BIG DATA:
Computational Fluid Dynamics (CFD): Basic Science and Engineering Applications (Cinvestav-Abacus & ININ Projects)

Cinvestav-Abacus: Centre for Applied Mathematics and High Performance Computing

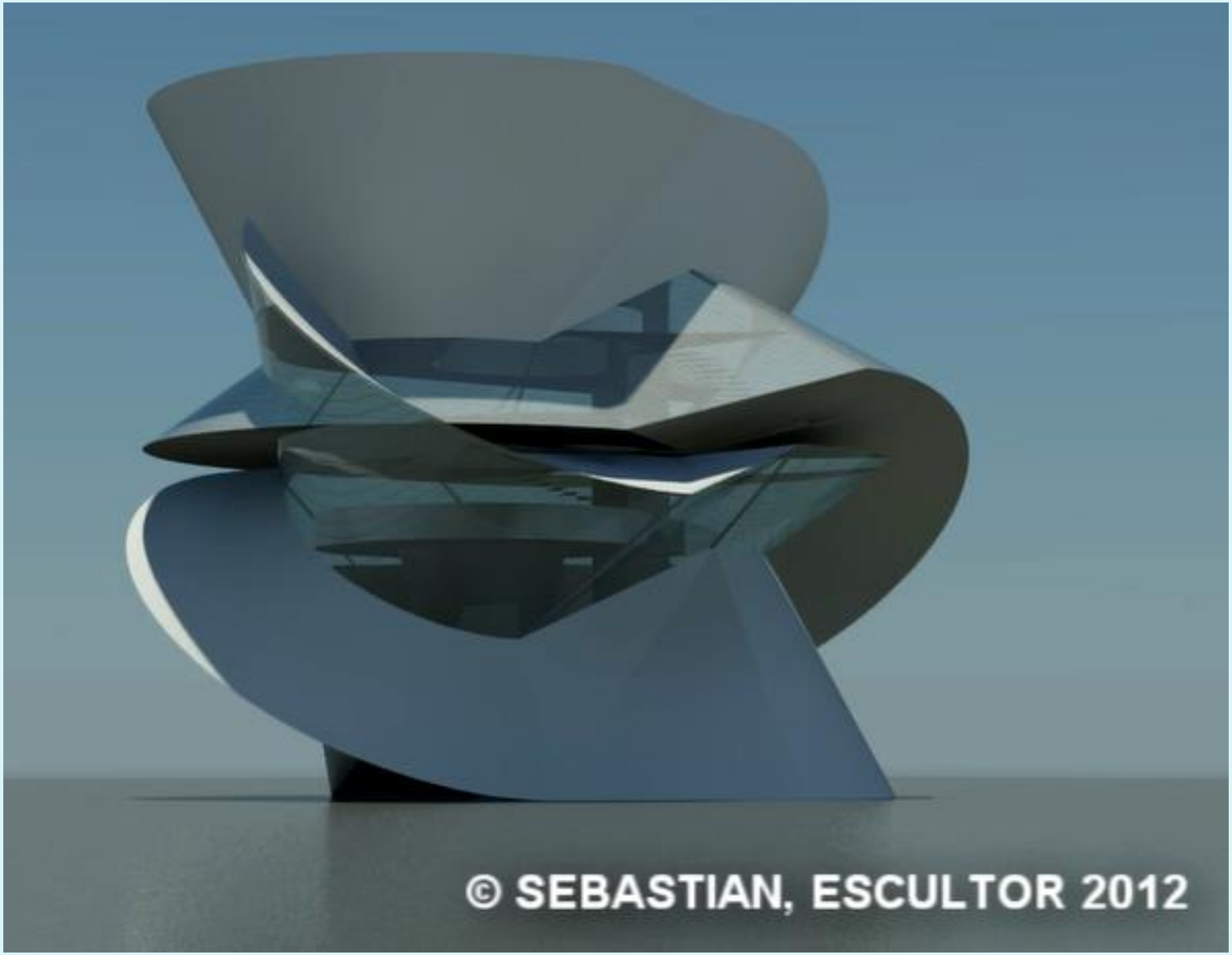
ININ: National Institute for Nuclear Research

Jaime Klapp

Cinvestav-Abacus e ININ

National Institute for Nuclear Research

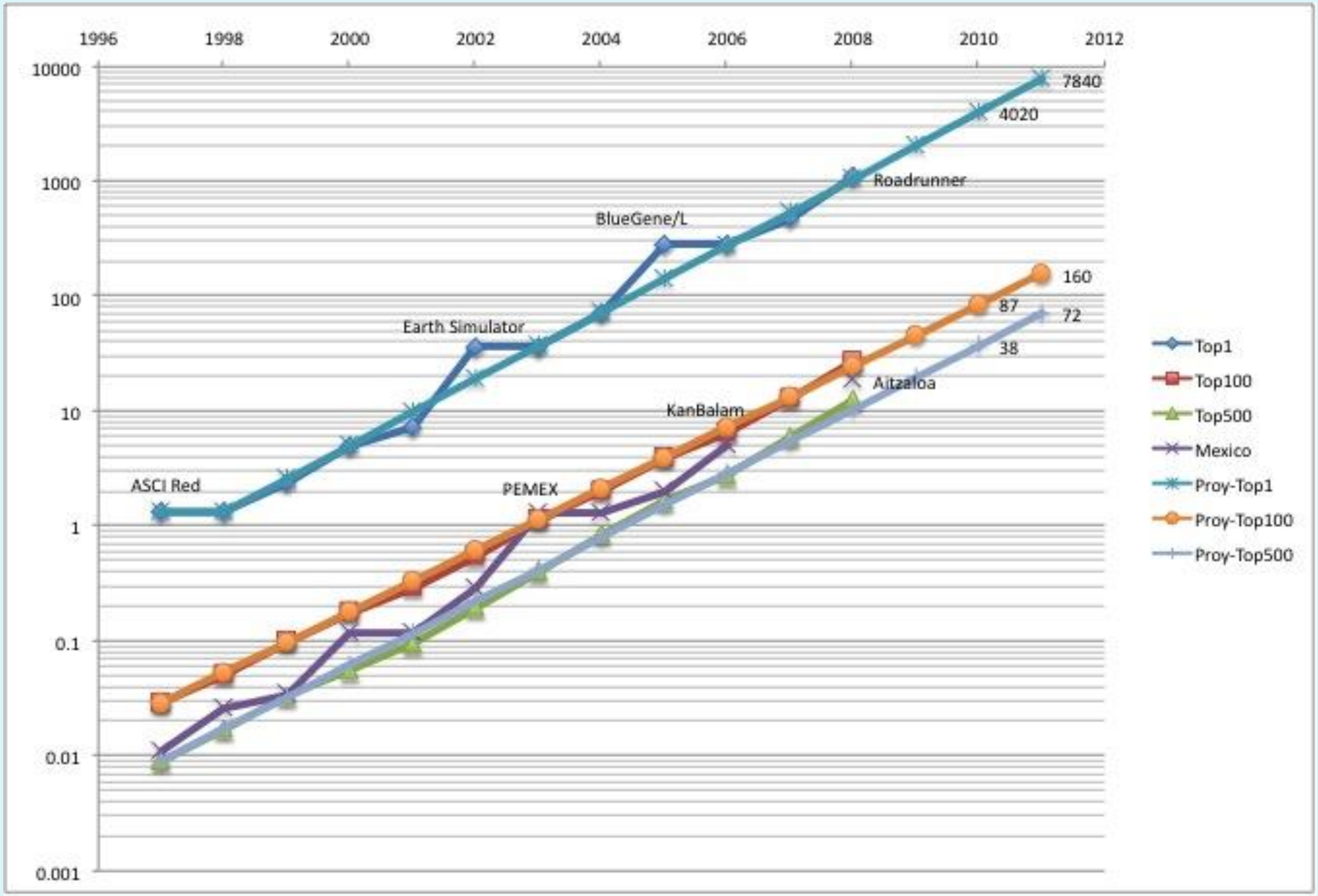
Cinvestav-Abacus



Abacus Centre for Applied Mathematics and High Performance Computing
(supercomputer will start operation by mid may 2015, this is in 3 weeks!!)

