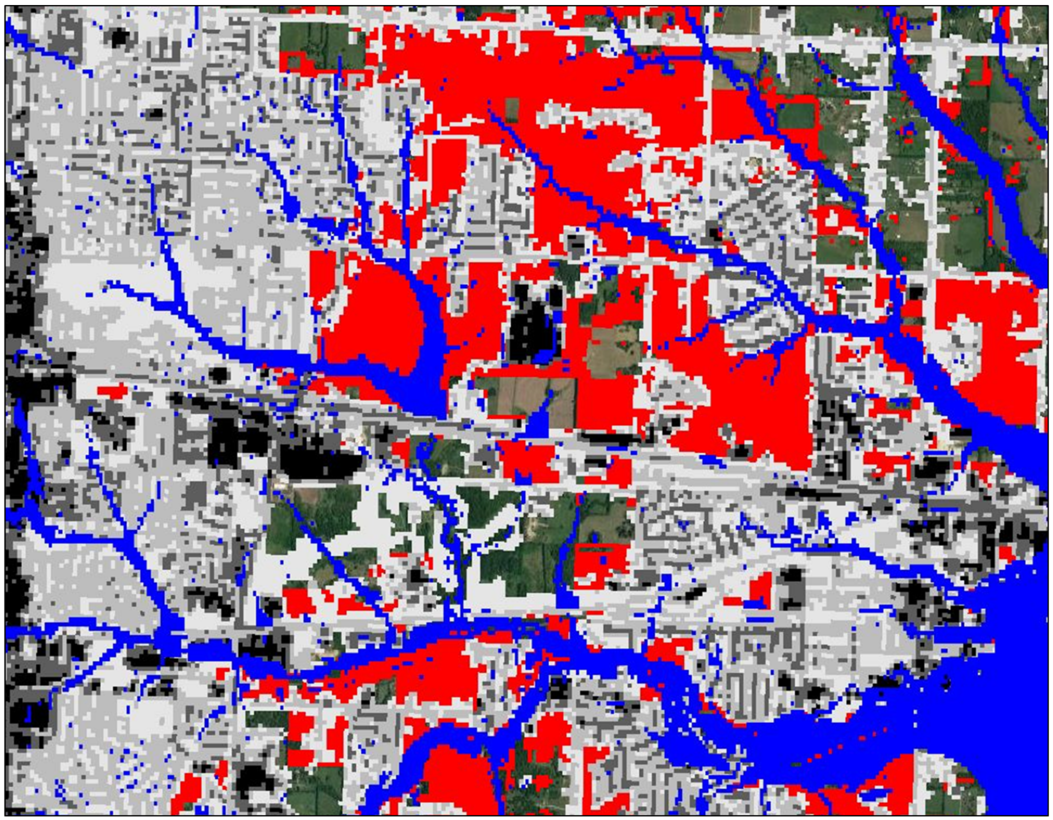


- SSP2 2050

Return Period	Dev. Area (km²)	Exposed assets ($tn)	Potential dmg ($tn)
1 in 50	174,989	6.9	1.5
1 in 100	**195,981**	**8.1**	**1.7**
1 in 500	251,702	11.3	2.7

