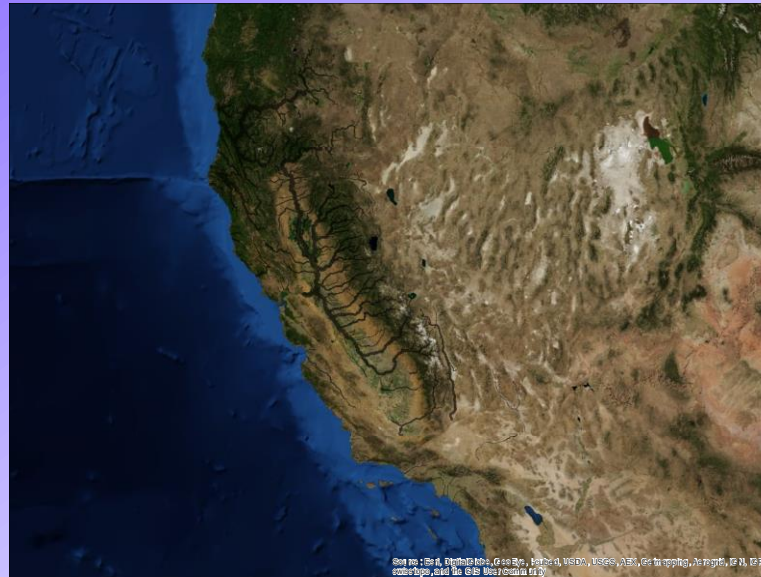


On the Use of Technology to Enable the Flow of Digital Rivers at the Continental-Scale



Cédric H. David

University of California Center for Hydrologic Modeling

NSF EarthCube Modeling for the Geosciences

Boulder, CO

23 April 2013

An EarthCube vision...



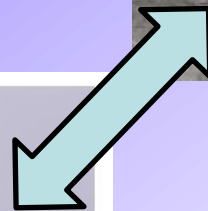
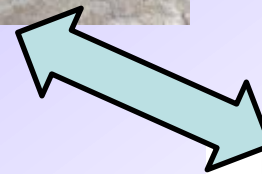
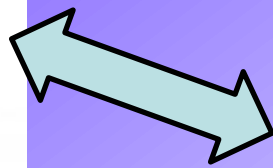
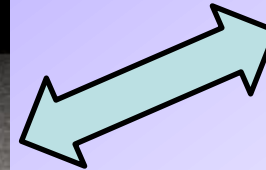
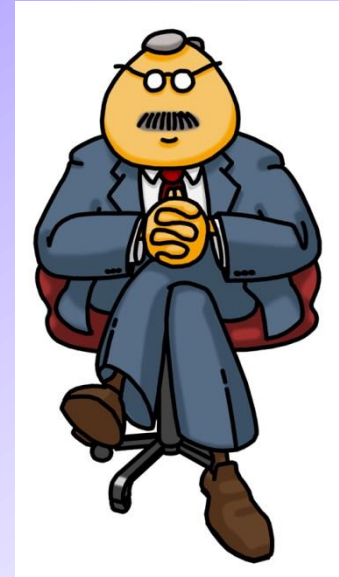
Making **all** geosciences accessible to **everybody**,
in an **exciting** way!!!

...not science fiction!