- Campus in Km 38.5 México-Toluca highway
- 15 mins from Santa Fe (Mexico City) and Toluca, 5 mins from ININ.
- **Computer with: (430 Tflops)**
- 9,000 Haswell Intel cores 2.6 Ghz
- 4 GB RAM per core (36,000 GB)
- 1,500 TB HD
- 100 GPUs Kepler K40 (200,000 cores)
- Water cooled

Evolution of the largest Supercomputers in Mexico



Abacus–ININ Wide Band connection

SCT SECRETARÍA DE COMUNICACIONES Y TRANSPORTES

RED NIBA URBANA

UNIVERSIDAD DE GUADALAJARA

Estudio de Factibilidad para el Desarrollo de la Red NIBA Urbana en el Estado de México

Coordinación de la Sociedad de la Información y el Conocimiento
"e-México"
Secretaría de Comunicaciones y Transportes

Solicitante:

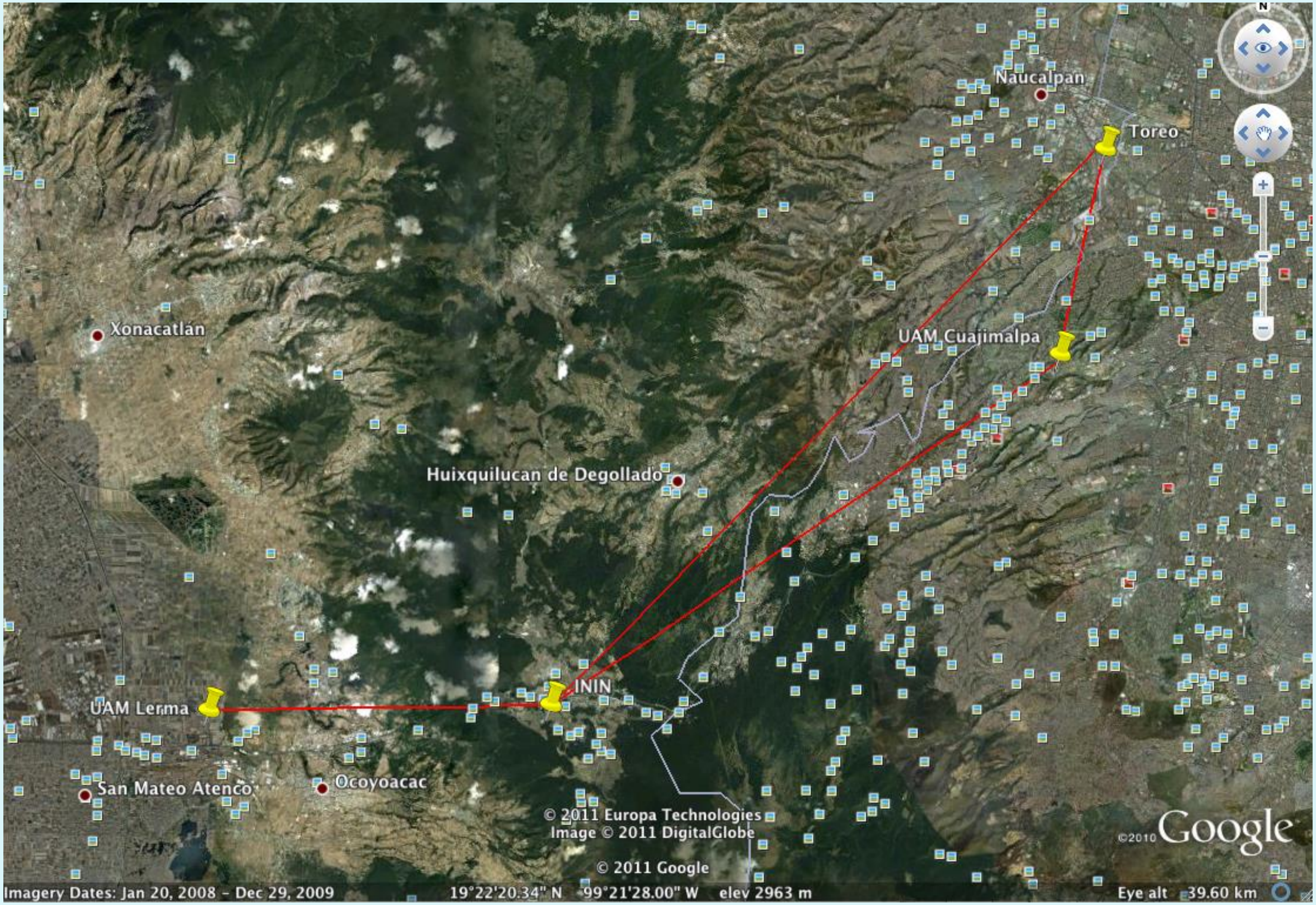
Mtra. Mónica Aspe Bernal
Coordinadora

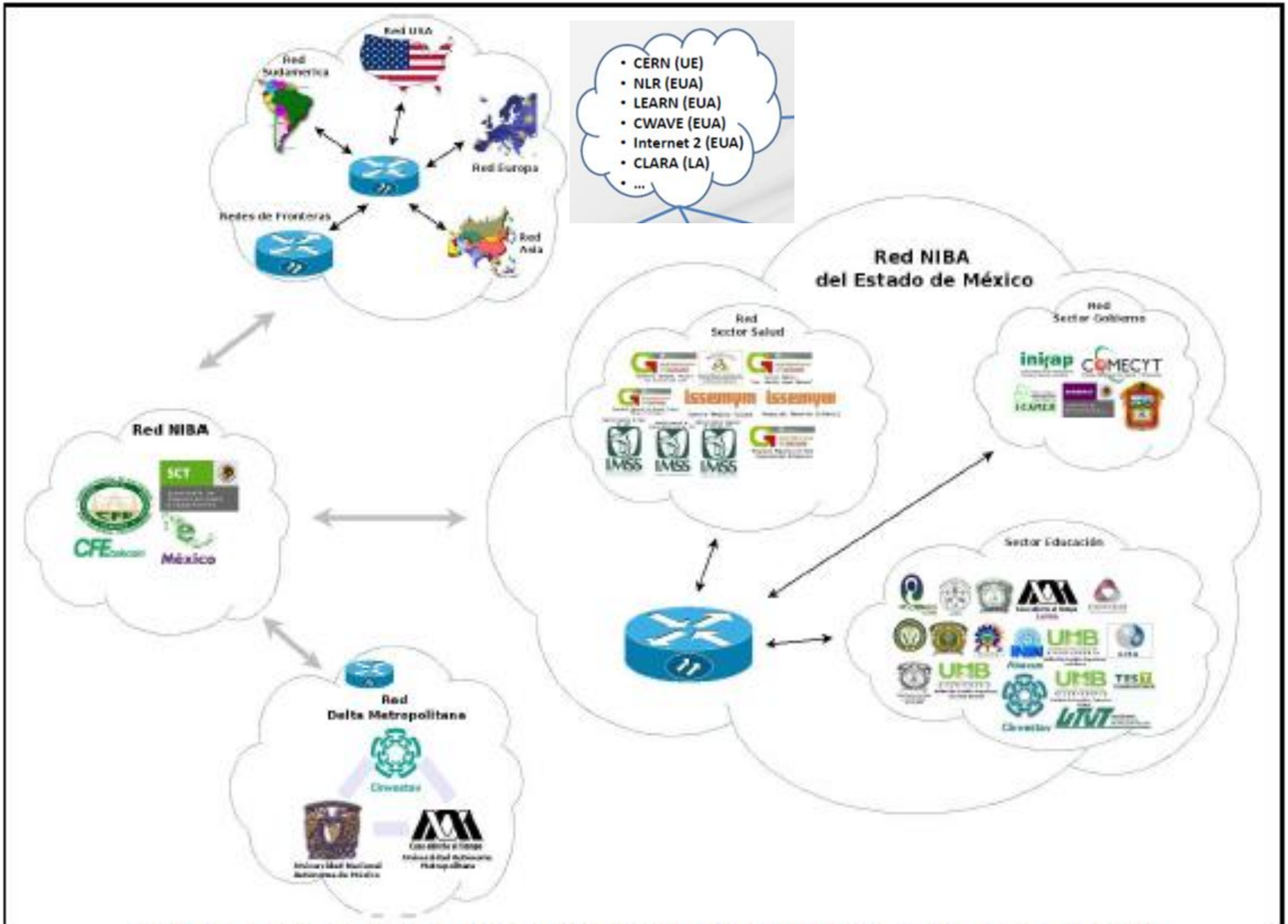
Responsables Globales del Proyecto