- 100-year developed floodplain = approx. the land area of **S. Dakota**

Asset exposure and risk

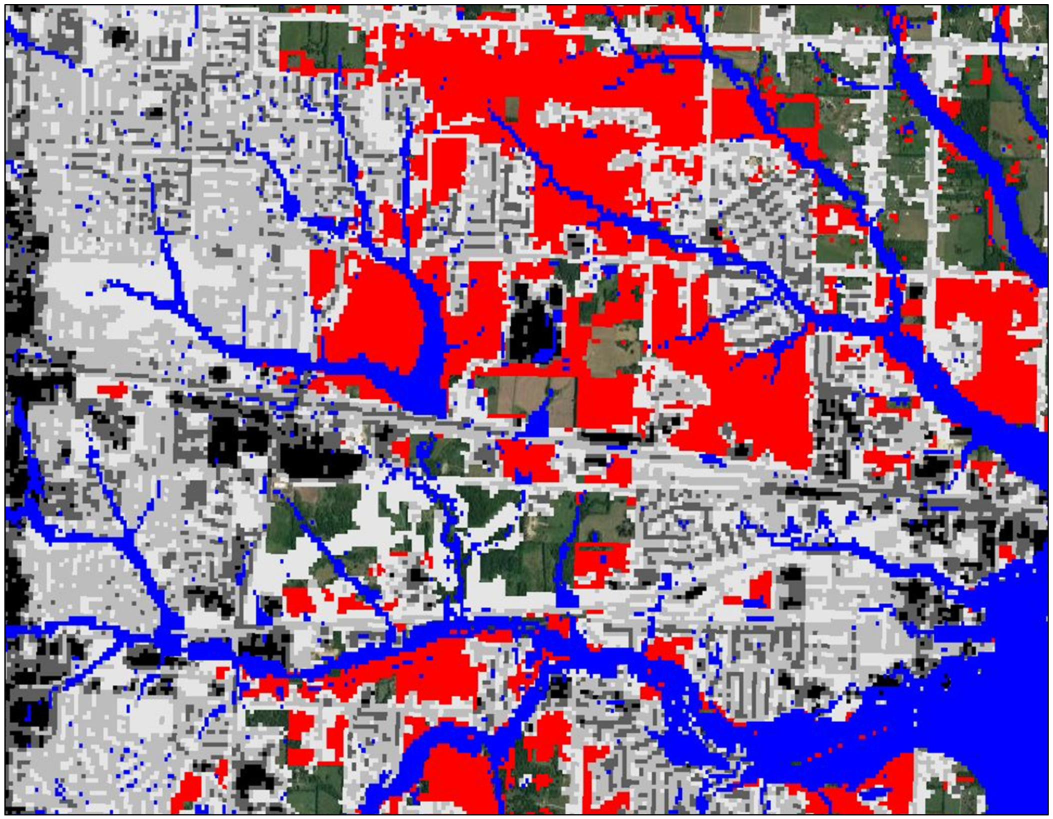


- SSP2 2100

Return Period	Dev. Area (km^2)	Exposed assets ($tn)	Potential dmg ($tn)
1 in 50	192,417	8.3	1.7
1 in 100	**215,900**	**9.8**	**2.1**
1 in 500	276,956	13.6	3.2

- 100-year developed floodplain = approx. the land area of **Kansas**

Asset exposure and risk



- SSP2 2100

Return Period	Dev. Area (km²)	Exposed assets ($tn)	Potential dmg ($tn)
1 in 50	192,417	8.3	1.7
1 in 100	**215,900**	**9.8**	**2.1**
1 in 500	276,956	13.6	3.2

- 100-year newly developed floodplain = approx. the land area of **W. Virginia**

Conclusions

- Developed a whole US flood inundation model with skill approaching that local bespoke simulations
- Intersecting simulations with high resolution population data shows exposed population and assets are ~3x higher than previous estimates
- Socio-economic change alone will increase the proportion of US population at risk during the C21st
- Climate change will undoubtedly amplify these effects further
- Now need to move away from ‘constant return period in space’ hazard layers to properly estimate flood risk

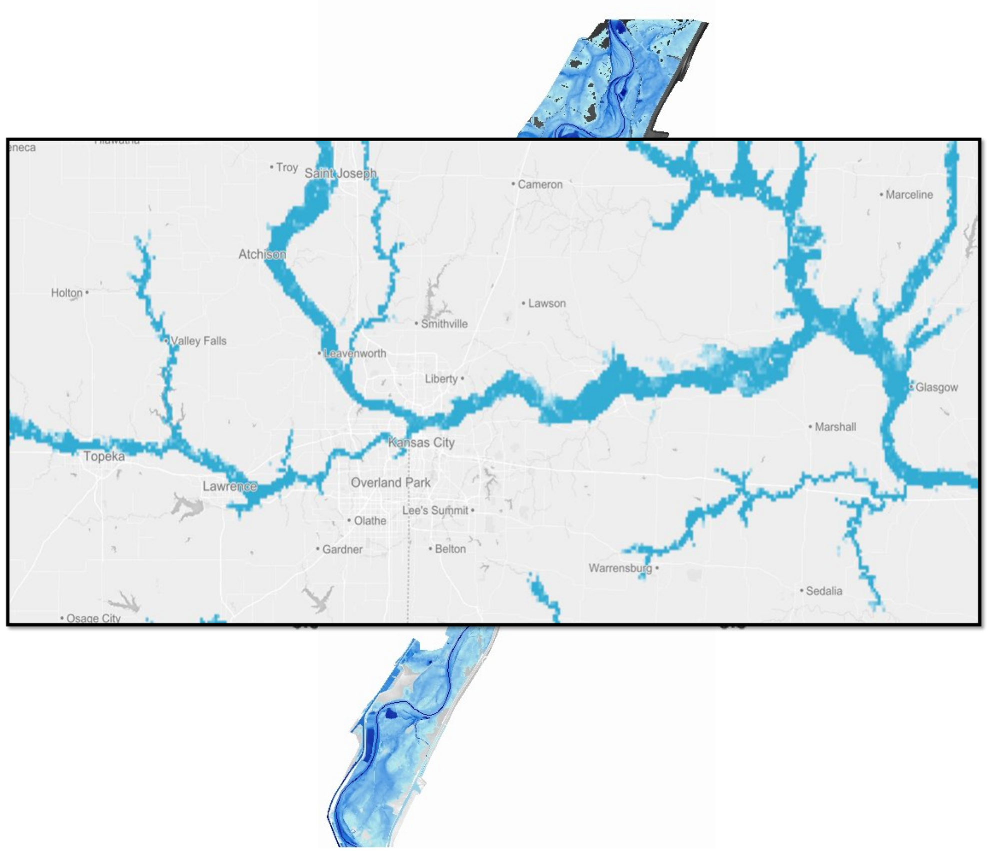
Spare slides

Catchment-scale validation: USGS 1D models

LOCATION	HIT RATE (%)				
	1 in 5	1 in 10	1 in 50	1 in 100	1 in 500
Battle Creek, MI	**-**	**81.9**	**84.3**	**88.1**	**90.7**
Harrisburg, PA	**-**	**95.1**	**92.3**	**88.7**	**86.1**
Killbuck, OH	**78.1**	**81.8**	**90.2**	**91.9**	**94.3**