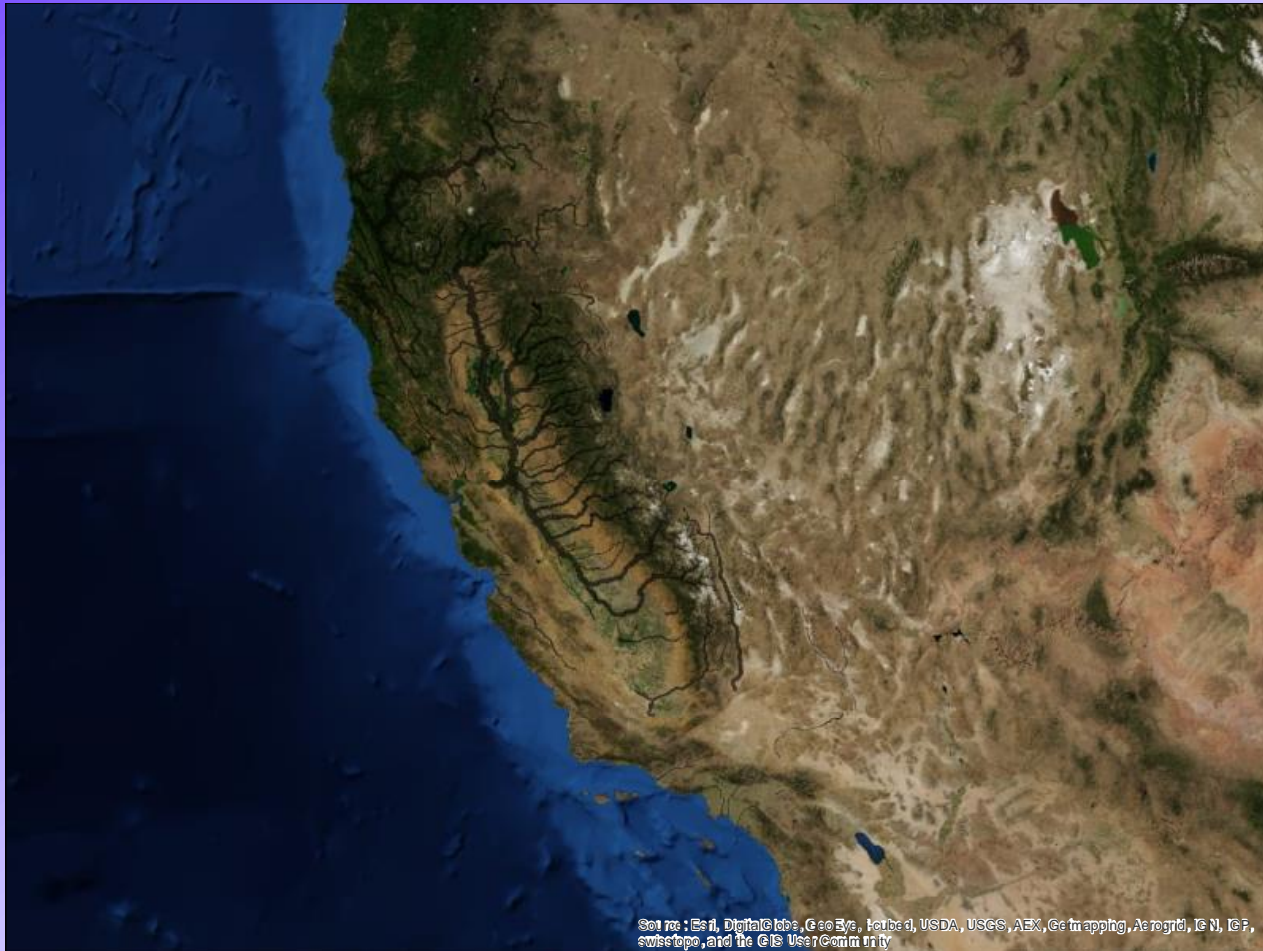


Holland Family Student Center – Univ. Texas at Austin

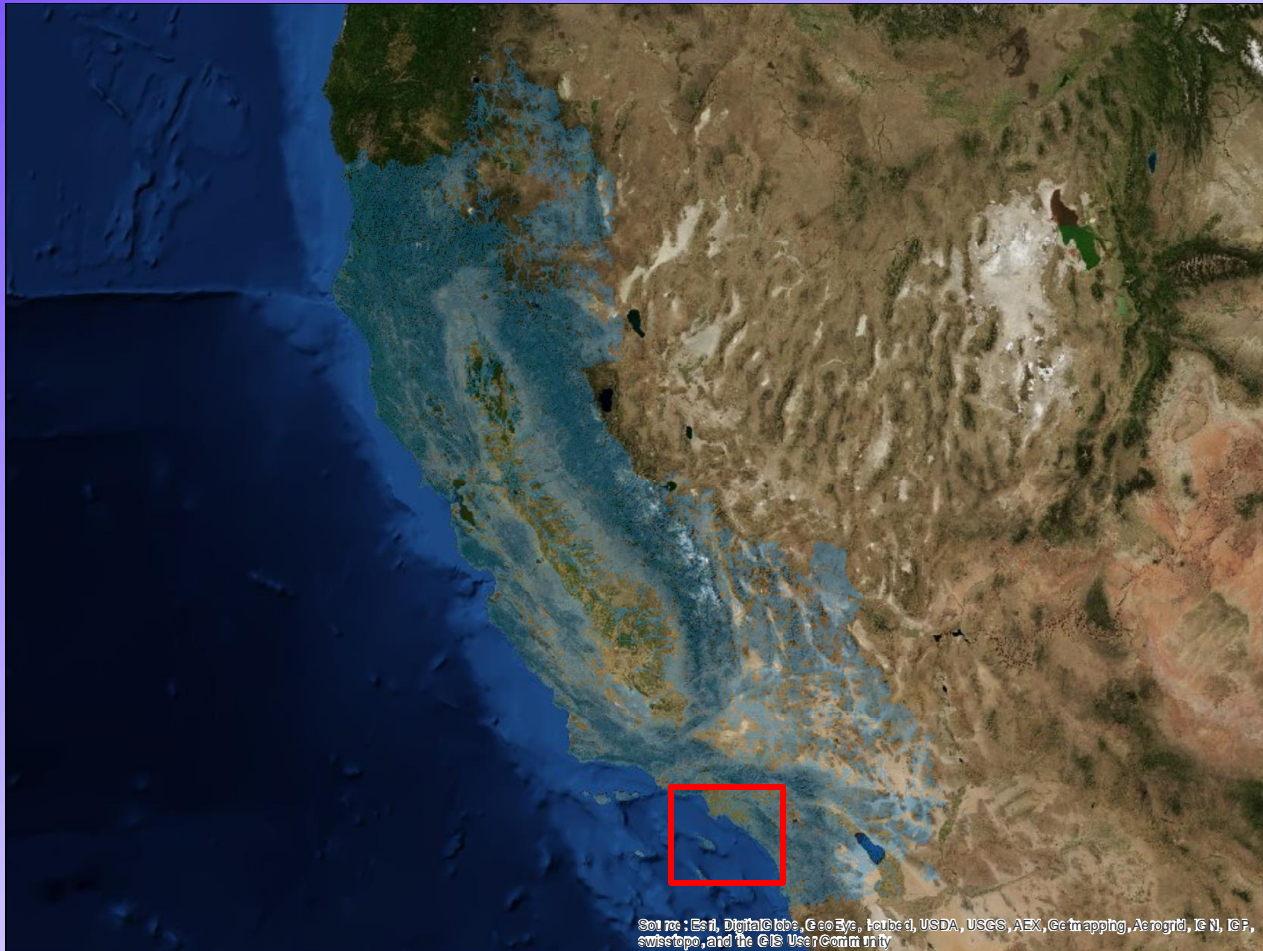
Help managers



Major rivers of California



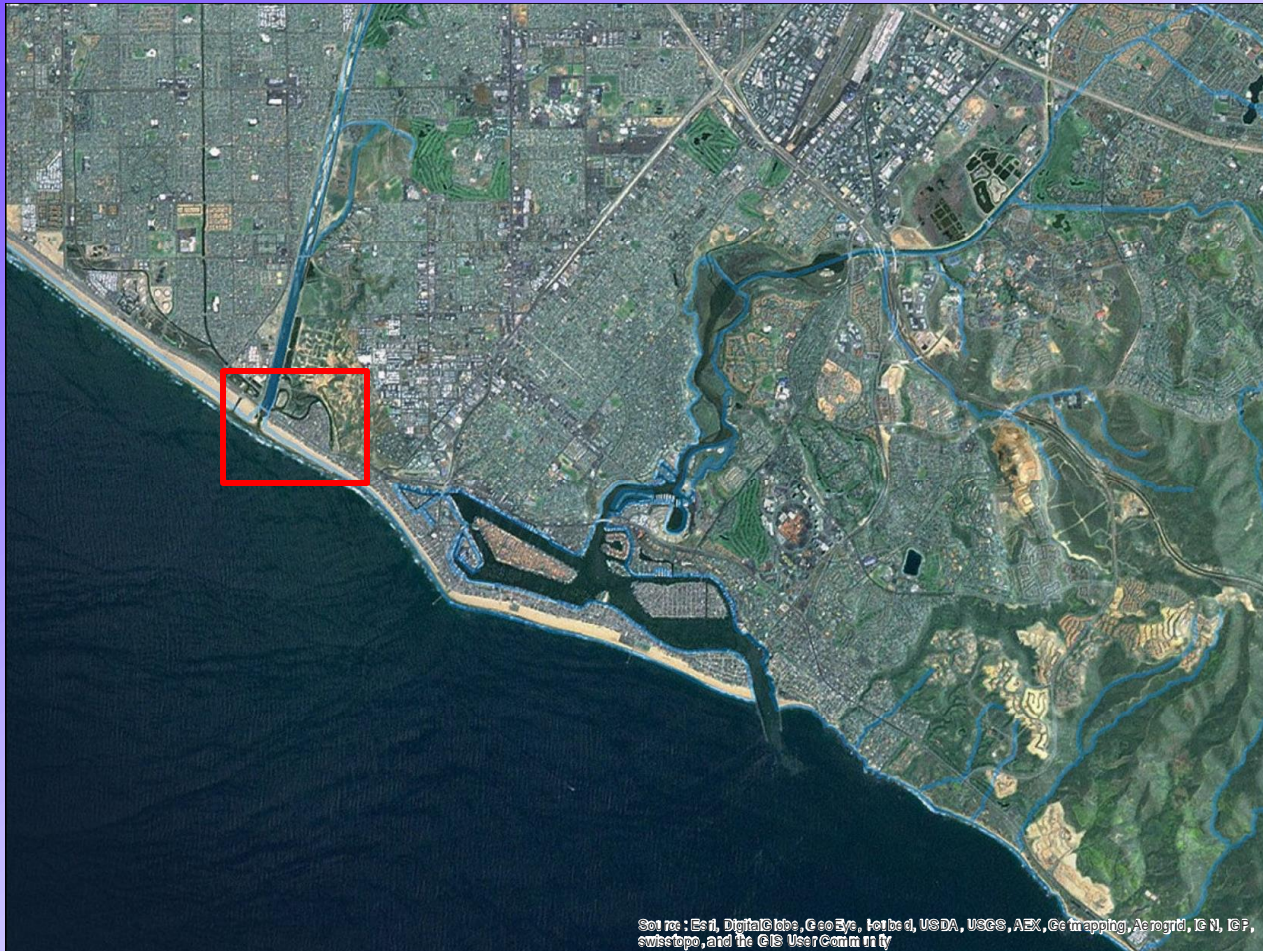
Mapped rivers of California



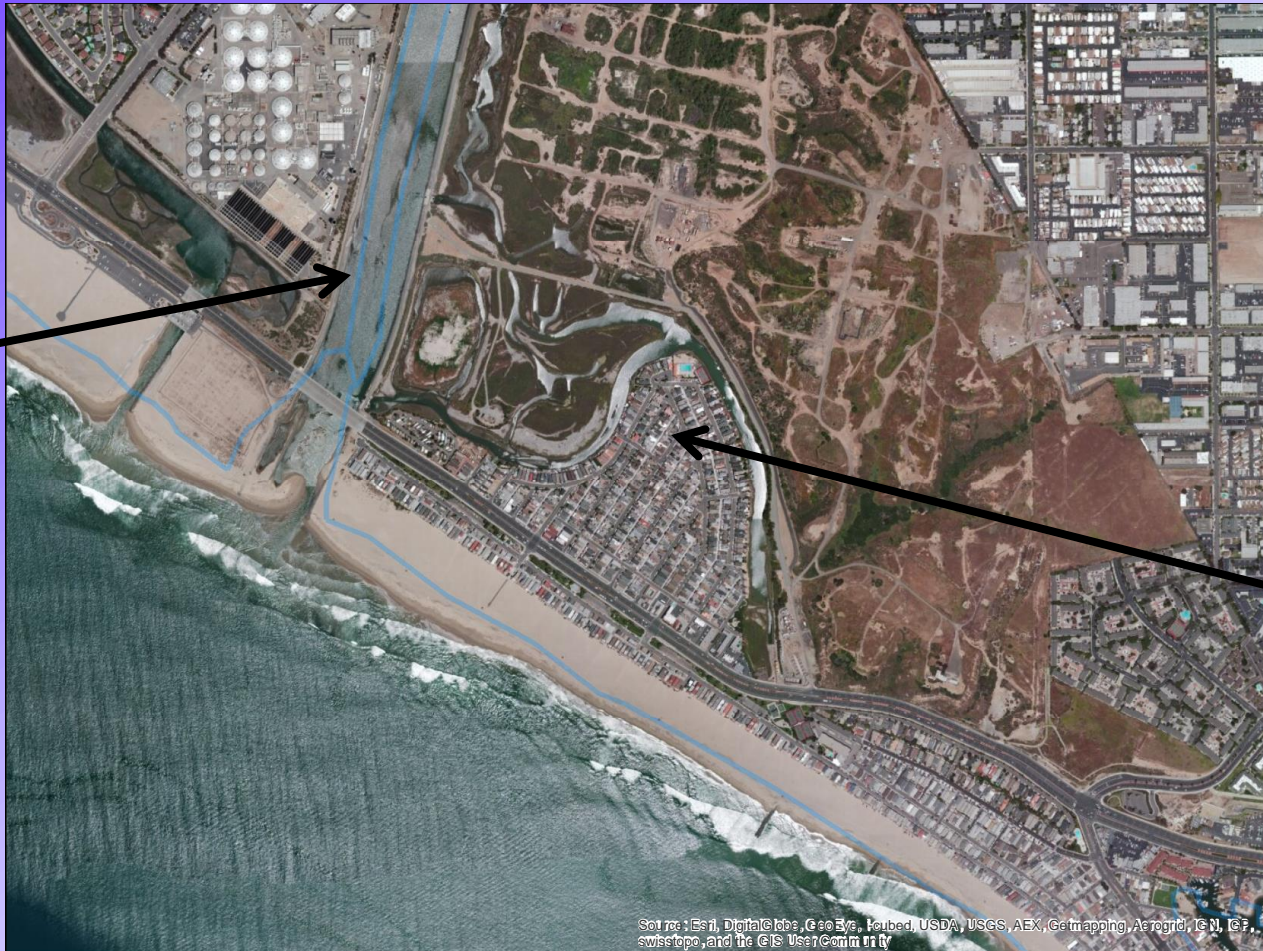
Mapped rivers of Orange County



Mapped rivers of Newport Beach



Santa Ana River

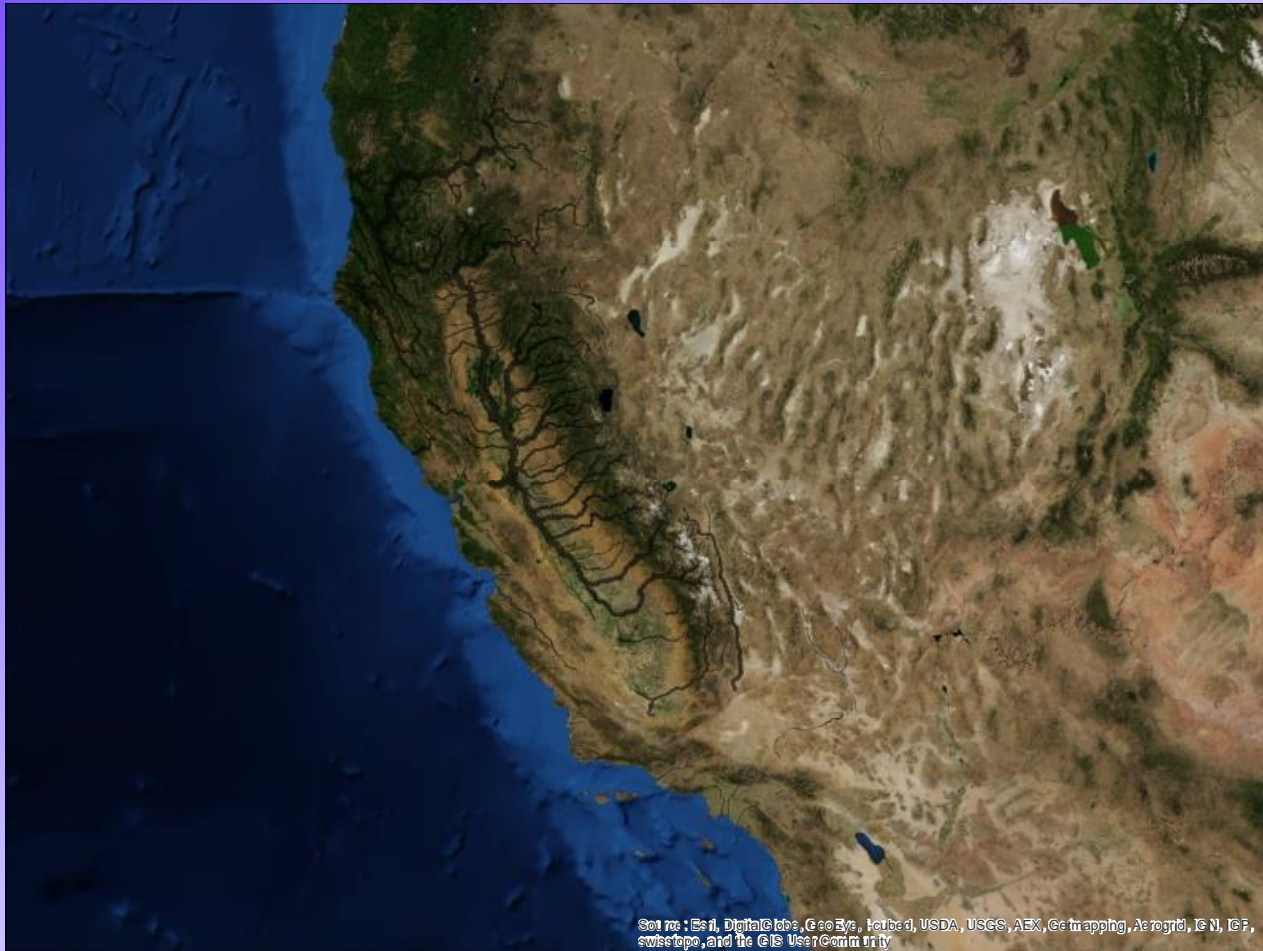


No gage
available
here

Where I live

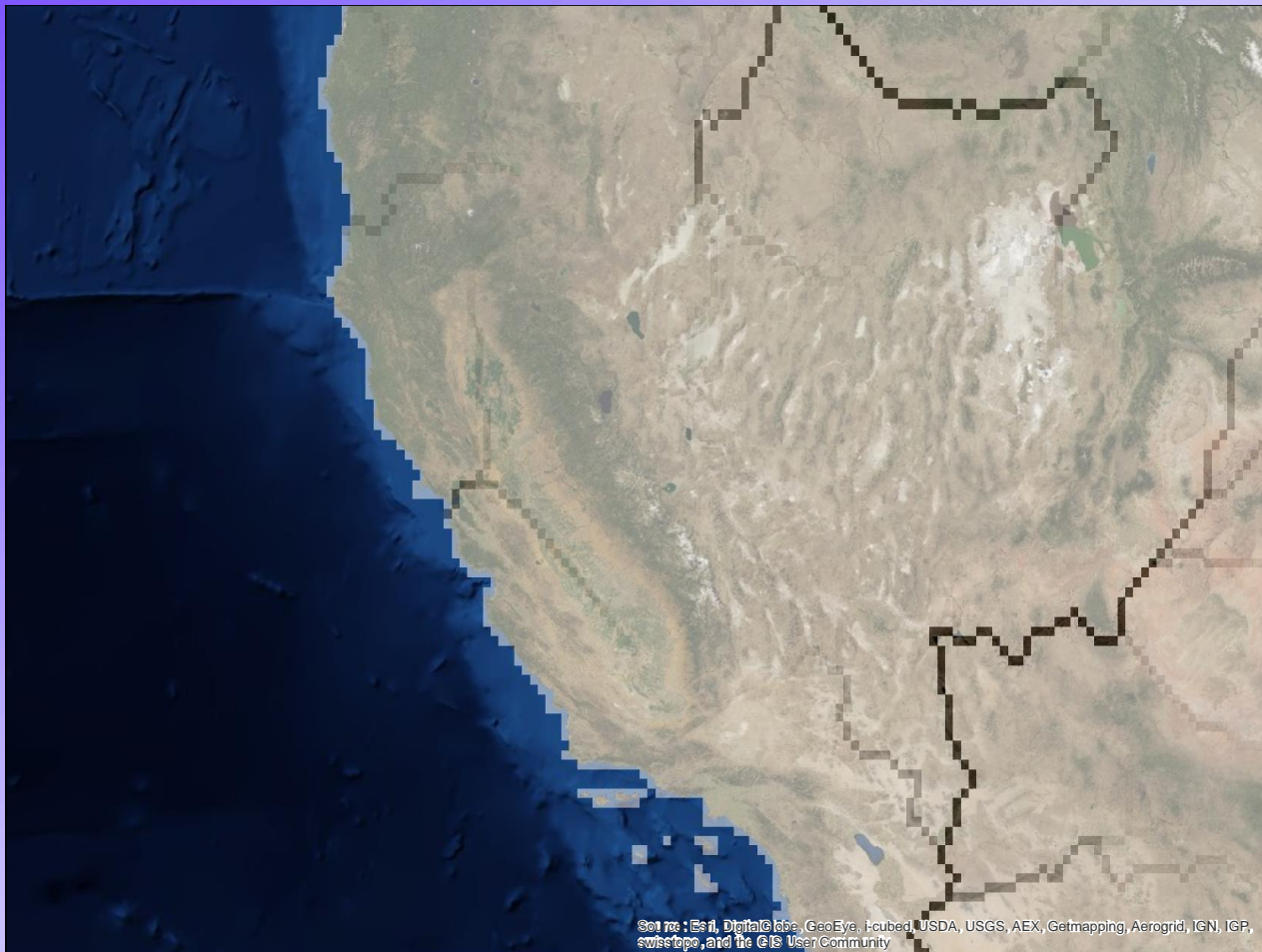
What if **anyone, anywhere** could have access to
the flow in the **river next to their home?**

Need access to continental-scale model results



Unified management for a nation
General understanding by its citizens

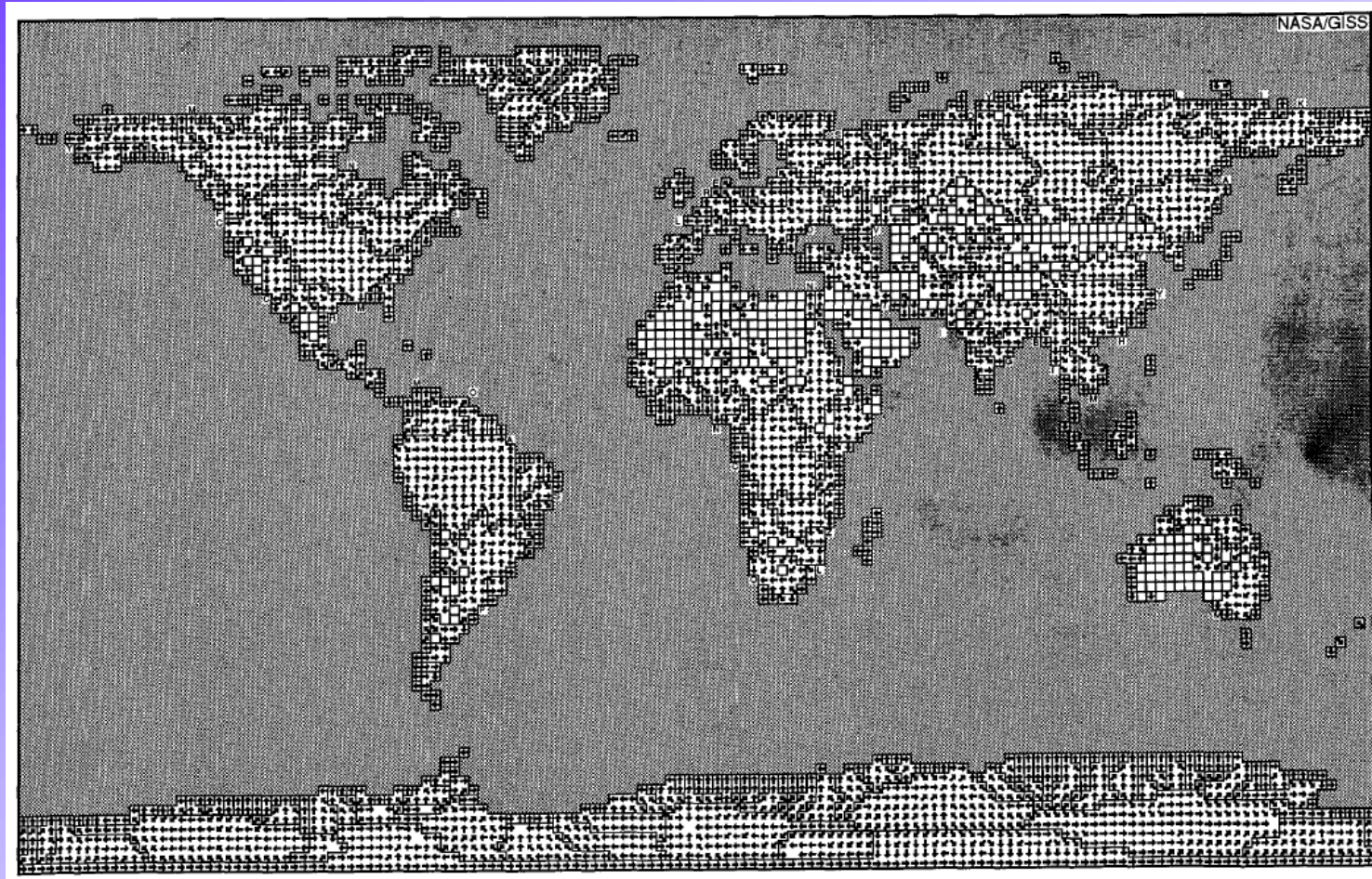
State-of-the-art continental-scale river modeling



Why so coarse?