Cinvestav
Institución Responsable del Proyecto de Conectividad en el Estado de México
Dr. José Pablo René Asomoza
Director General

Responsable del Proyecto de Conectividad
Dr. Jaime Lazaro Klapp Escribano
Cinvestav-Abacus e ININ

Abacus and ININ location: we have Red Niba 10 GBs





Red Institucional del Estado de México, Red NIBA, Delta Metropolitana y Otras Redes de Clase Mundial

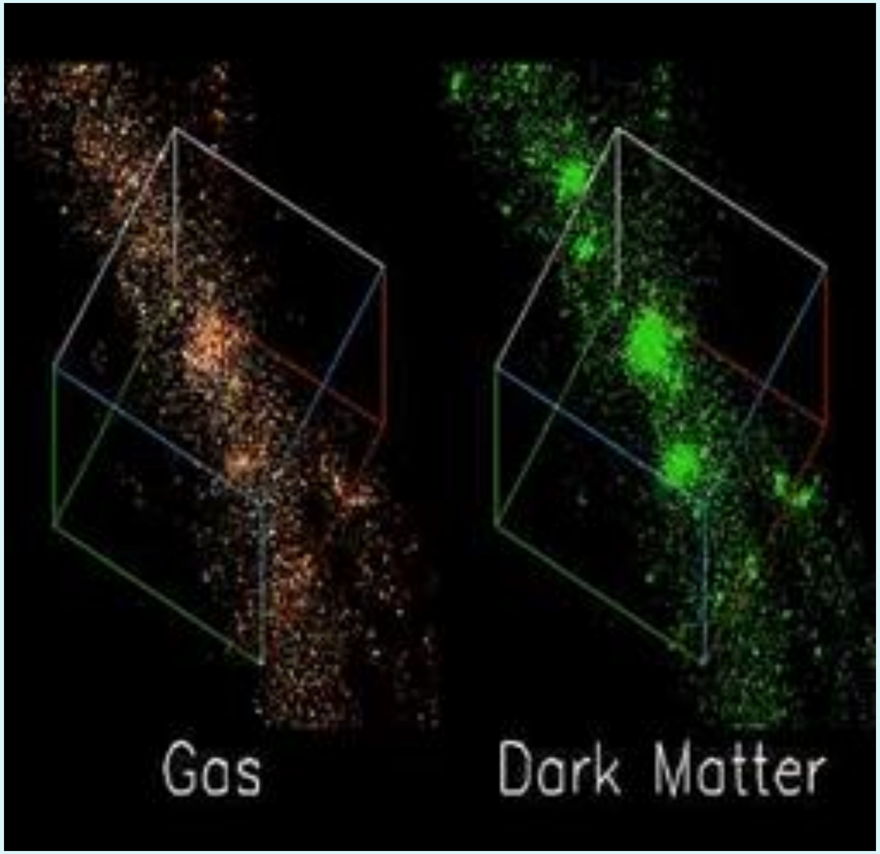
What is a fluid?

- Universe, galaxies, stars, etc.
- Atmosphere
- Ocean, tides and rivers
- Blood flow
- Car flow in cities and highways
- Electric energy flow
- Electron flow in a metal
- Population migration
- Oil through a fractured porous media
- Large amount of systems found in science, engineering and many human activities

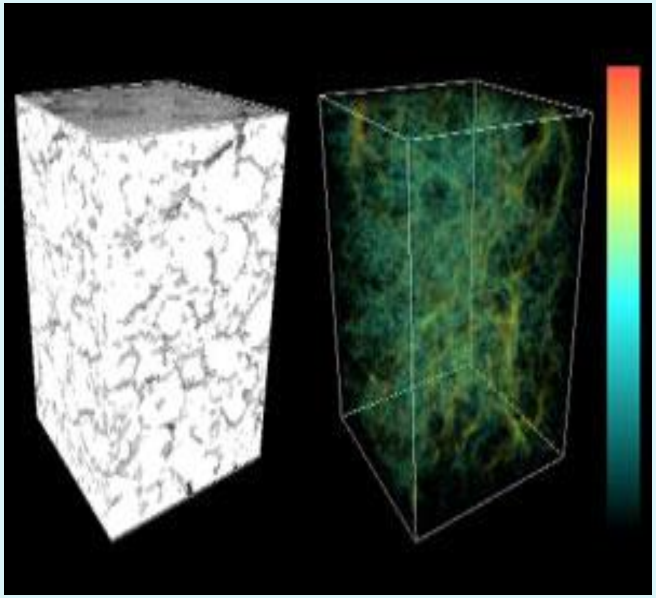
Why supercomputing?