Traditional flood hazard assessment

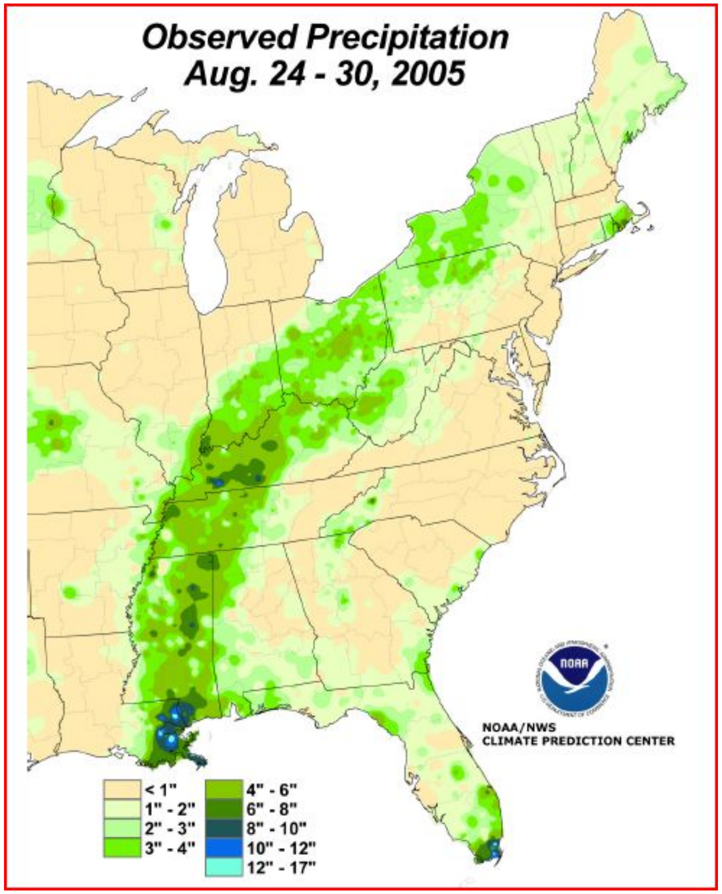
- Return period flows from gauge data
- Reach scale hydraulic models
- Large scale hazard maps then ‘stitch together’ the results of many local studies
  - E.g. FEMA Special Flood Hazard Areas



- Spatially invariant return period assumption breaks down at large scales

Event footprints

- Over large scales the event return period varies in space
  - Known as the ‘flood footprint’
- Often ignored in many large scale analyses
- Gauge time series not long enough to sample all possible footprints
- Use conditional stochastic simulation to generate a bigger sample of plausible event footprints

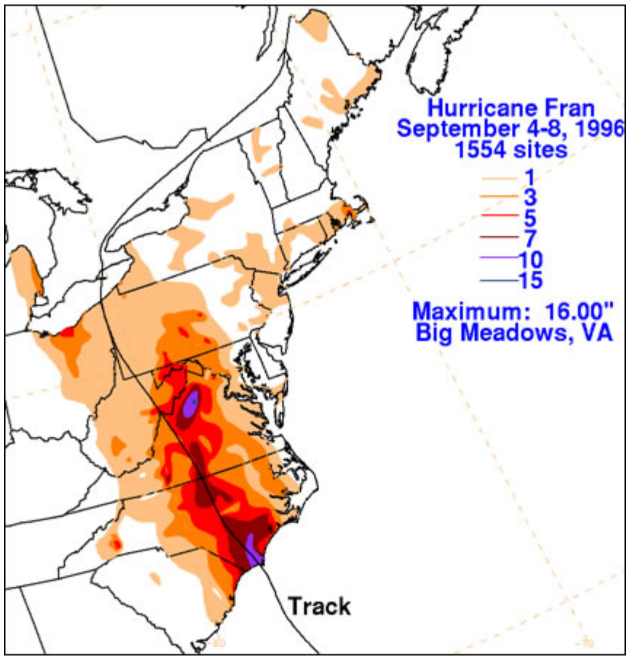


The problem

- For a T year return period flow at gauging site X, Q_X^T, what is the probability distribution of flow at gauge Y, i.e. $Pr(Q_y | Q_X^T)$ for all $Y_{1,2 \ldots\ldots n}$

- Multi-site conditional probability statistical methods are well known, but not previously applied at continental scales for thousands of gauges
  - Cross-correlation between all gauges gives a large compute problem
  - Large climatic differences at continental scales
  - Multiple flood generating mechanisms

Method



Heffernan, J.E. and Tawn, J.A. (2004). A conditional approach for multivariate extreme values. J. *R. Statist. Soc. B*, **66**, Part 3, 497–546.