River network of Lohmann et al (1996), Mitchell et al. (2004), and Xia et al. (2012)

Scientific legacy - We've come a long way!

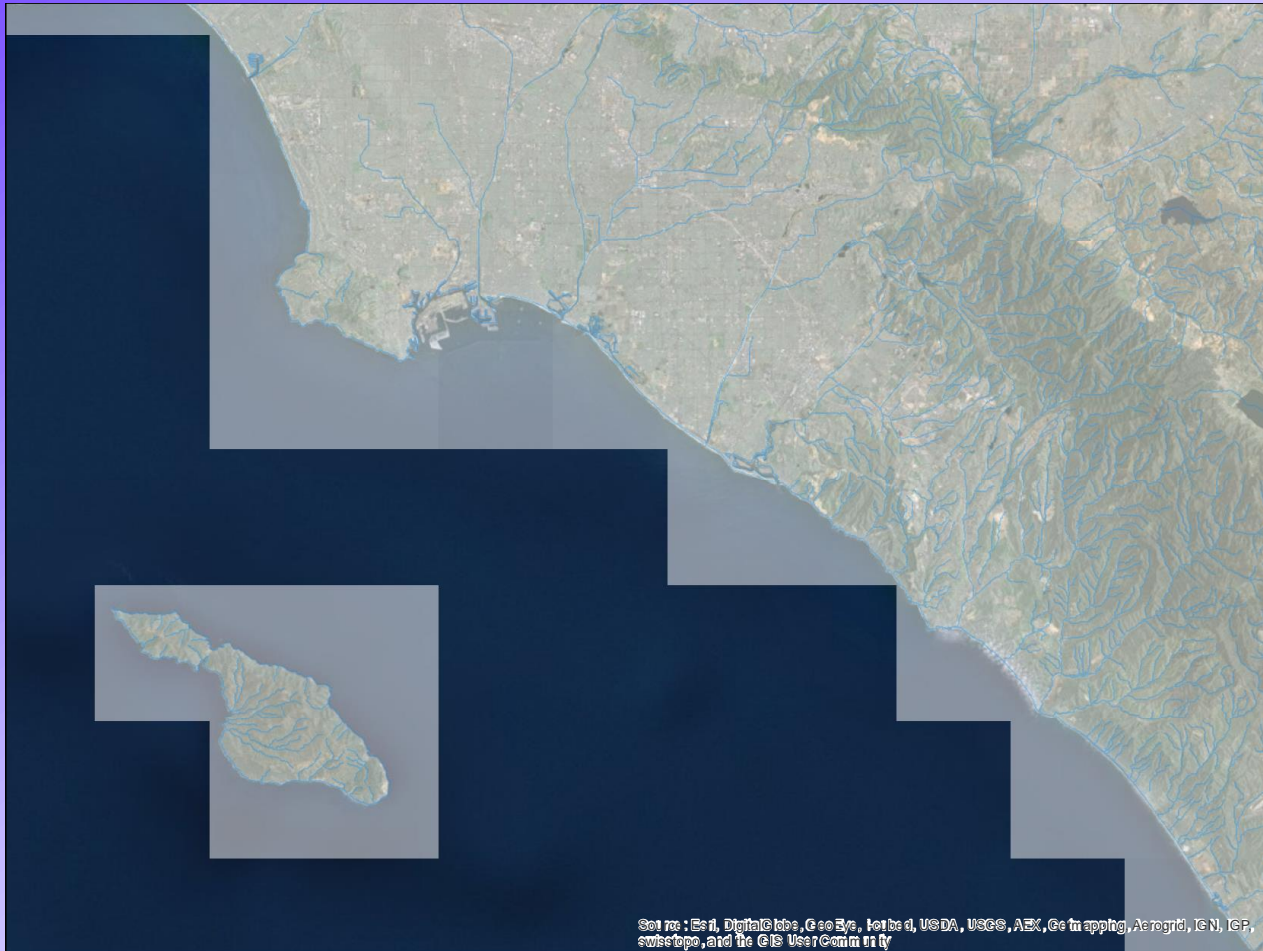


Miller et al. (1994) J. Clim.

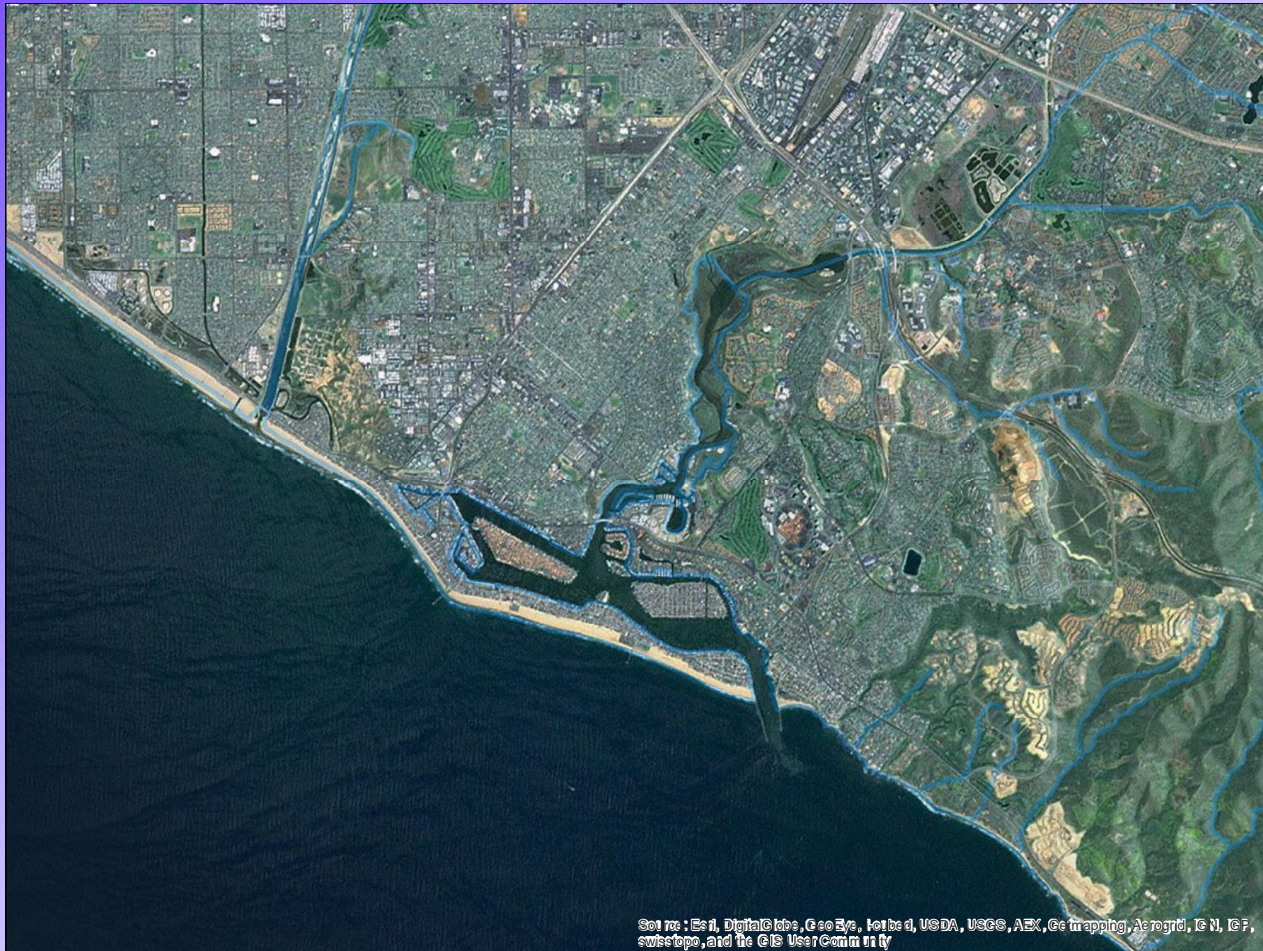
Mapped rivers of Orange County



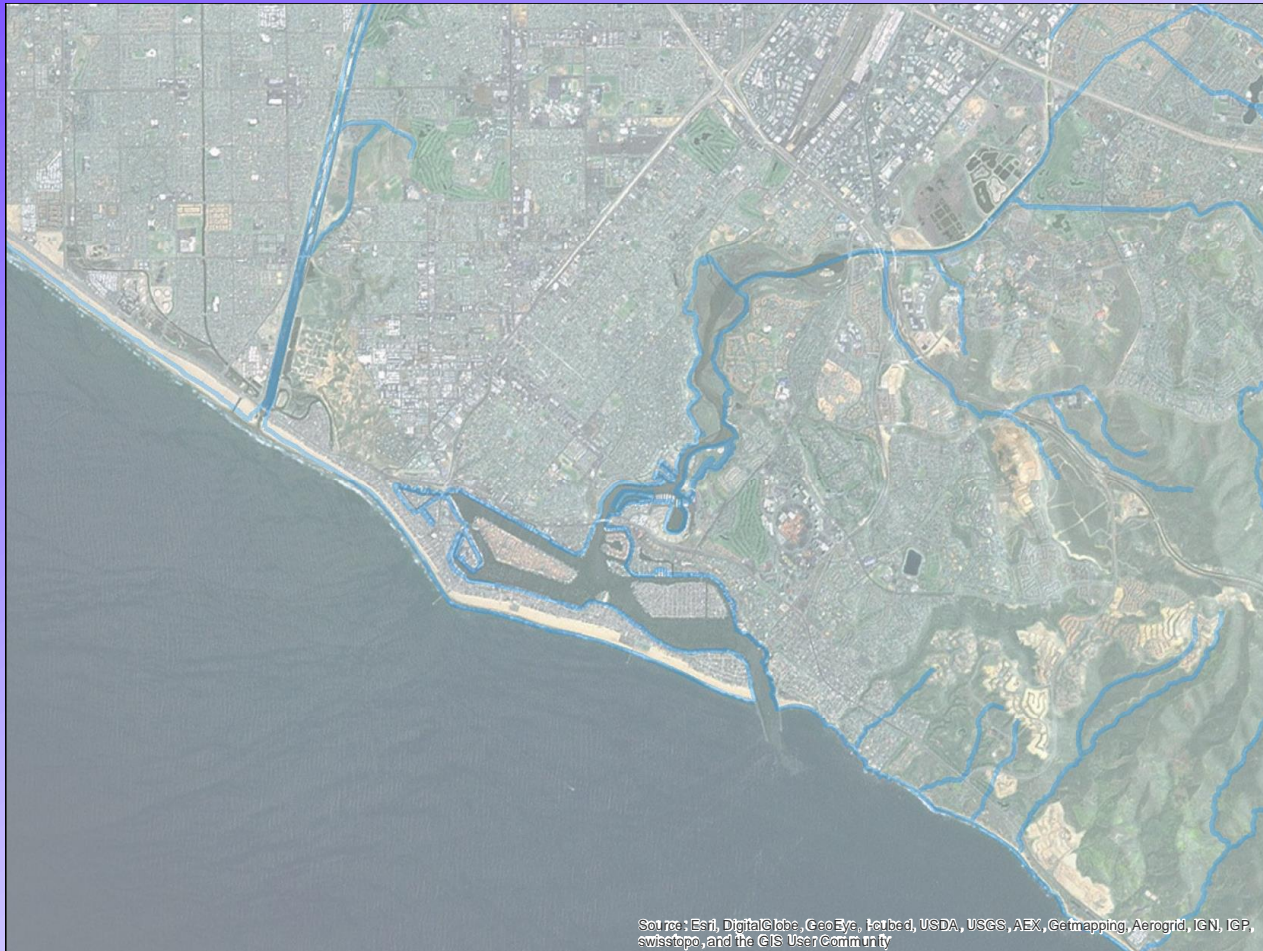
State-of-the-art river continental-scale modeling