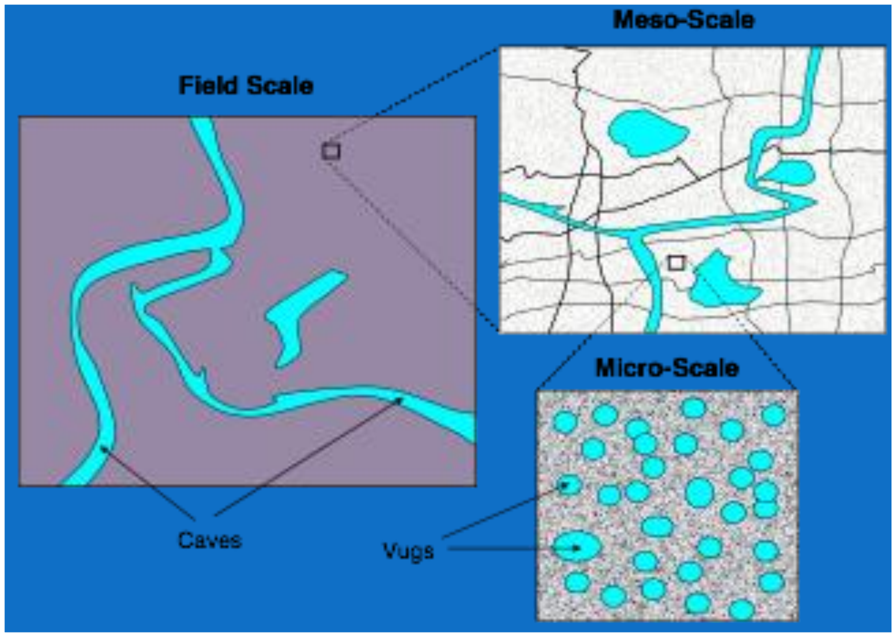
- Consider a computational 3D region divided in $1000 \times 1000 \times 1000 = 10^9$ cells.

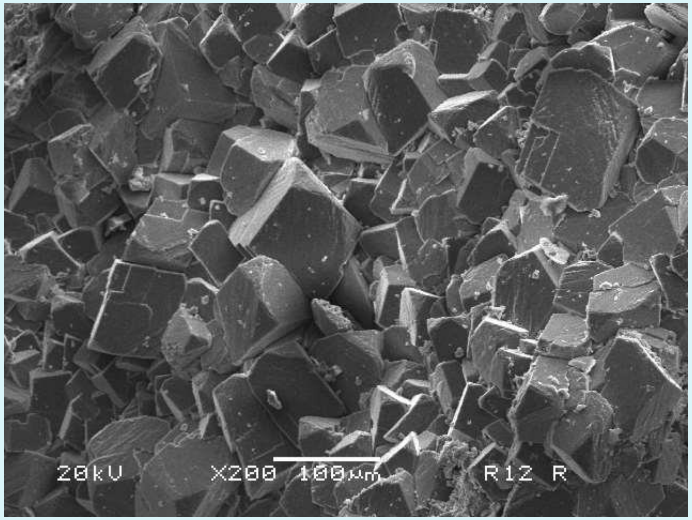
- If we have 10 equations, we have 10 variables that we have to solve at each point, this is 10^{10} variables or equations to solve at each time. A time evolution could require to solve the system 10^{4-5} times.