Gauge Preparation

Import gauges

QC gauges

Infill missing data

Filter gauges

Spatio-temporal dependence

Marginal thresholds

- Define threshold
- Produce semi-parametric distribution

Dependence thresholds

- Define threshold
- Define multivariate dependence structure

Generate event catalogues

Define expected number of events per year within region

Utilise dependence information to simulate *N* event footprints

- Likelihood of a given gauge being the largest is used to define the number of times of condition upon that gauge

Generate areal flood depths and losses

Interpolation

- Point data to accumulation units

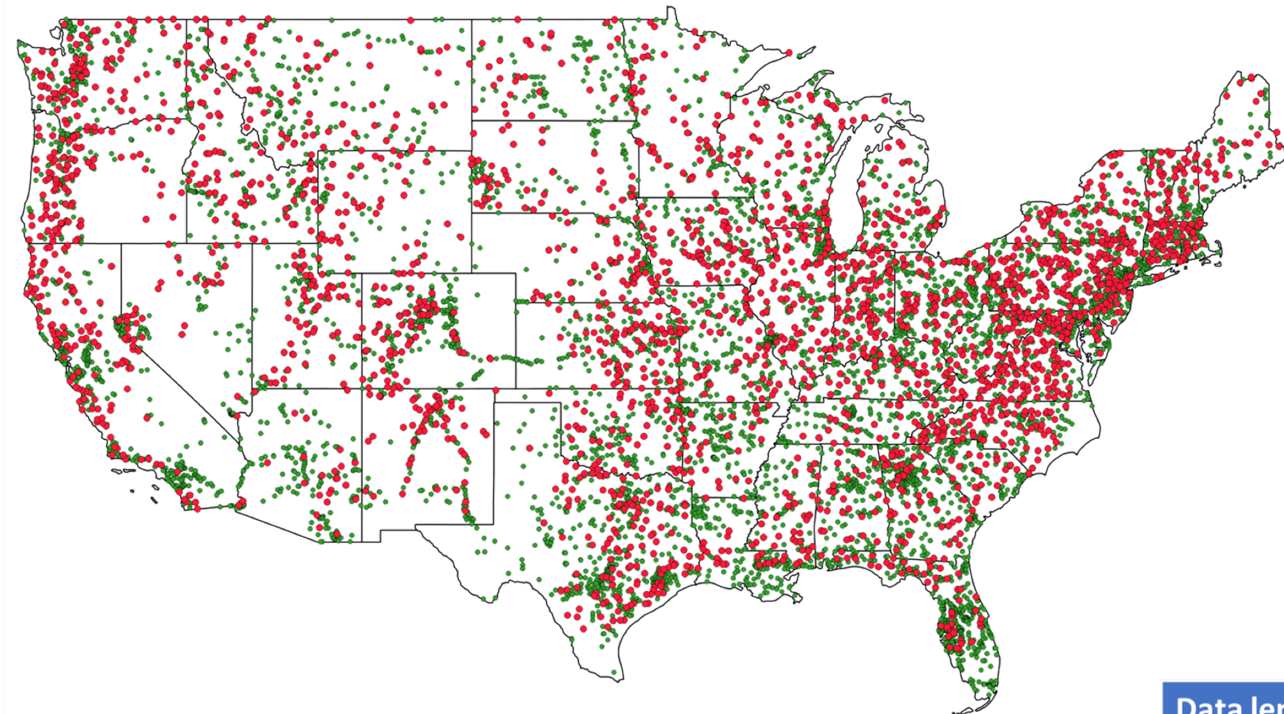
Extract Fathom hazard layers

- Corresponding to interpolated R period surfaces

Extract economic losses

- USA asset information and depth damage functions
- Overlay with extracted inundation extend / depth