Mapped rivers of Newport Beach

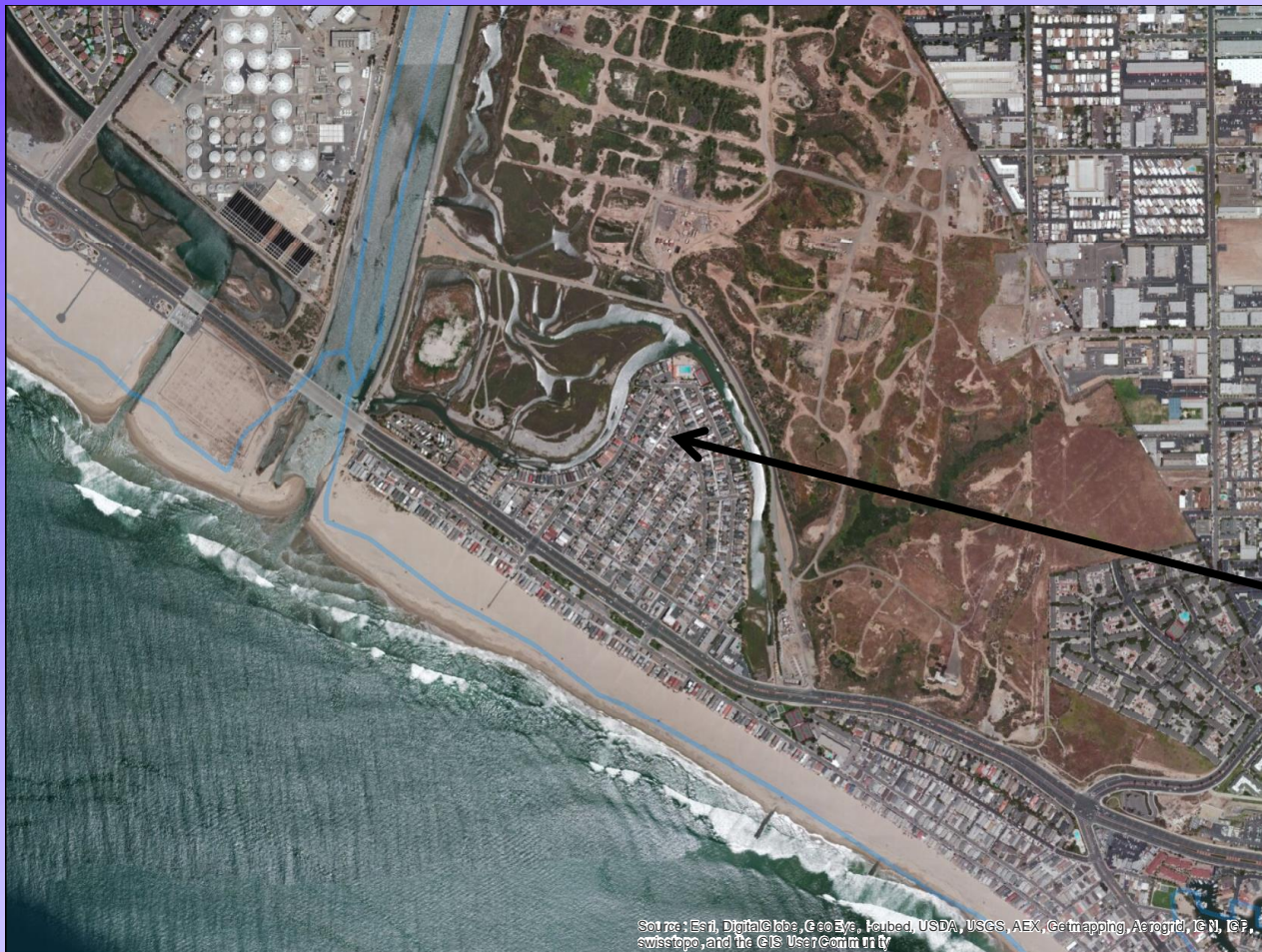


State-of-the-art river continental-scale modeling



Source: Esri, DigitalGlobe, GeoEye, Earthstar, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

Santa Ana River



Where I live

Source: Esri, DigitalGlobe, GeoEye, Earthstar, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

State-of-the-art river continental-scale modeling

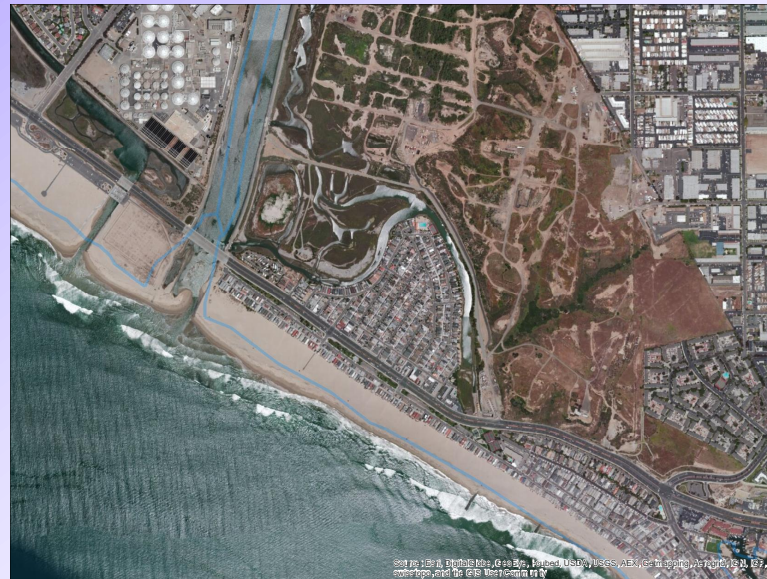
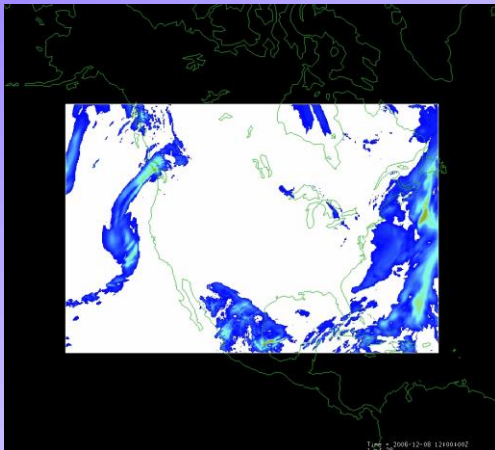


Where I live

Source: Esri, DigitalGlobe, GeoEye, i-cubed, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

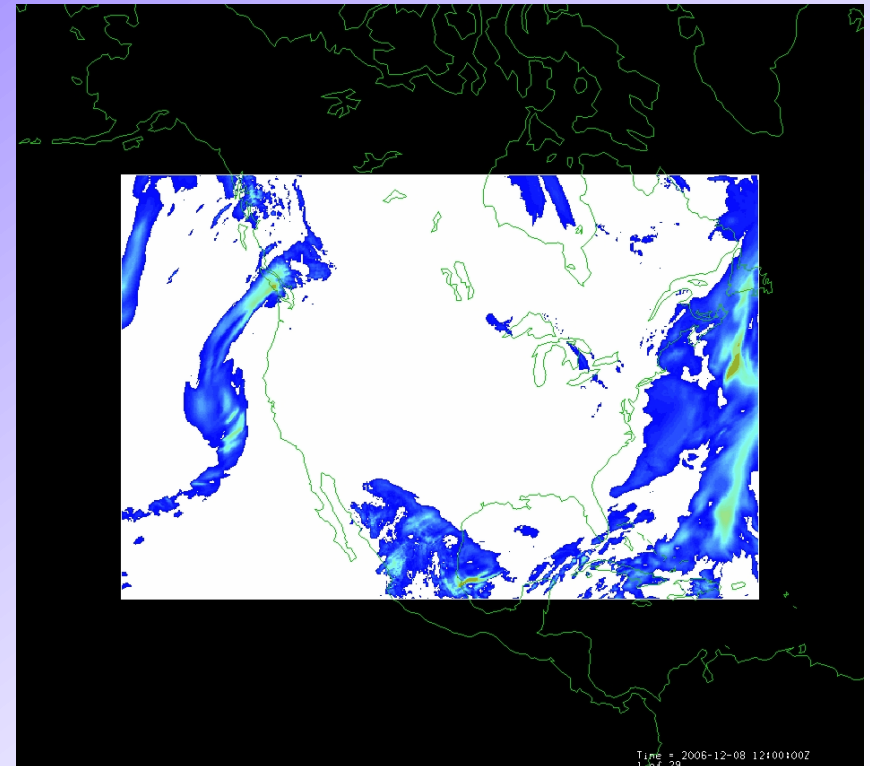


What technology can we use to enable the flow of digital rivers at the continental scale?



Atmospheric modeling

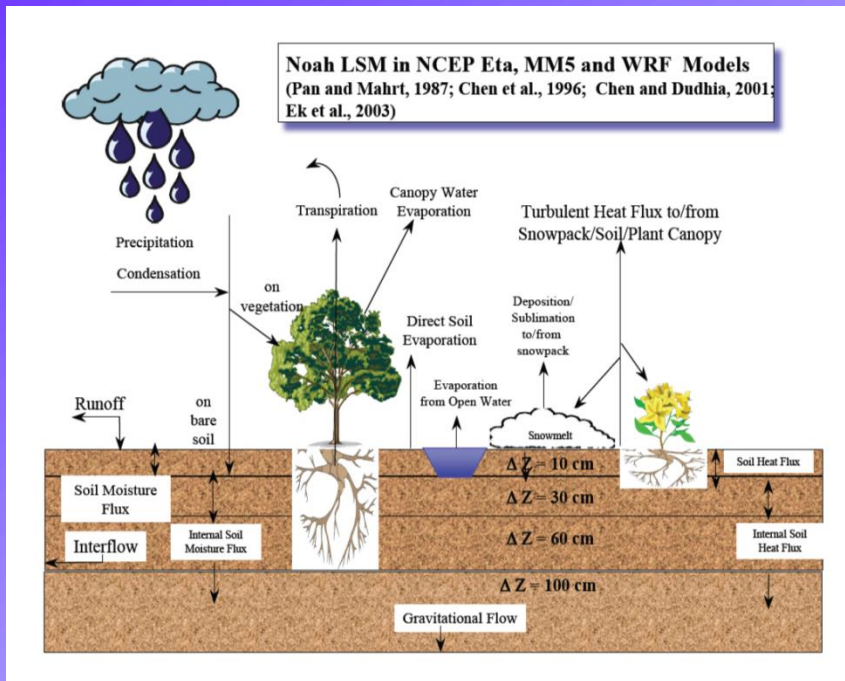
- Equations of fluid mechanics and thermodynamics of the atmosphere commonly solved everyday by computer models
- Temperature, pressure, winds, precipitation, etc. are available
- For past, present and future
- Dynamic maps



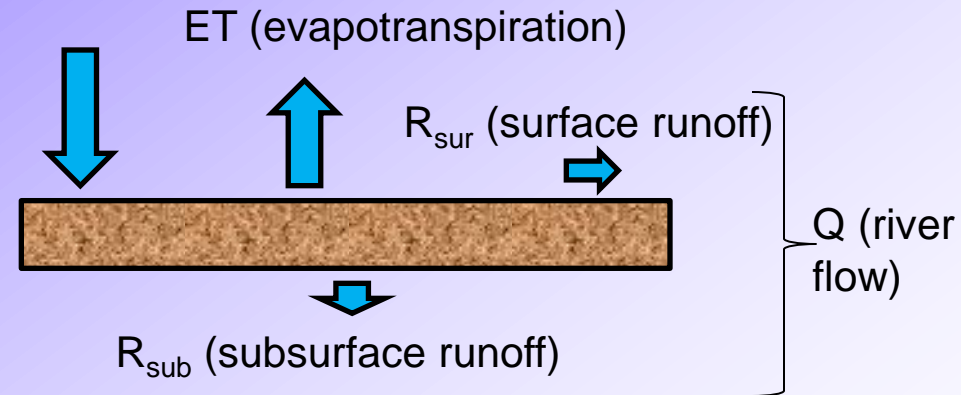
Animation of precipitation over the U.S.

Models and datasets available cover the entire U.S.

Land surface modeling



P (precipitation)



Noah land surface model (first version in 1999), serves as the land model for operational weather prediction in North America

**Bottom boundary conditions for atmospheric models
Models and datasets also cover the entire U.S.**

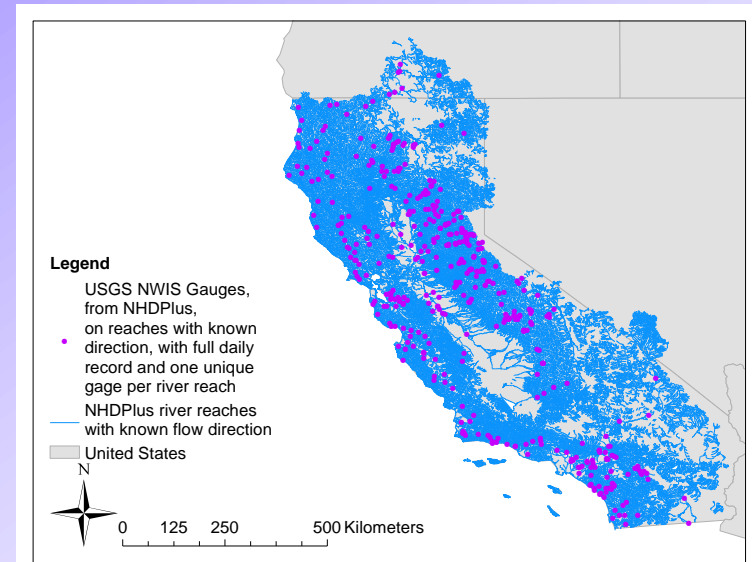
NHDPlus – River and Catchment Network for the Nation

Entire dataset



3 million river reaches

Region 18 = California



137,401 river reaches
420,000 km²

Connectivity
information

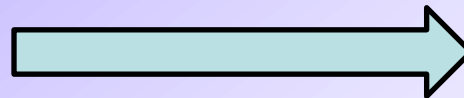
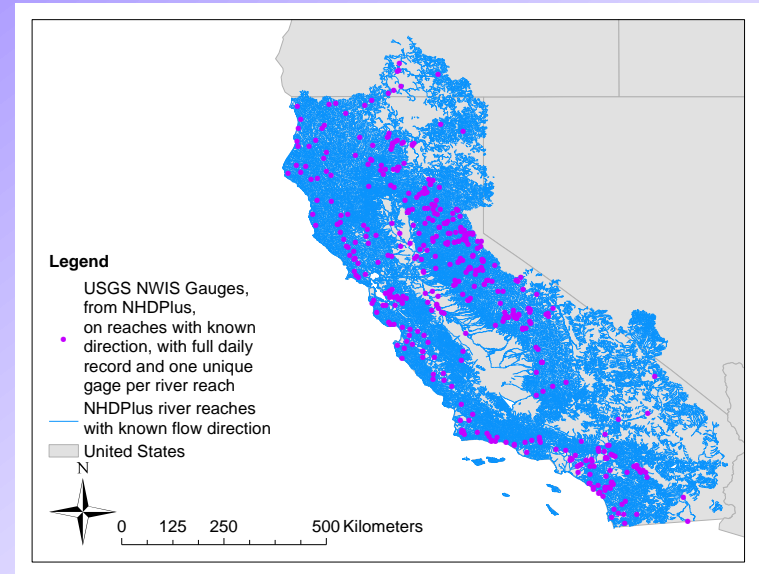
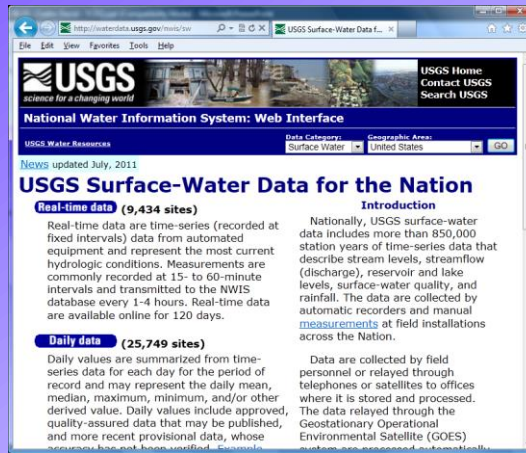
Integration of the National Hydrography Dataset, National Elevation Dataset and National Land Cover Dataset completed by EPA in 2006

“Blue line” rivers available for the entire U.S.