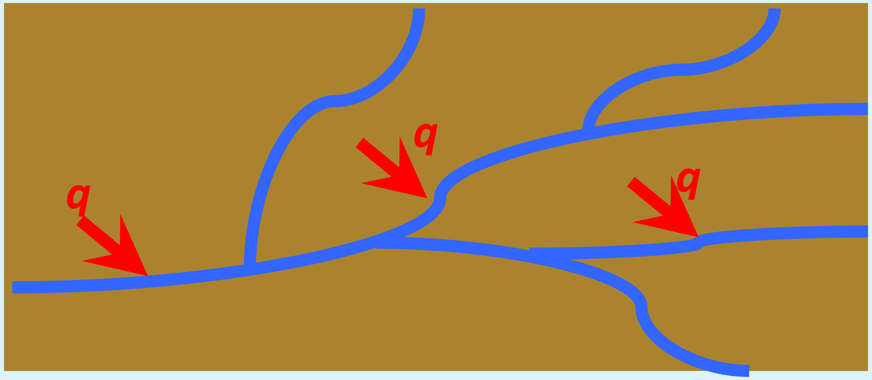


An example of our calculations: simulation of fractured oil reservoirs

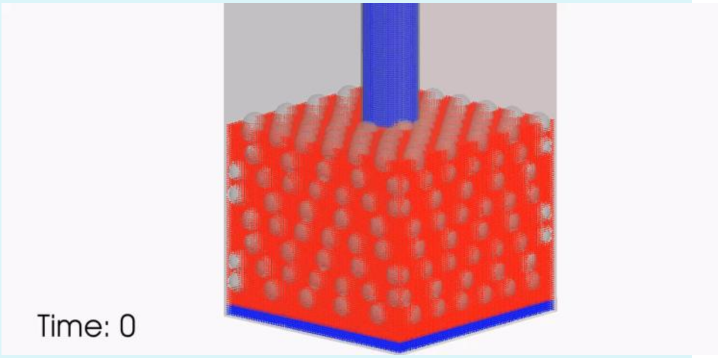


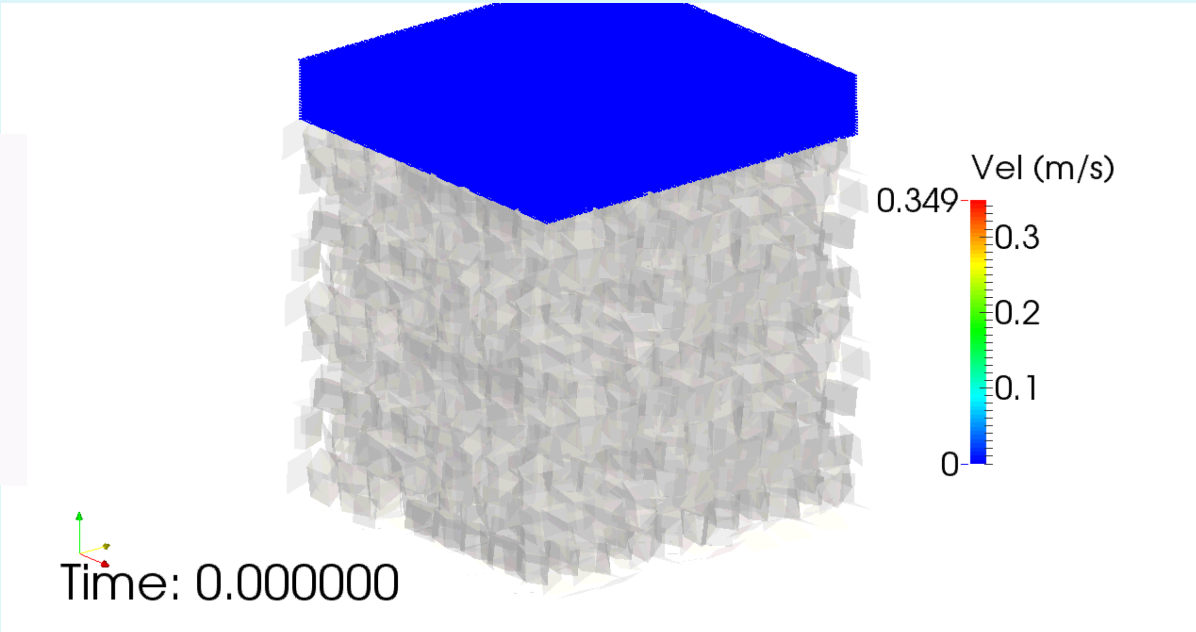


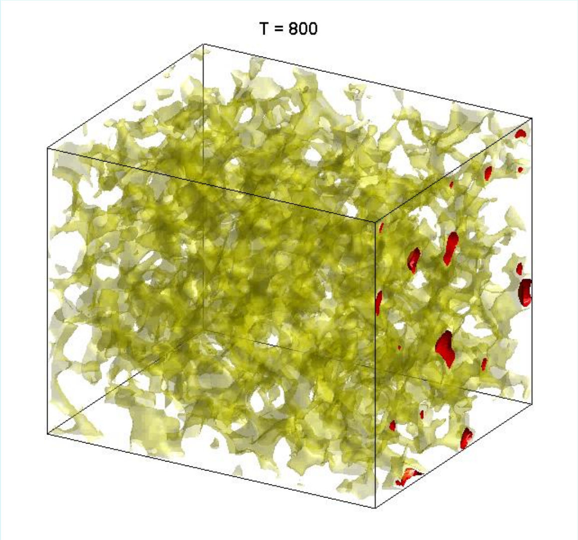


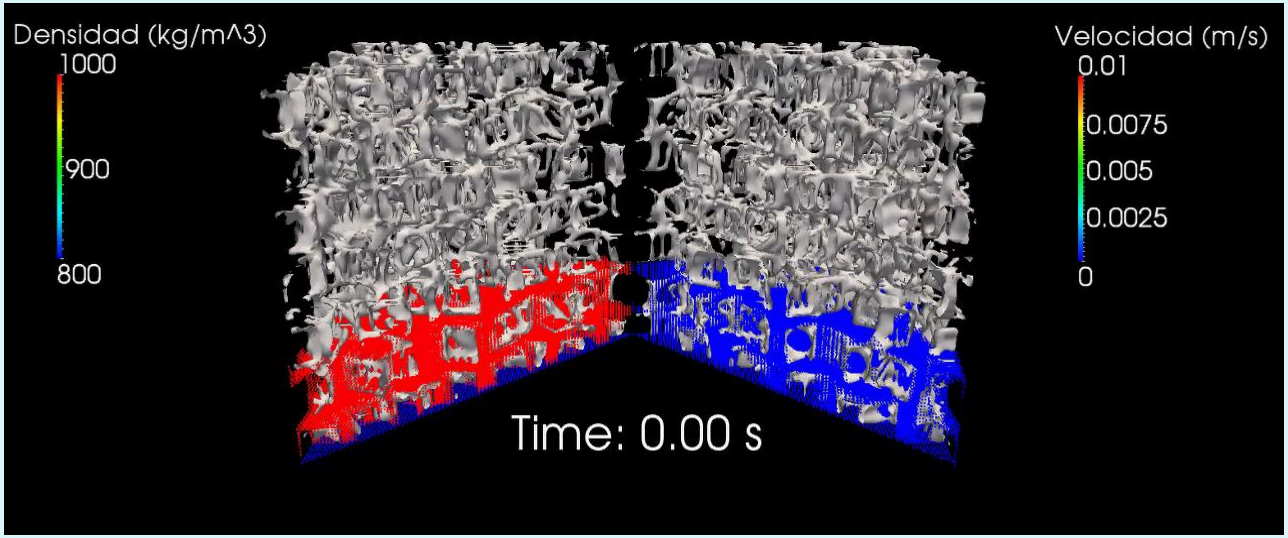


An example of our calculations









More examples: climate change, global warming and brain simulations

Example of Big Data: Sloan Digital Sky Survey

SDSS

The Data Volume of Data Release 12 is about 116 TB:

http://www.sdss.org/dr12/data_access/volume/

Another example: The High-Altitude Water Cherenkov Gamma-Ray Observatory

- The HAWC Observatory was officially inaugurated on Mar. 20, 2015, Sierra Negra volcano near Puebla, Mexico
- HAWC will produce massive amounts of data – upwards of

  10 Terabytes per day (3,650 TB per year)

  (http://www.rh.gatech.edu/news/390961/hawc-captures-evidence-powerful-cosmic-visitors)
- http://www.hawc-observatory.org/

Example (Abacus): Large CFD calculation with 500-1000 million particles

1 model (calculation): 10-100 TB

For a paper we require, say 5 runs: 50-500 TB.

For example, if in Abacus we produce 10 papers like this per year, we require 500-5000 TB per year. I estimate we will produce 1-5 PB per year. **Similar to HAWC.**

In 1-2 years we will run out of space (only 1500 TB), or get the data out of the computer to the final users. A large portion of our data can be erased or cleaned after some time.

Solution:

Leave all the data in the supercomputer, plot and visualize in the same machine, **only transfer images or videos. Many users want to analyze the data locally.**

Problem: Supercomputers usually ask users to move the data out from the machine in 2 weeks to few months.

We have to solve the data problem.

CFD projects applied to Science and Engineering in Cinvestav-Abacus and ININ.