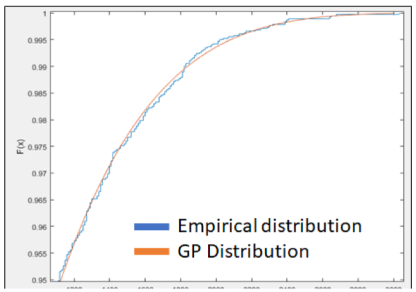
Gauge preparation

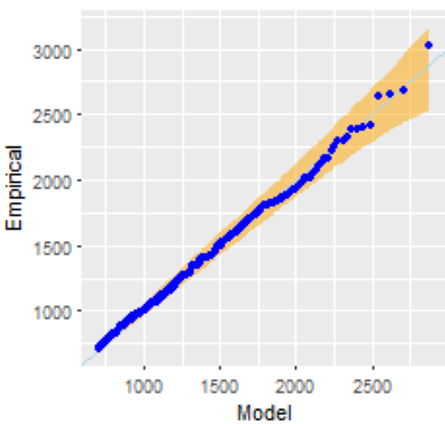


Data length (years)	# Suitable Gauges
>30	4450
>40	3625
>50 ✓	2800

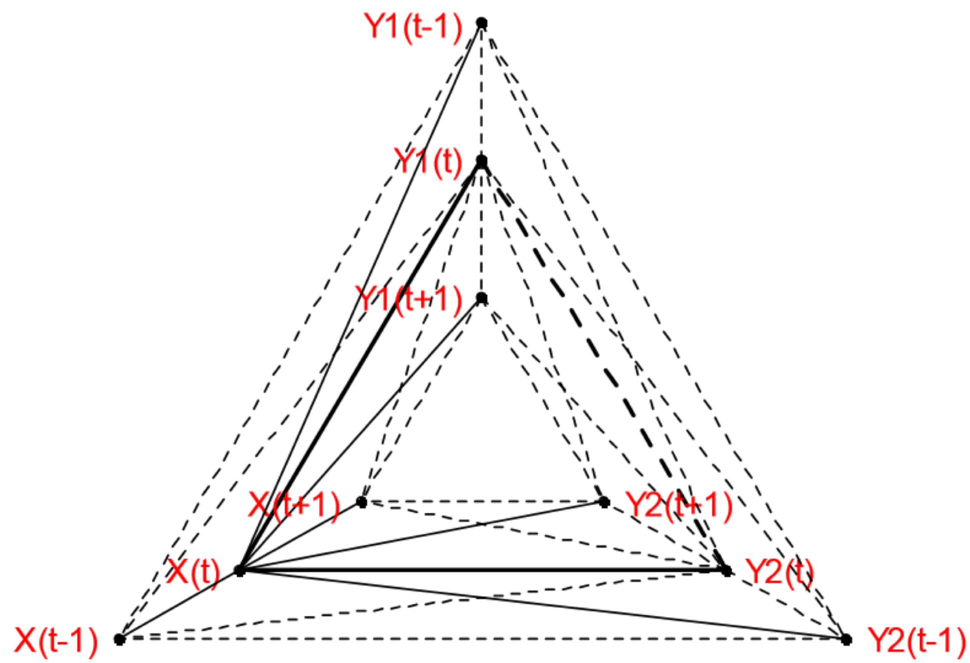
Spatio-temporal dependence

a. Define marginals at each gauge





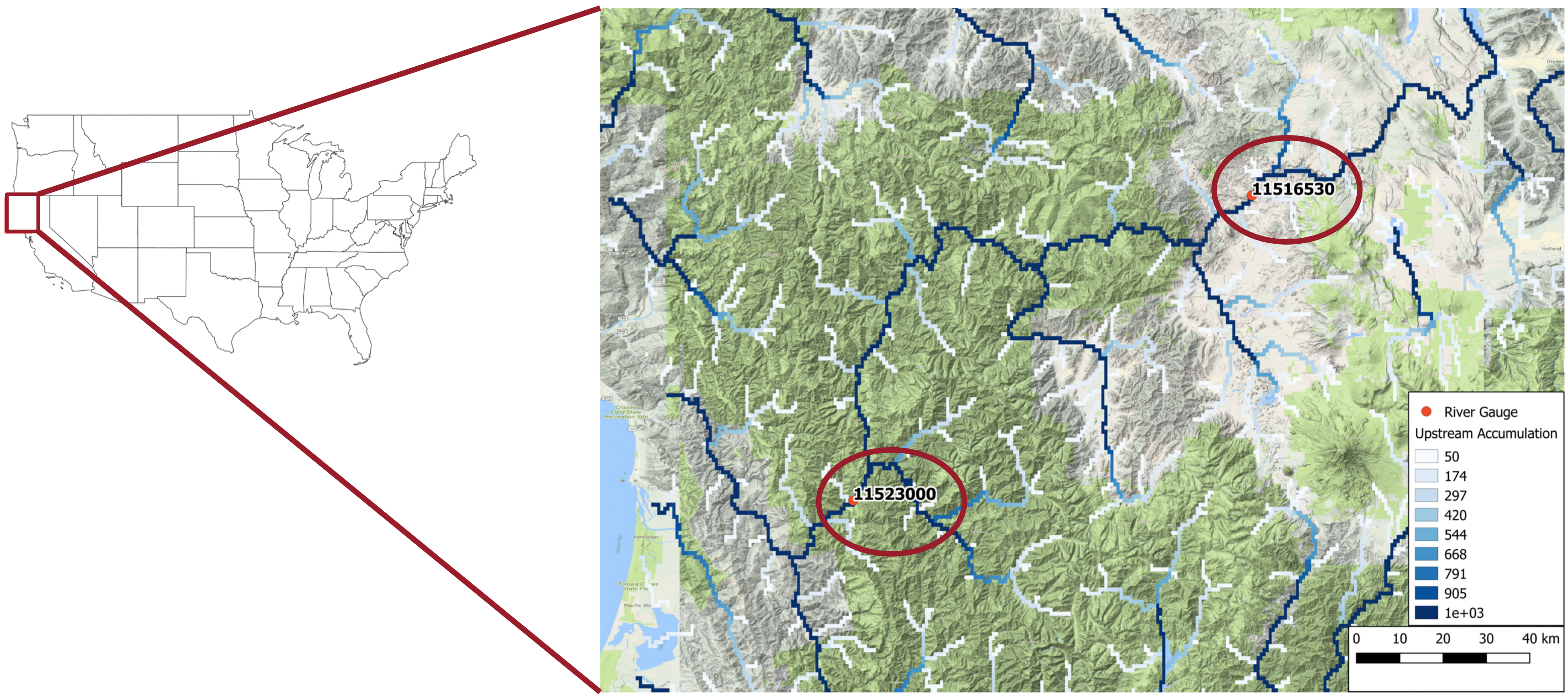
b. Compute spatio-temporal connections between gauges for flows Above Q99 threshold