USGS National Water Information System

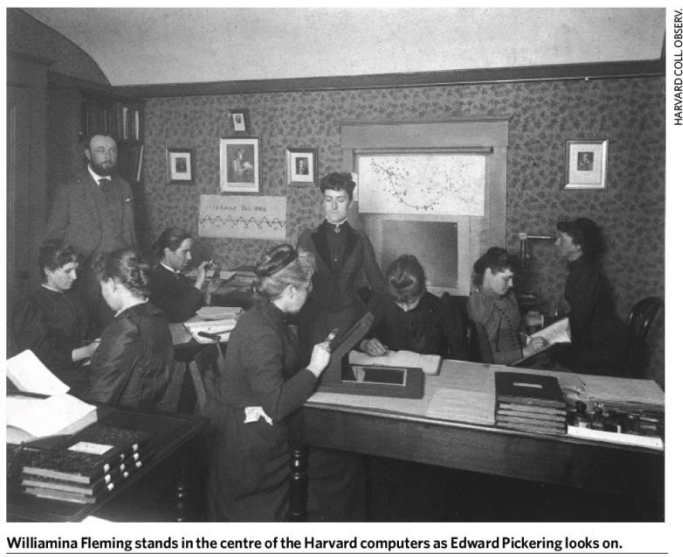


20,000+ gauges available for the United States



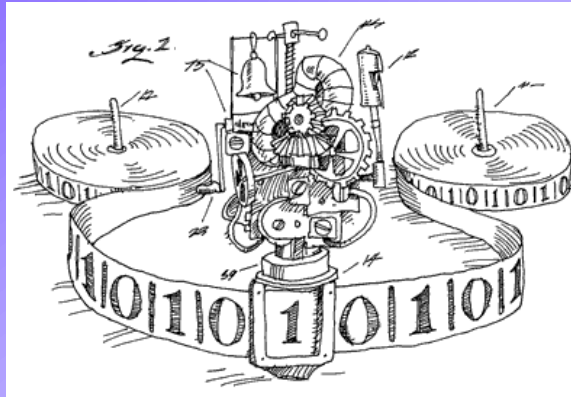
2311 gages total in Region 18. 433 with full daily data for 2000/01/01 – 2007/12/31 and on reaches with known direction

Computers and computing tools have evolved



Williamina Fleming stands in the centre of the Harvard computers as Edward Pickering looks on.

People in 1880s (Nature, 2008)



Turing machine (1936)



IBM Portable PC 5155 (1984)



Sun Supercomputer (Ranger) World's fastest computer for open science when released (2008)



Cray 2 Supercomputer World's fastest computer when released (1985)



Computing Centers



Scientific computing libraries

River modeling and Computing

What's being used:
Desktop computers



What could be used:
Parallel computers



Lonestar
5,840 processors



Ranger
62,976 processors

- Today's computers are as powerful as supercomputers ten years ago
- Most computers come with multiple cores
- Parallel computing is becoming increasingly accessible
- Need to know parallel computing
 - A learning curve

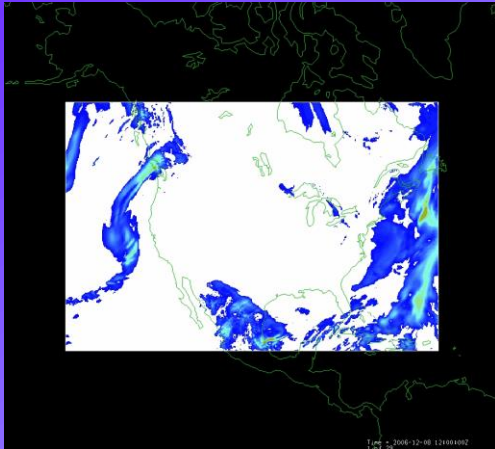
Existing Large-scale River Routing Models

- *Miller et al. (1994)*
- *Lohmann et al. (1996)*
- *Olivera et al. (2000)*
- *Oki et al. (2001)*
- *Goteti et al. (2008)*

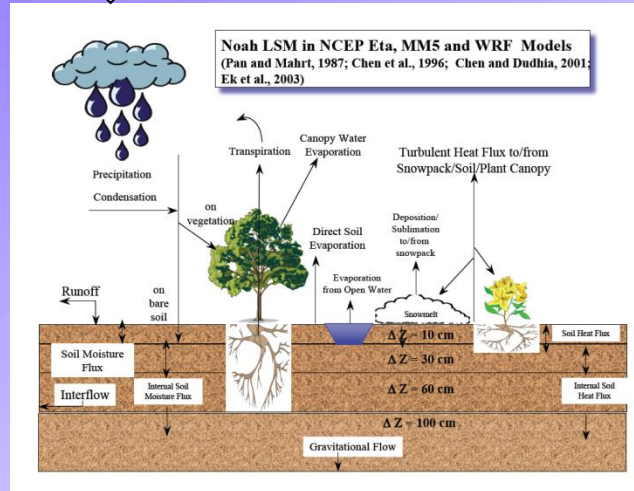
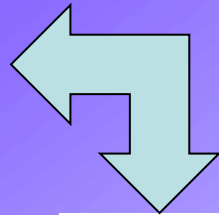
Almost no vector-based (GIS) hydrography, Almost no parallel computing



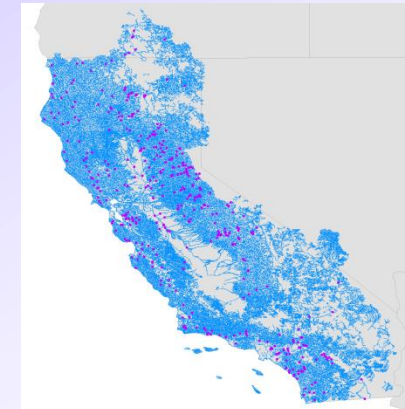
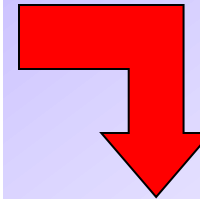
Integrated River Modeling



Atmospheric Model
or Dataset



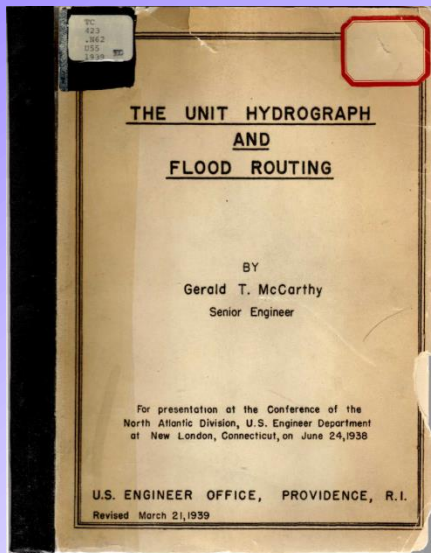
Land Surface
Model



“Blue Line” River Network -
High-Performance Computing
River Network Model



From available technology to applications

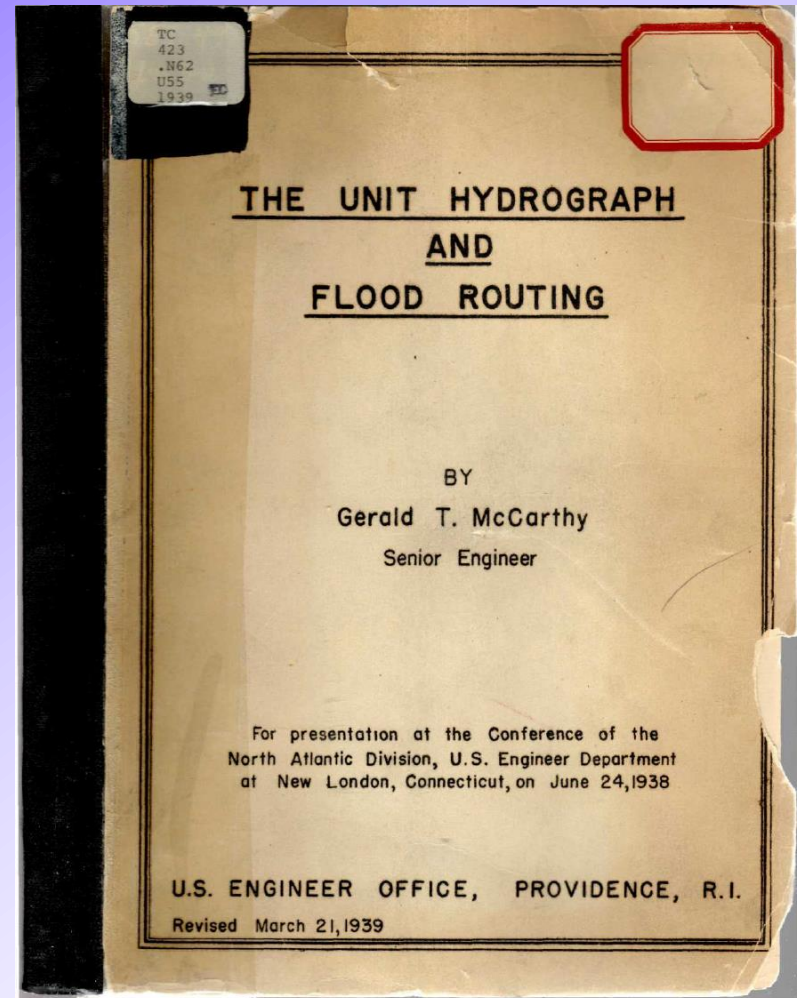


The Muskingum method

$$Q_j(t + \Delta t) = C_{1_j} \cdot Q_j^{up}(t + \Delta t) \\ + C_{2_j} \cdot Q_j^{up}(t) \\ + C_{3_j} \cdot Q_j(t)$$



IBM Portable PC 5155 (1984)
Price : \$4,225 (256 - 640K RAM)

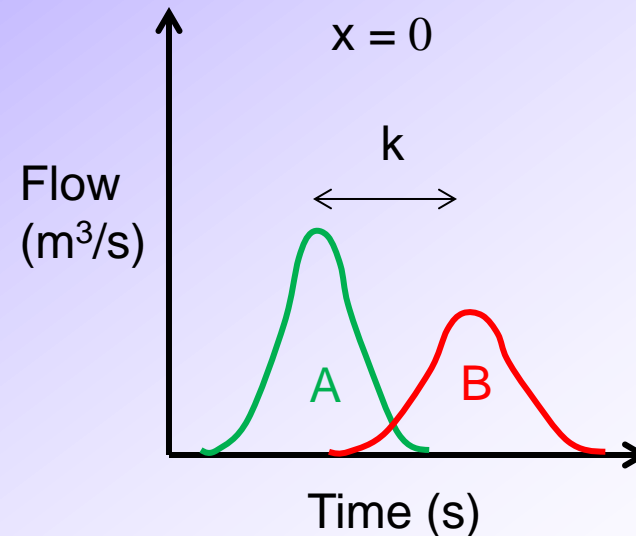
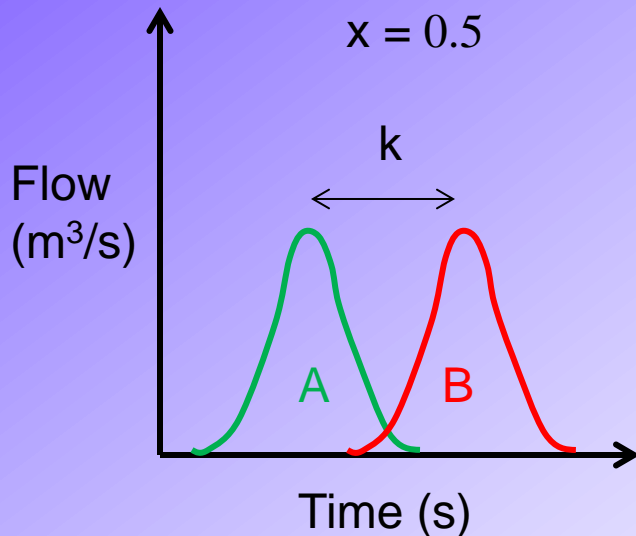
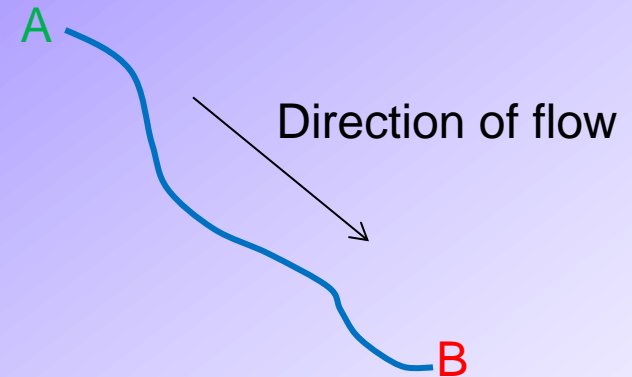


McCarthy (1938)

Meaning of the Muskingum parameters

k is a time ($k \geq 0$) related to the celerity of the flow wave

x is a non-dimensional parameter ($0 \leq x \leq 0.5$) related to diffusion of the flow wave



A vector-matrix version of the Muskingum method

$$(\mathbf{I} - \mathbf{C}_1 \cdot \mathbf{N}) \cdot \mathbf{Q}(t + \Delta t) = \mathbf{C}_1 \cdot \mathbf{Q}^e(t) + \mathbf{C}_2 \cdot [\mathbf{N} \cdot \mathbf{Q}(t) + \mathbf{Q}^e(t)] + \mathbf{C}_3 \cdot \mathbf{Q}(t)$$

I Identity matrix

N Network matrix, computed based on network connectivity information

Q^e Vector of flow rates from outside the network into upstream of each river reach

C₁, C₂, C₃ Parameter matrices, computed based on the values of k and x

Q Vector of flow rates at the outlet of each river reach (output of RAPID)



(First formulated in 2007)

River model

- **RAPID** (Routing Application for Parallel computation of Discharge)
- **Computes flow and optimizes model parameters**
- Model code, input data and animations are available online
- Can run on supercomputers



Home
Résumé
Publications
RAPID
RAPID Training Course
Contact

RAPID

Description
RAPID (Routing Application for Parallel computation of Discharge) is a river routing model. Given surface and groundwater inflow to rivers, the model can compute flow and volume of water anywhere in river networks made out of many thousands of reaches. The design of RAPID allows it to be adapted to any river network, if given basic connectivity information. RAPID uses a matrix version of the Muskingum method, and has an automated parameter estimation procedure that allows finding optimal model parameters based on available gage measurements. This model uses the Fortran programming language and can be run on personal computers, as well as on massively-parallel supercomputers, with actual parallel speedup. RAPID has the ability to run and/or optimize model parameters on any subbasin included in its computing domain. If major man-made infrastructure are present on the river network, RAPID allows to locally substitute upstream flow by gage measurements during both computation of river flow and optimization of parameters. Detailed information on RAPID can be found in the related [publications](#).