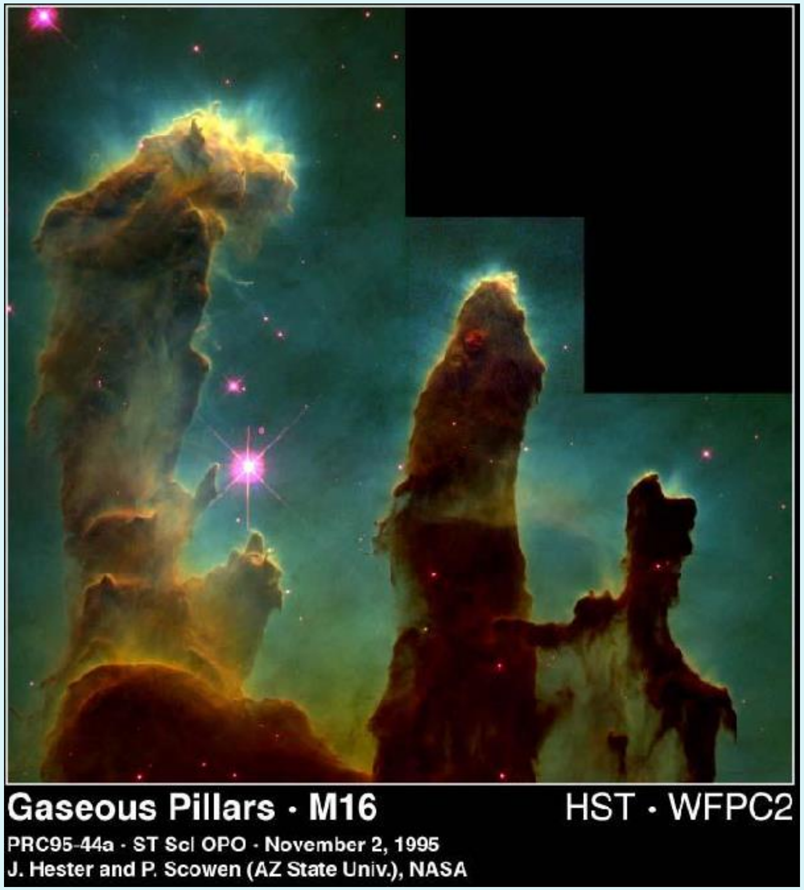
Astrophysics and Cosmology

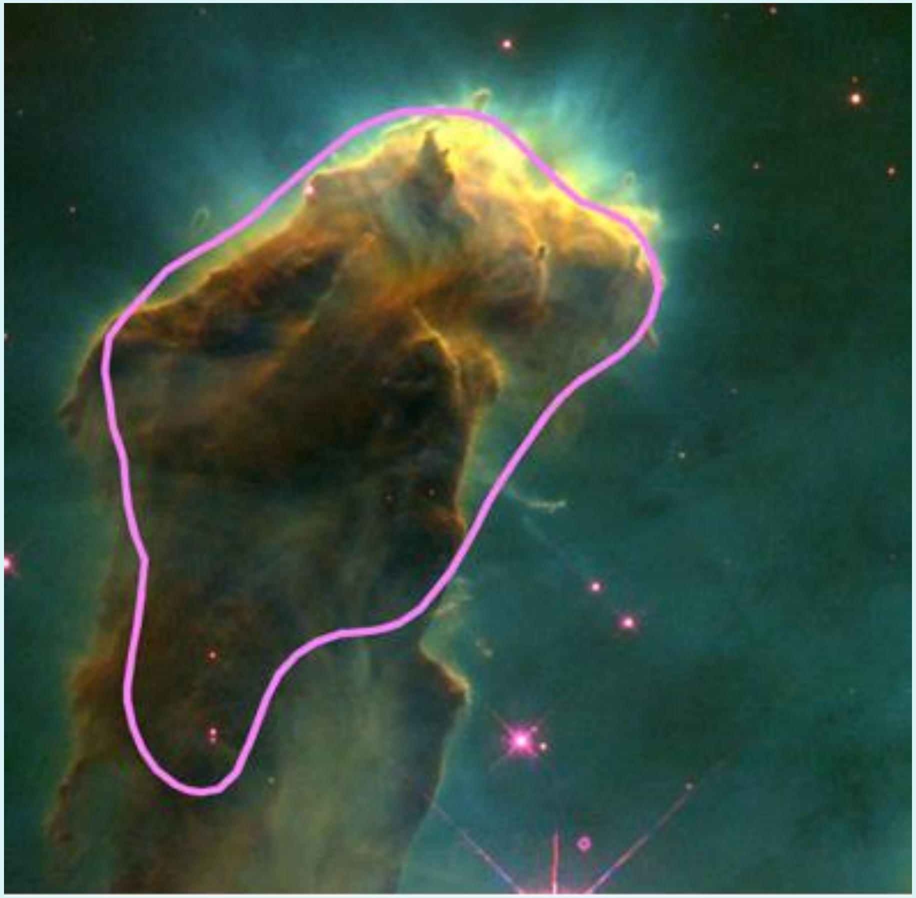
- Stellar evolution

- Galaxy collisions

- Cosmology

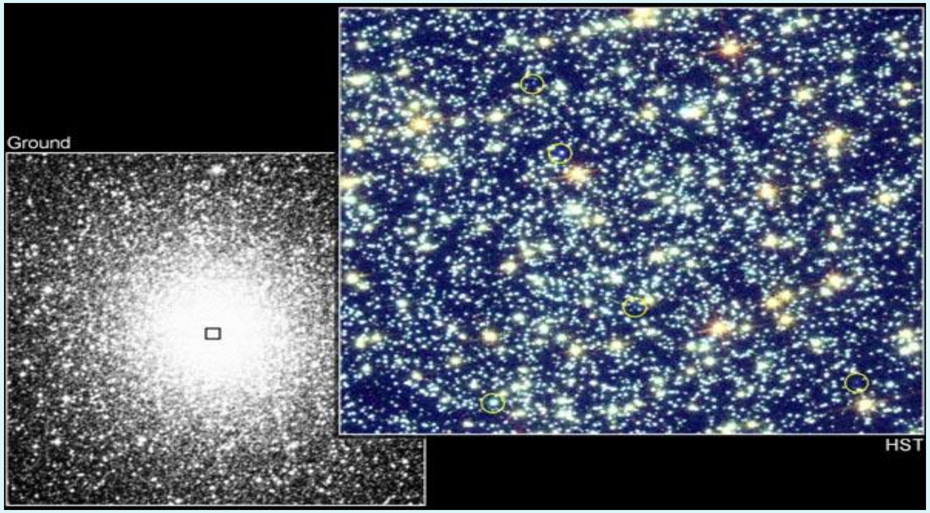
Star formation: molecular clouds





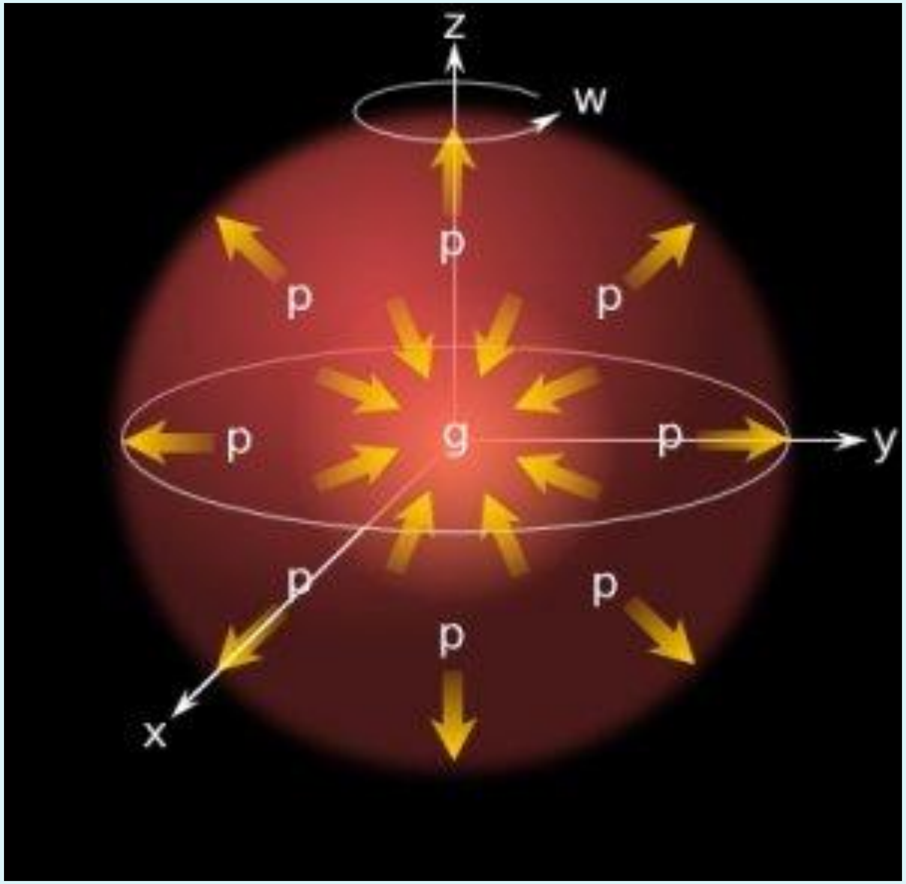

New
solar systems
forming in
Orion
Hubble Space Telescope photographs of
proto-planetary disks in the star-forming
nebula of Orion. http://hubblesite.org