See also poster H21J-1615 by Quinn et al

Spatio-temporal dependence

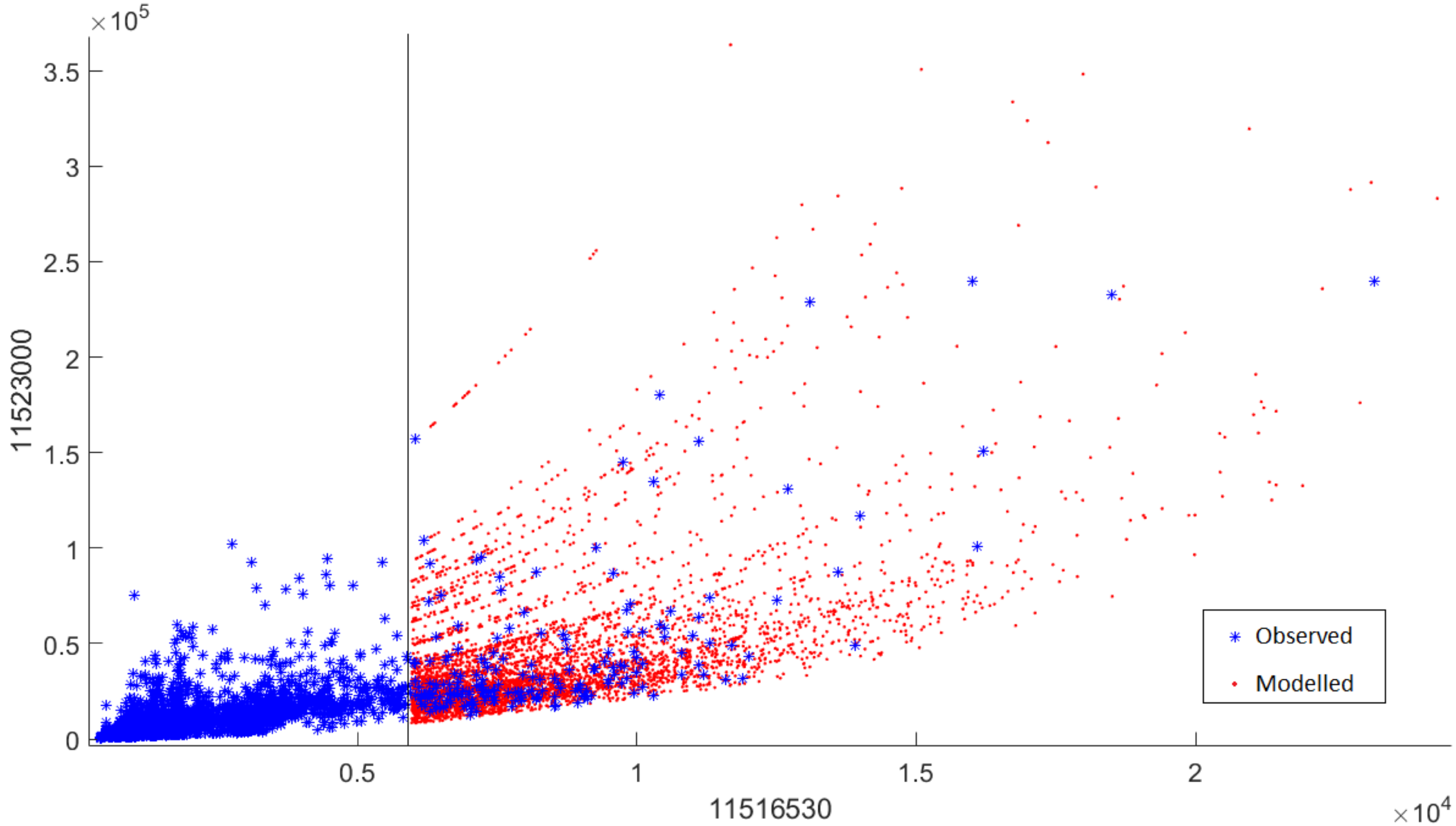
c. Define dependence structure



See also poster H21J-1615 by Quinn et al

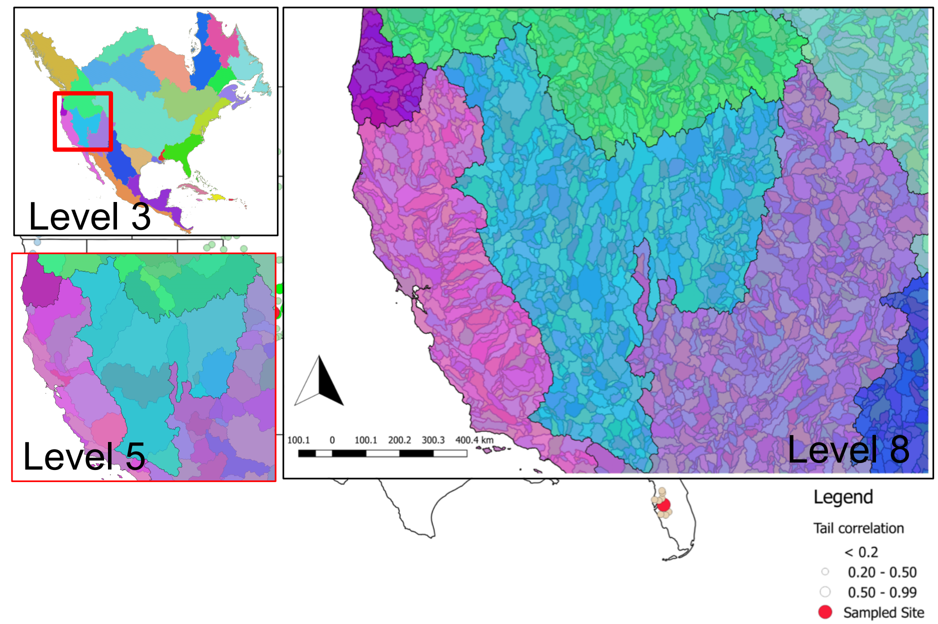
Spatio-temporal dependence

c. Define dependence structure

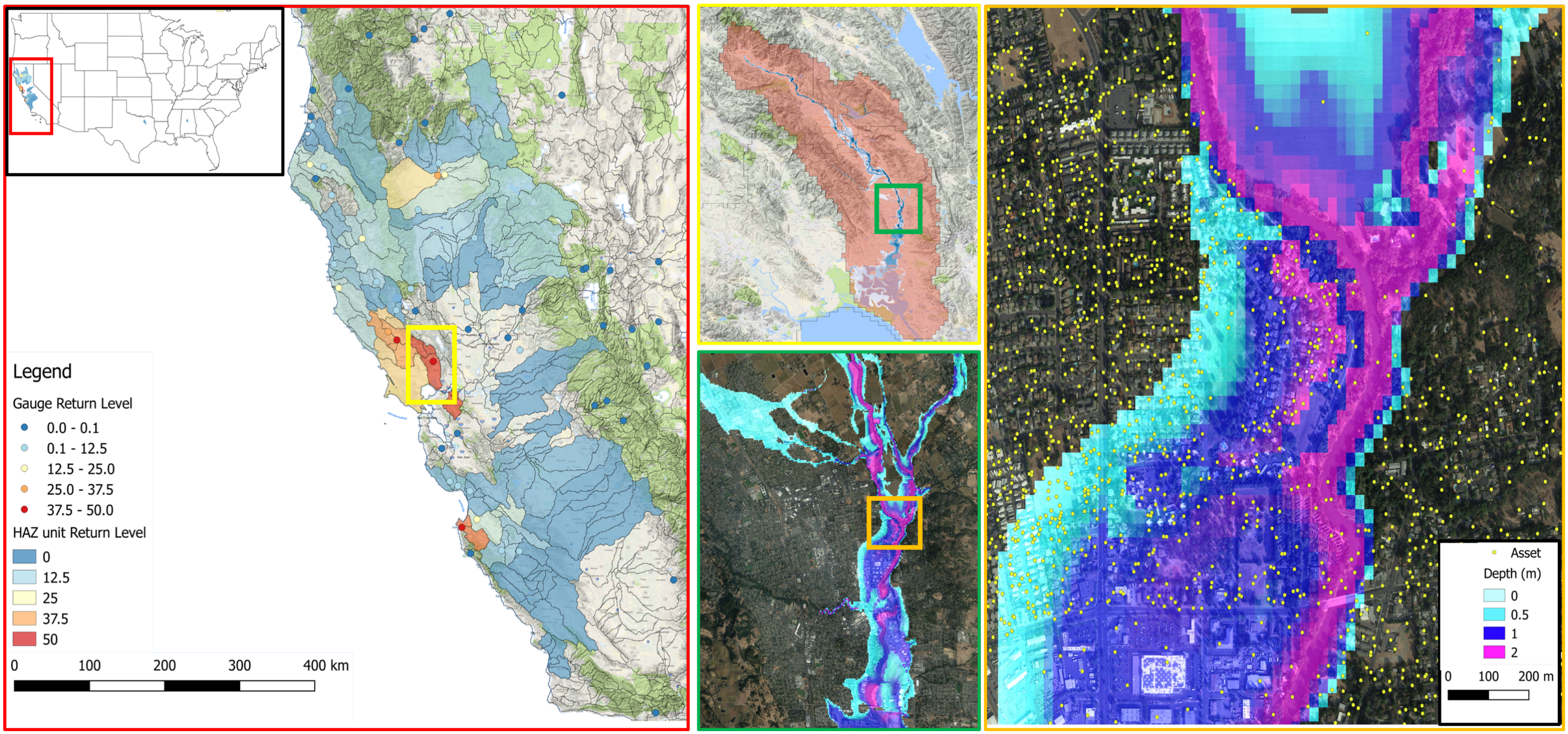


Generate event catalogues

- Given gauge dependence structure we can simulate event footprints at any given conditioning site
- Need a structure to interpolate over
  - Use HydroBasins Level 8 and 10 units
  - Interpolate gauge return period values to these units
  - Build footprint using pre-computed set of return period hazard layers from the Fathom Global US 30m hydraulic model