Development history
Designing, developing and testing RAPID were a large part of my Ph. D. work at the Center for Research in Water Resources at the University of Texas at Austin. The development of this model started in September 2007 as I joined the Center for Geosciences at Ecole des Mines de Paris (Mines ParisTech), France for a 6-month visit. RAPID was originally developed as a substitute for the river routing scheme of SIM-France, the operational hydro-meteorological model used by Météo France (the French weather service). The code has since then been adapted to run on the NCEP/NCAR dataset that provides a "blue line" description of the river networks in the USA. For the current NCEP/NCAR version of RAPID, several land surface models can be used to compute inflow of water to the river network.

Cite
Download the [RAPID source code](#).
Download [input data](#) from David et al. (2011, IJDR). These include NCEP/NCAR river network connectivity, lateral inflow from the land surface (computed with NCEP/NCAR and gage measurements from USGS NWIS) for a 4-year run (between 2004-01 and 2007-12) in the San Antonio and Guadalupe River Basins in Texas.
Download [output data](#) from David et al. (2011, IJDR).
Download [input data](#) from David et al. (2011, IJDR). These include SIM-France river network connectivity, inflow from the land surface (computed with IS9A), lateral inflow from groundwater (computed with MODCOOL) for a 10-year run (between 1995-08-01 and 2005-07-31) in the domain of SIM-France.
Download a [script to get NCEP/NCAR land model data](#) and convert them to RAPID water inputs and [download other corresponding RAPID input data](#).
Download an [iMovie script to flow maps](#) from a [shapefile](#) and corresponding RAPID outputs and [save as an image](#). This images created with this script are used to make the animations below.

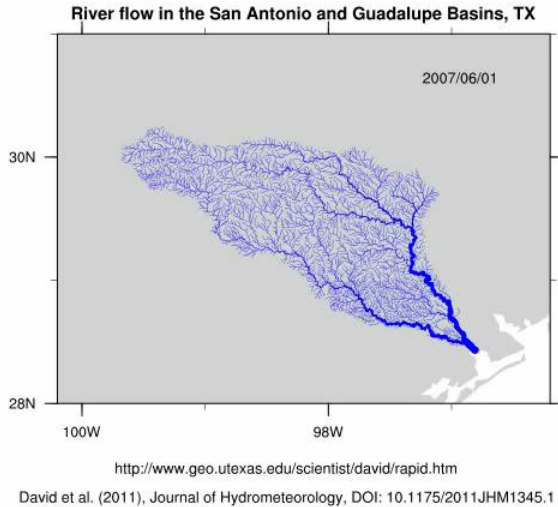
Documents
A few slides explaining [how RAPID works](#).
A few slides on [parallel computing in RAPID](#).
A [tutorial guide on how to compile and run RAPID](#), and a [detailed guide on how to compile and run RAPID](#).
Some information on the [data model used in RAPID](#).
Explanations on the [input and output files used in RAPID](#).
A guide on [how to download NCEP/NCAR shapefiles and format them for use in RAPID](#).
A tutorial on [how to display a map and a hydrograph of RAPID flow rate computations in ArcGIS](#).

Animations
In the following animations of stream flow, the thickness of river reaches varies with the magnitude of the flow rate going through them. One can see the flow waves propagating downstream. All these animations were prepared using 3-hourly outputs from RAPID. They can be played directly from this webpage (full-screen mode available) or downloaded in .avi, .flv and .mp4 formats. The latter can be uploaded to and played on YouTube.

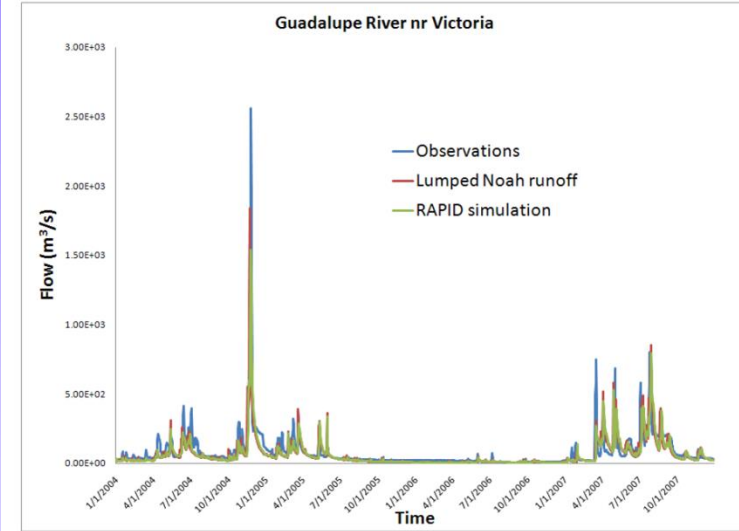
This first animation is of the San Antonio and Guadalupe Basins in Texas, USA, over four months (between 2007-06-01 and 2007-09-30). 3-hourly surface and subsurface runoff was produced by the VisVAP land surface model using a combination of NCEP/NCAR and NCEP/NCAR for atmospheric forcing. RAPID was run at a 15-minute time step. Download [avi](#), [flv](#) or [mp4](#). Or [play directly from YouTube](#).

<http://www.ucchm.org/david/rapid.htm>

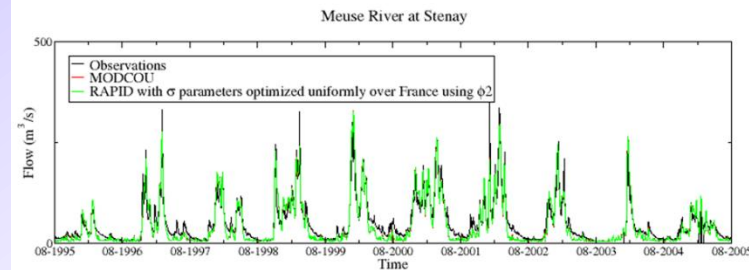
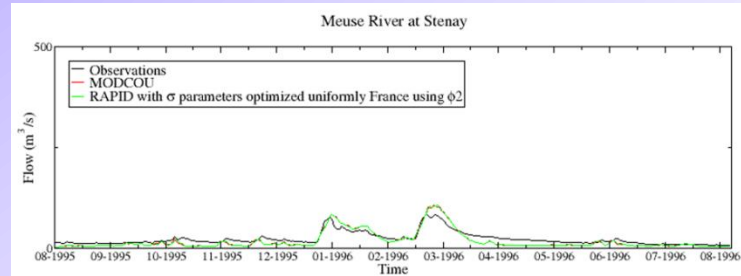
Existing Applications (1/3)



David et al.
2011a
J. Hydromet.

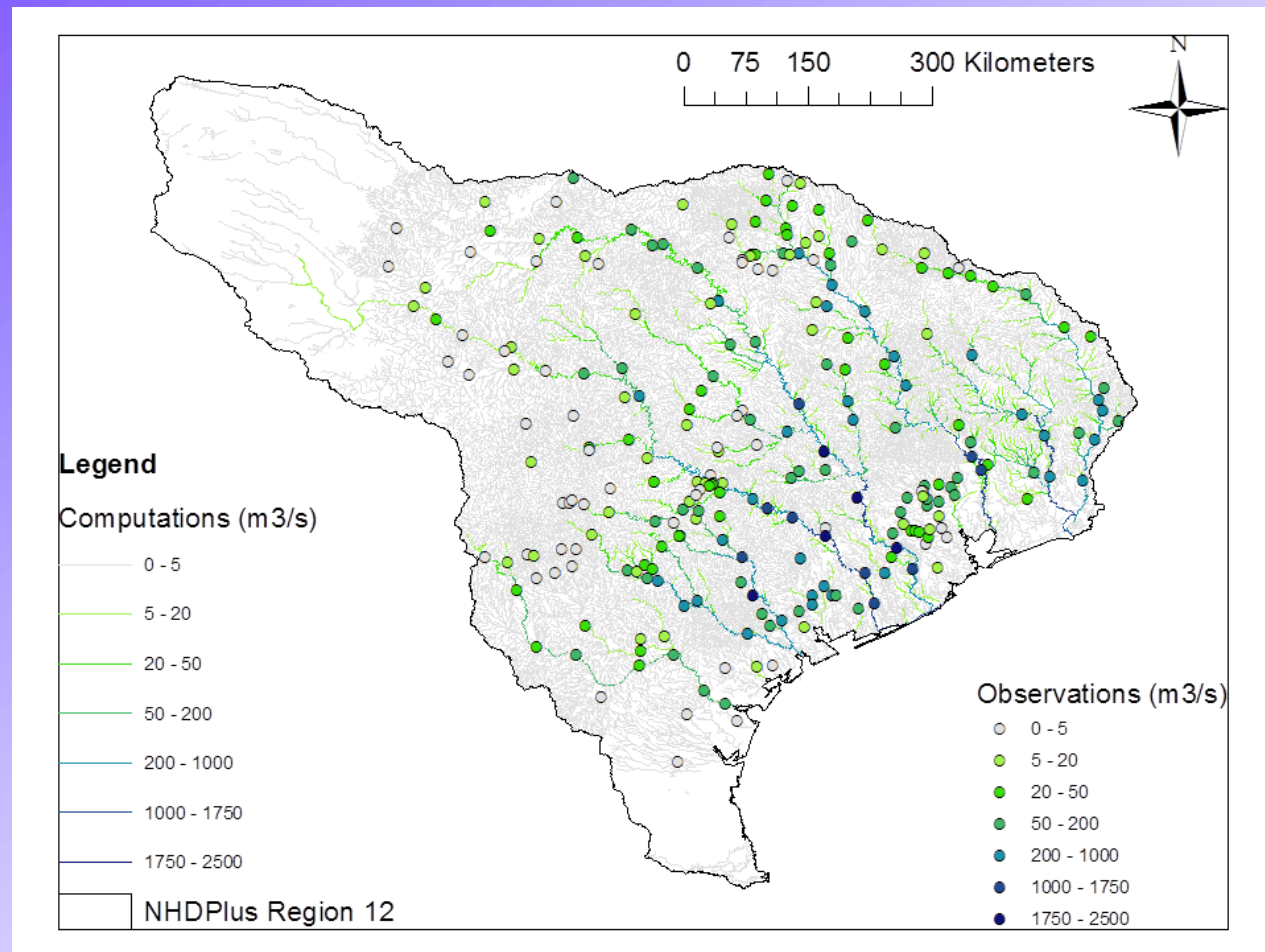


David et al.
2011b
Hydrol. Proc.



Existing Applications (2/3)

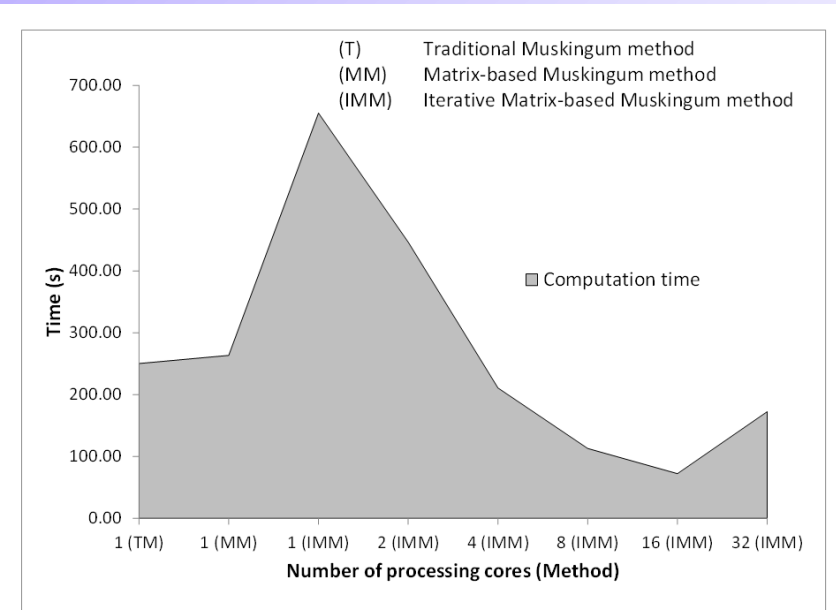
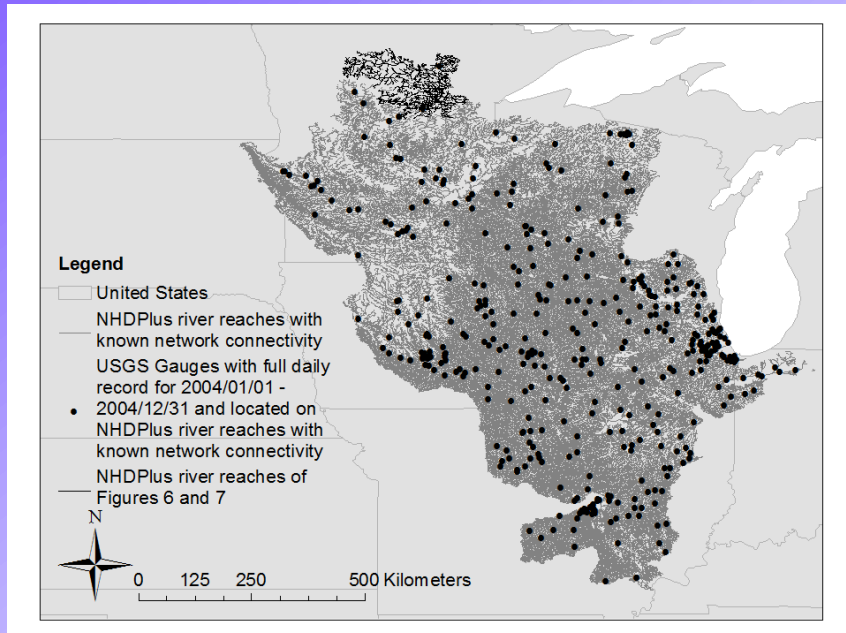
David et al.
2013a
Env. Mod. & Soft.