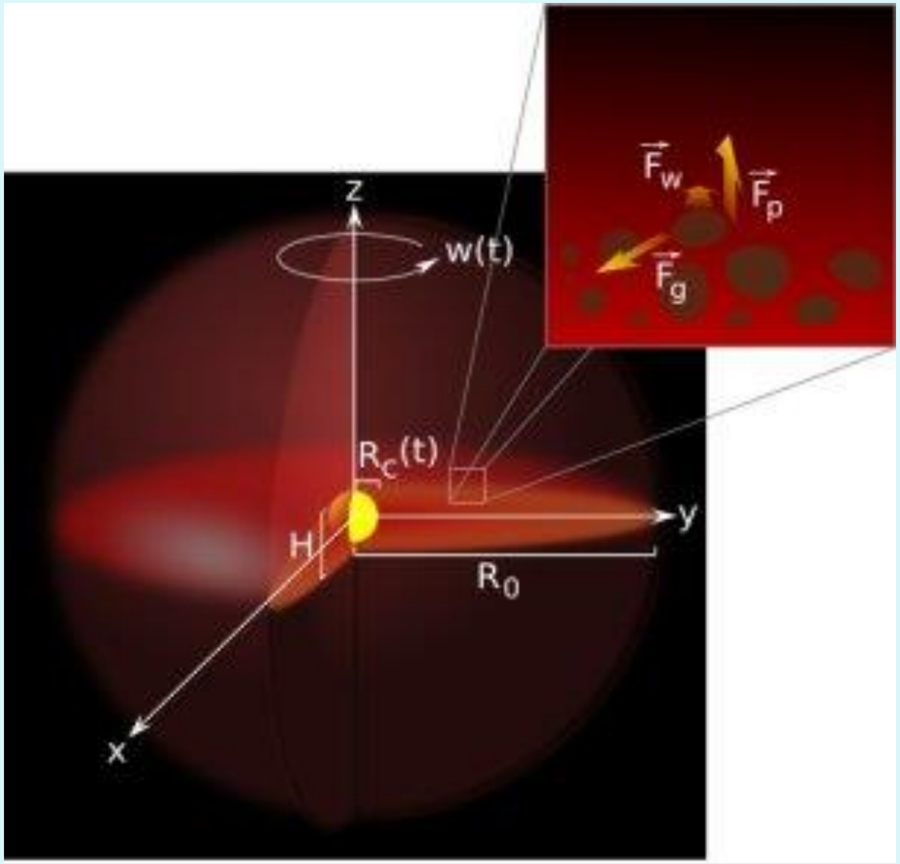
Molecular clouds → stellar clusters



Open clusters

Gravitational collapse that leads to fragmentation (star formation) occurs if gravity is more intense than pressure and other forces (magnetic)





Collapse and fragmentation of protostellar clouds

Galaxy collisions

- Galaxy collision simulations and comparison with observational data.

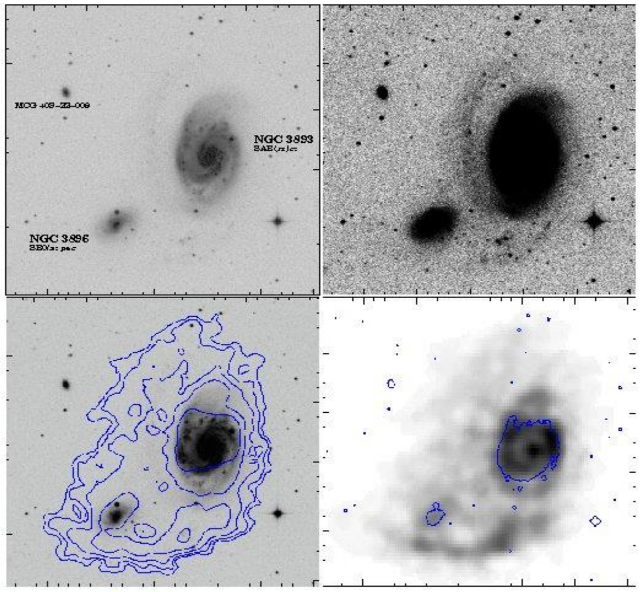
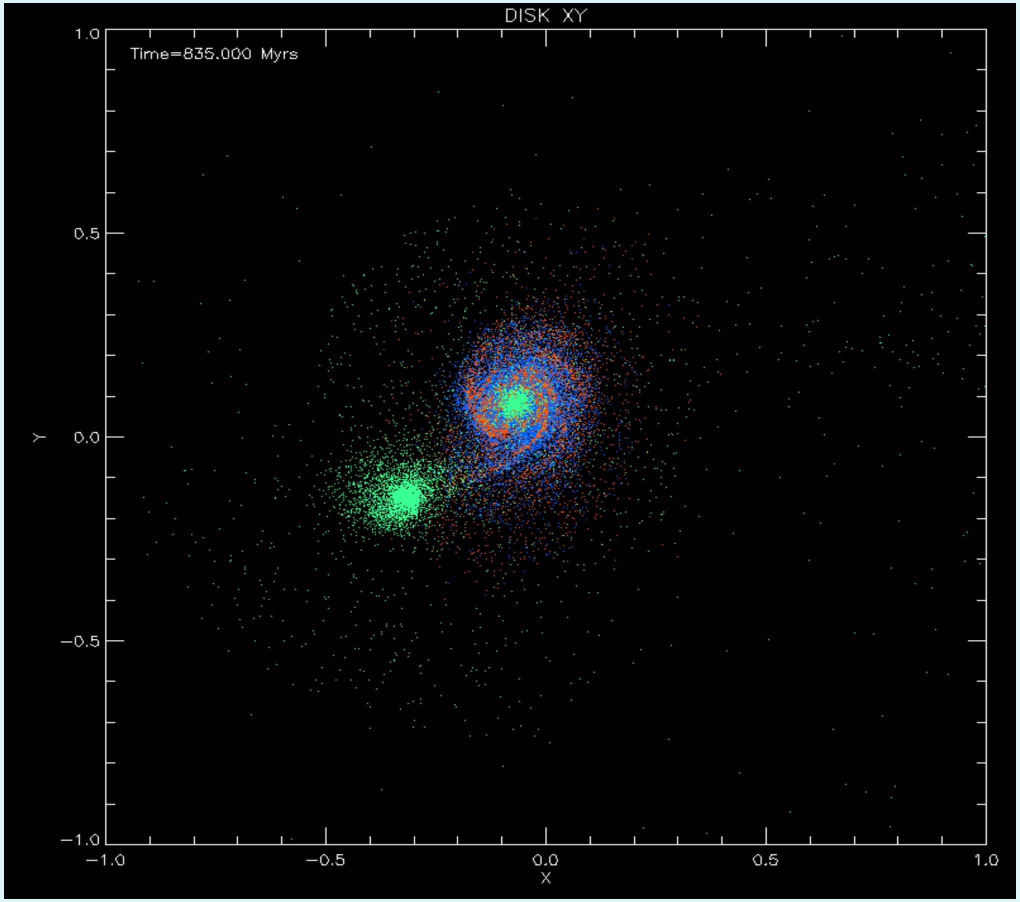


Figure 1. **The NGC 3893/6 System.**
HI: WSRT, 30″ resolution, contours=1×10^{20} cm^{-2} $\times 2^n$.
Optical: DSS, FOV=$10' \times 10'$.
Notes: Member of the Ursa Major Cluster. See also Fig. ??.
Reference: Verheijen, M. A. W., & Sancisi, R. 2001, A&A, 370, 765.