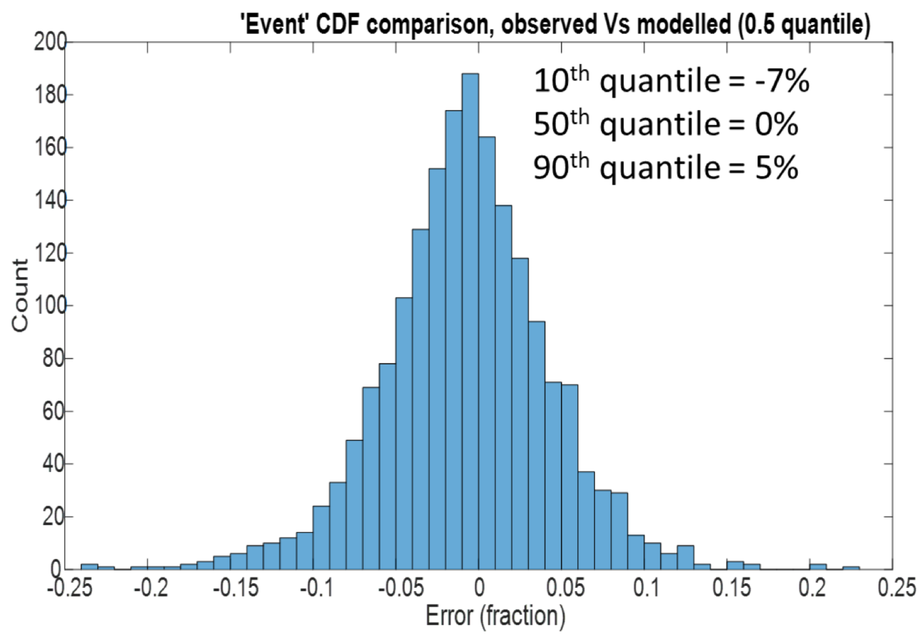


Example event footprint



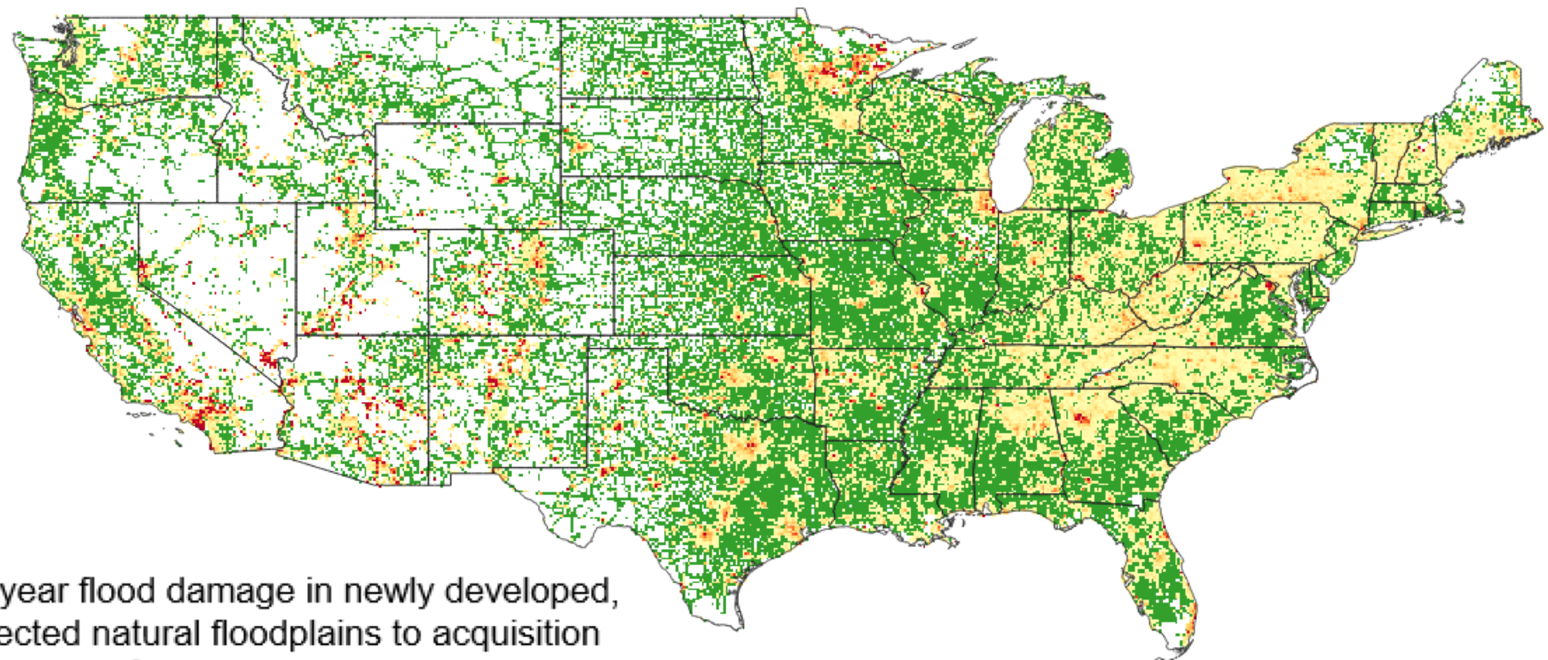
Validation

- Test 1: observed vs. modelled extreme value CDFs
  - For a given CDF quantile (0.5 in this case) does the dependence structure in the synthetic event ensemble match the observations?
- Test 2: Independent events
  - Extract all independent events > 1 in 5 year return period in USGS record
  - Simulate same record length many times using the stochastic method
  - Mean error in no. of events and no. of gauges hit is <5%



Event animation

Costs of avoidance



Ratio of 1 in 100-year flood damage in newly developed, currently unprotected natural floodplains to acquisition cost per 100 km² in 2050 under SSP2