Existing Applications (3/3)

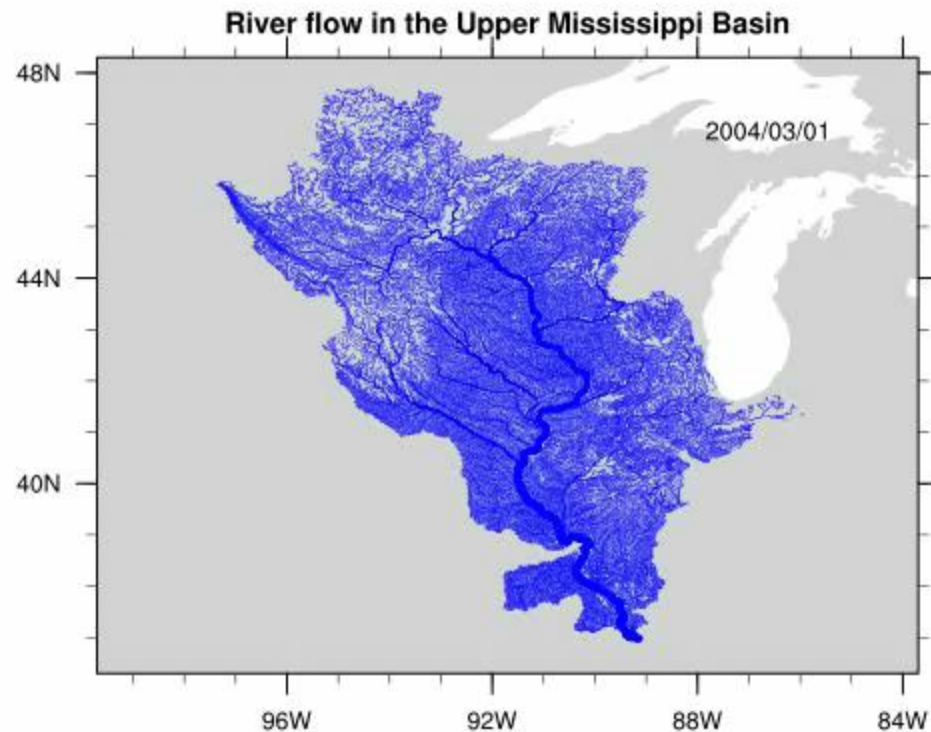


David et al.
2013b
Water. Resour. Res.



So fast theory had to be revised!!!

Animation of Upper Mississippi Basin

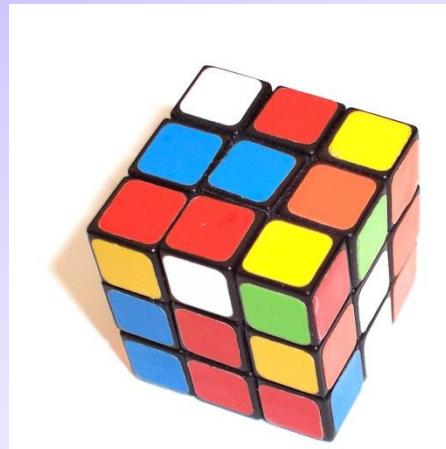


<http://www.geo.utexas.edu/scientist/david/rapid.htm>

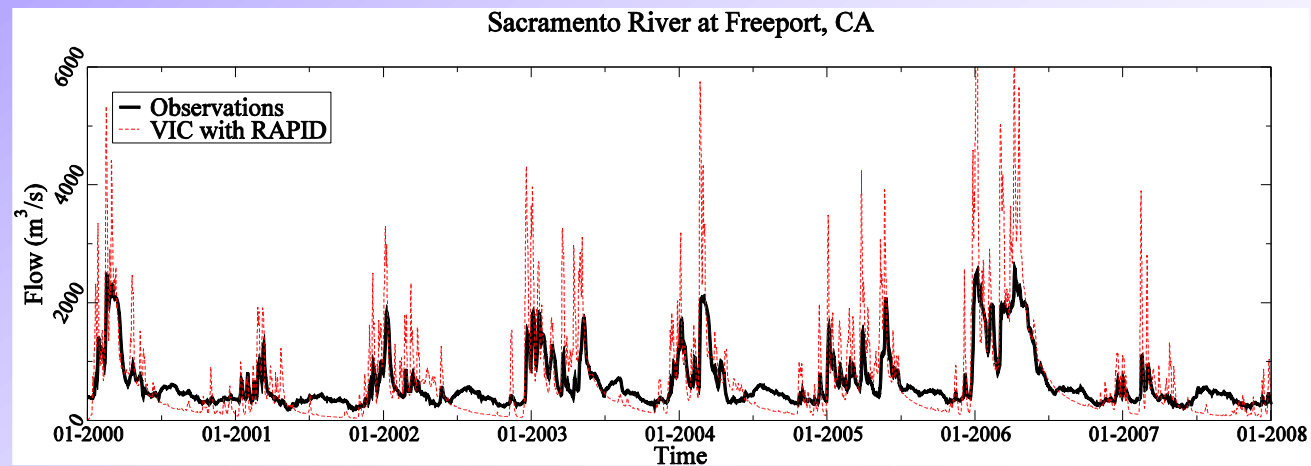
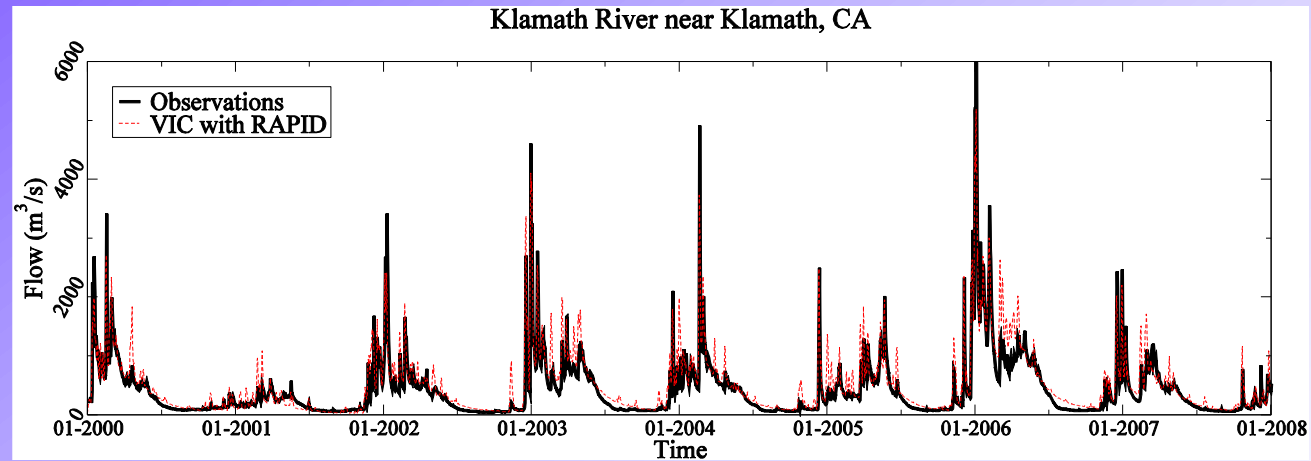
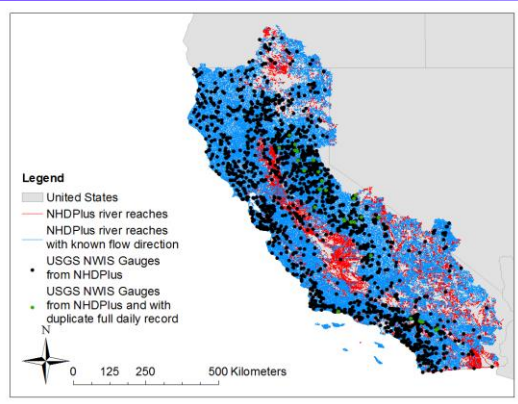
David and Yang (201x), in preparation



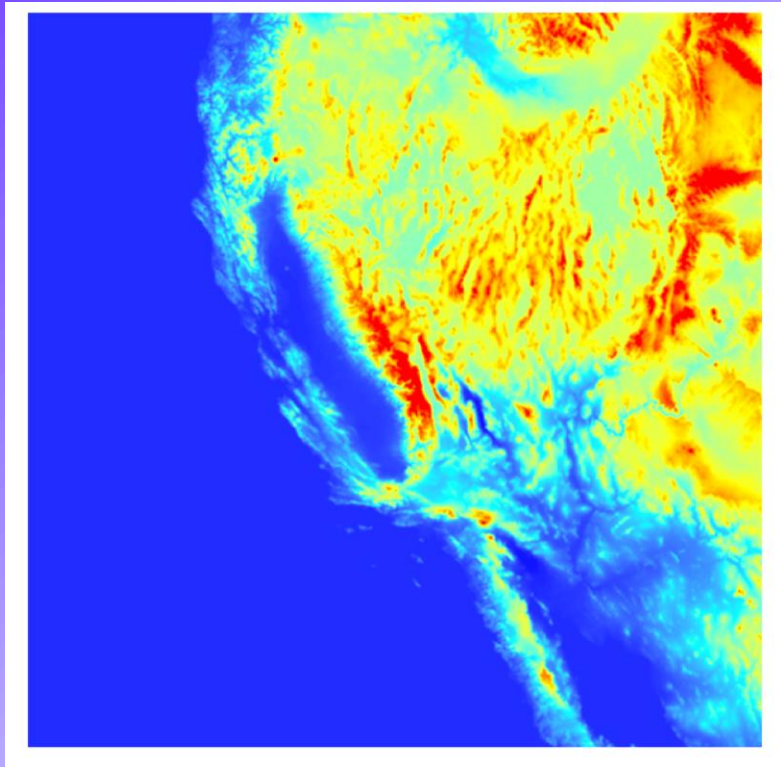
Challenges remain!!!



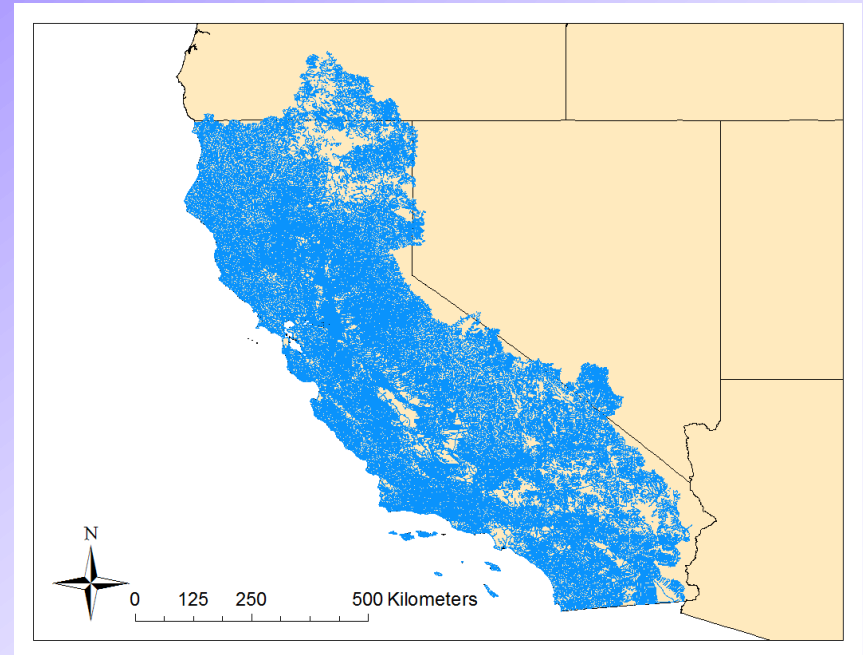
Human influence on surface water



New domain = New data processing

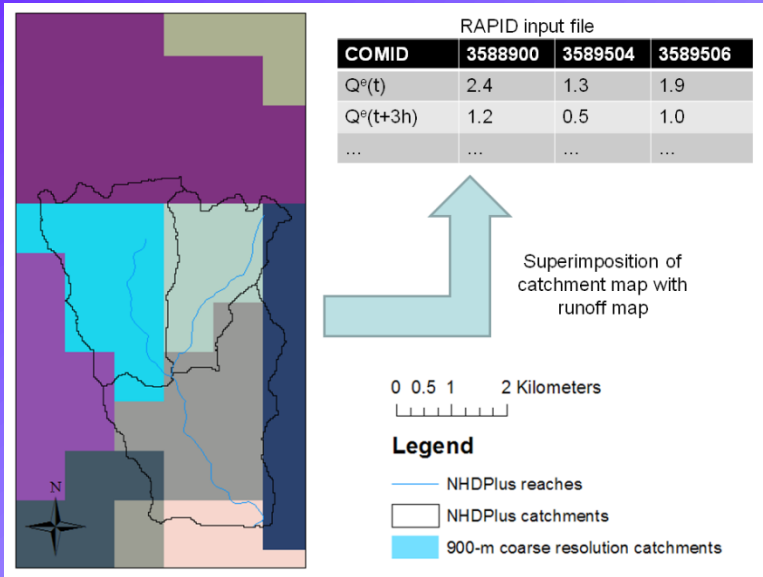


Existing WRF-Hydro domain
over California

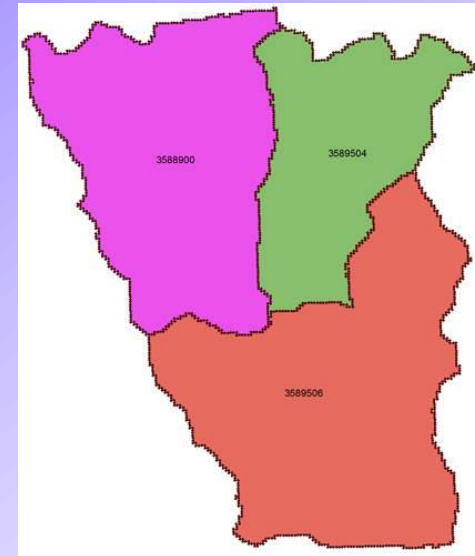


River networks of California

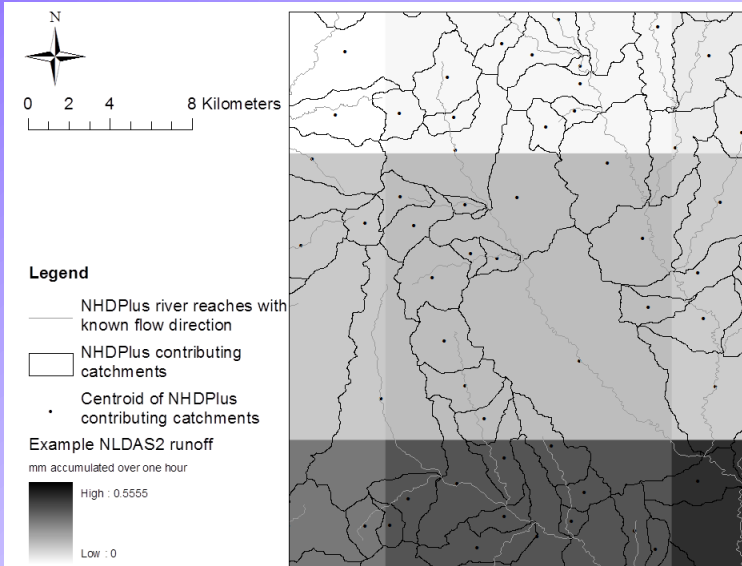
Land/river connection



David et al. (2011a)



Parallel re-gridders



David et al. (2013a)



Workflow opportunities

The screenshot shows a web browser window displaying the Earth Cube Workflow Community website. The page title is "Earth Cube Workflow Community" and the main heading is "Workflow Vignettes". The left sidebar contains a navigation menu with categories like "Home", "About Workflows", "About the EarthCube Workflow Community Group", "Questionnaire for the Community", "EarthCube Workflows Roadmap", "Weekly Meetings", and "Workshops". The "Workflow Vignettes" link is highlighted in orange. The main content area features a paragraph explaining the purpose of the vignettes, a list of requirements for submitting a vignette, and a preview of a vignette titled "A Software Marketplace Paradigm Highlight: River Modeling at the Regional Scale: An Example Using RAPID".

Earth Cube Workflow Community

Search this site

Workflow Vignettes

The EarthCube Workflows group is collecting "vignettes" that illustrate how workflows are or could potentially be used in geosciences. These vignettes illustrate graphically and concisely the role of workflows in a particular setting, and they are very easy to create.

We welcome vignettes from everyone in the community. Simply send:

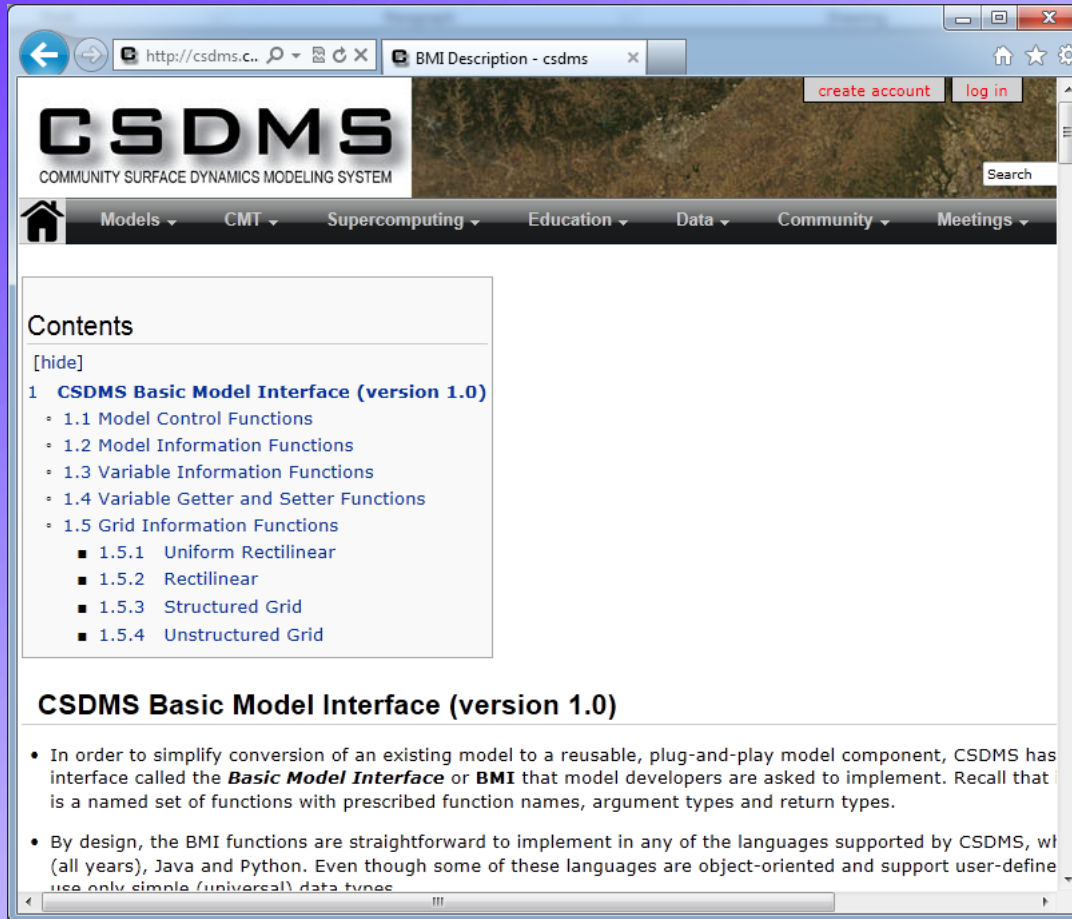
- 2-3 sentences describing the goal of using workflows environment, mentioning who are the users of the workflow and/or its results
- 2-3 sentences describing why workflows are useful (reuse, verification, provenance, etc)
- a graphic, e.g. a workflow sketch or a data product of a workflow
- a sentence mentioning the institutions involved, POC, and a URL if available
- (optional): a list of major steps involved in the workflows

Workflow Vignette Highlight

A Software Marketplace Paradigm Highlight: River Modeling at the Regional Scale: An Example Using RAPID

The Routing Application for Parallel computation of Discharge (RAPID) is a river routing model. Given surface and groundwater inflow to rivers, this model can compute flow and volume of water everywhere in river networks made out of many thousands of reaches. Running RAPID for a regional scale application involves: selecting a river network and preparing model parameters from the enhanced version of the National Hydrography Dataset (<http://www.horizon-systems.com/nhdplus/>), getting an estimate of the water inflow from surface and subsurface into the river network using data from the second phase of the National Land Data Assimilation System (NLDAS2), gathering discharge observations from the U.S. Geological Survey National Water Information System (NWIS), running the

Inclusion in existing coupling frameworks

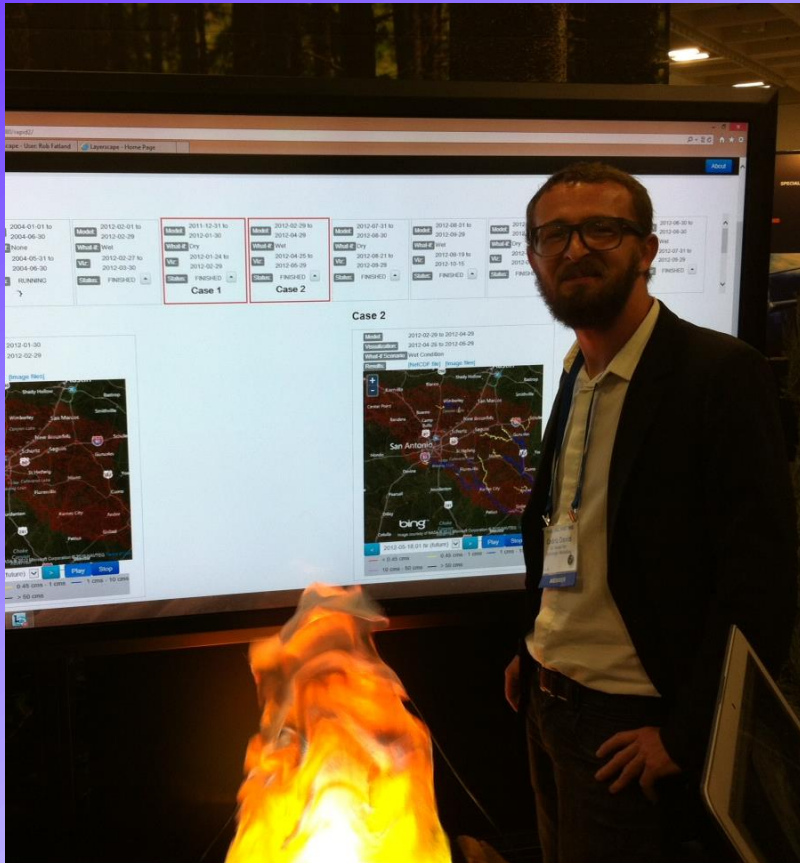


CSDMS-enabled RAPID

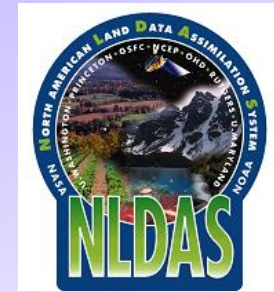
Re-think the organization of the model code

Teaches you model servitude!

Near real-time modeling???



Near real-time model input available

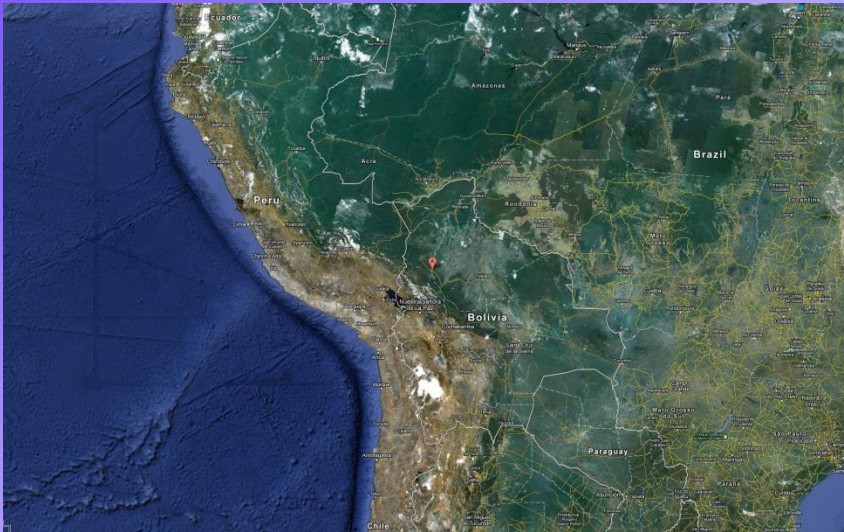


Can we automate all modeling steps?

- Pre-processing of data
- Model run
- Post-processing of data

Microsoft Booth at AGU Fall Meeting 2012

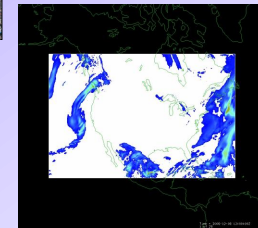
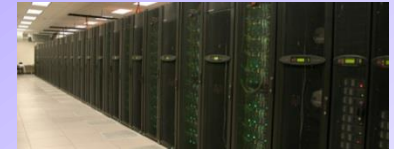
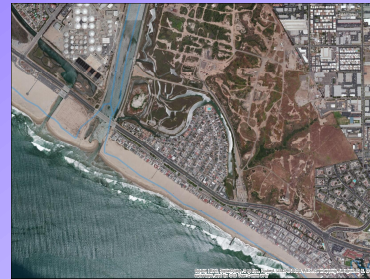
What if we did the whole world?



Rio Beni, Bolivia



Summary

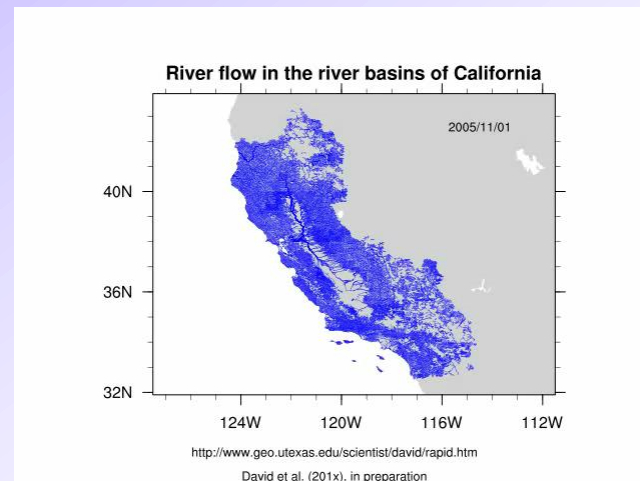


Technology

Vision

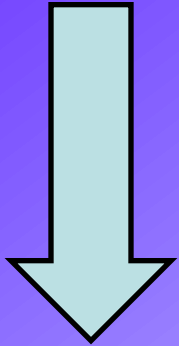


Society

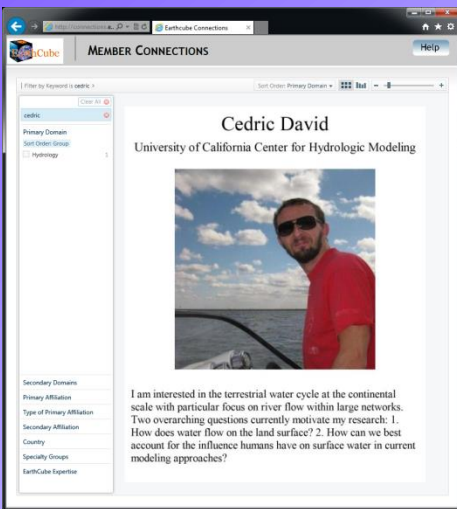


Science

Want to
collaborate???



Thank you!
Questions?



More information on RAPID at:
<http://www.ucchm.org/david/rapid.htm>
chdavid@uci.edu