Numerical simulations of a galaxy collision that produced the Galaxy pair KPG 302.

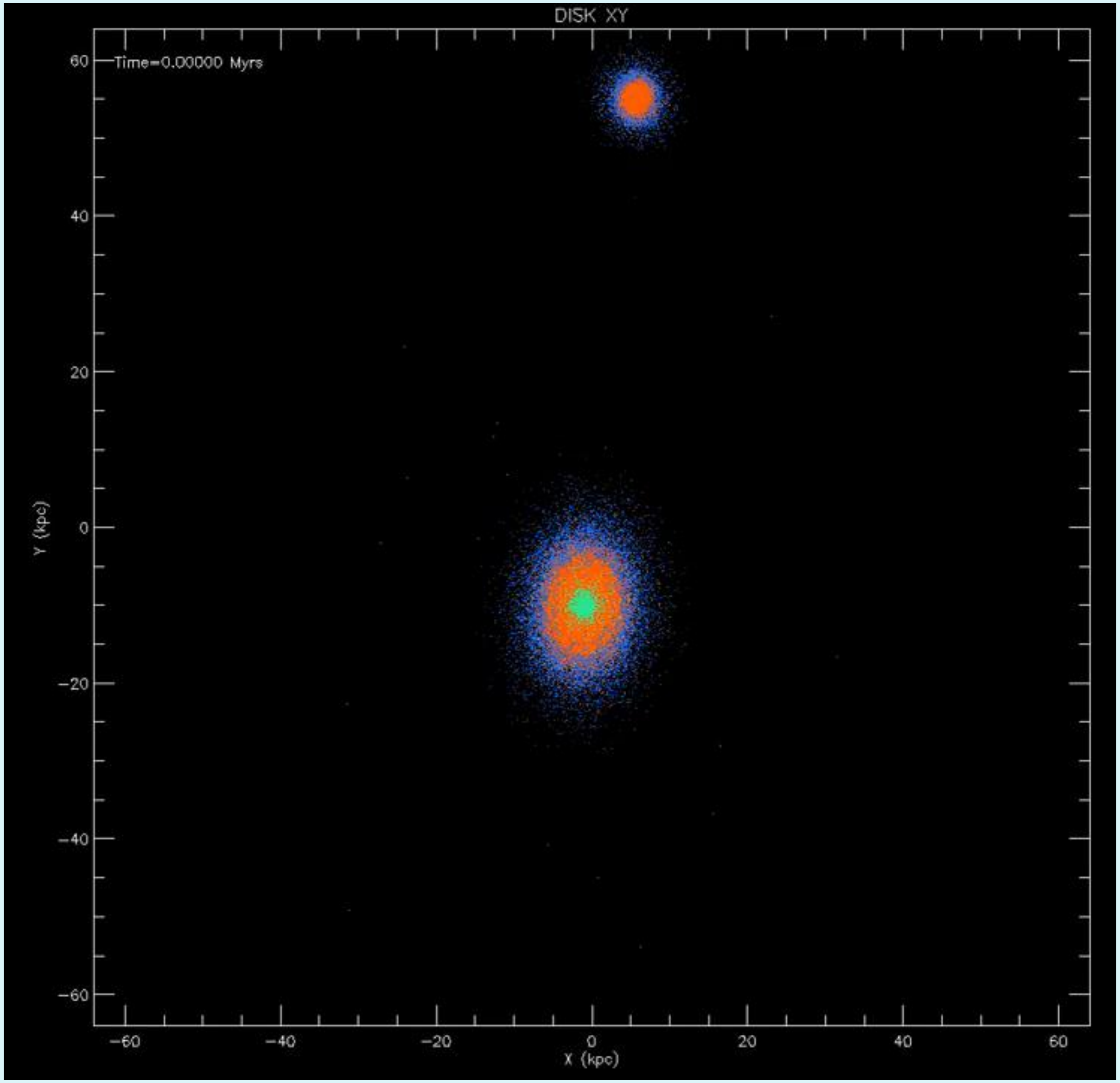
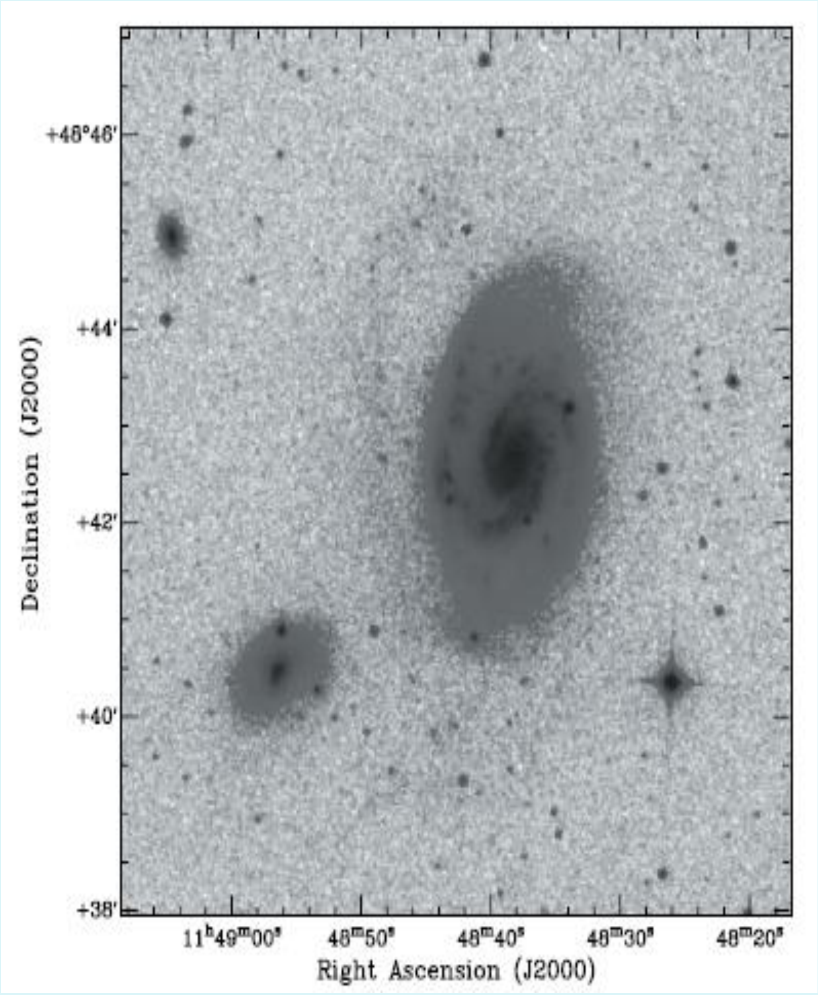
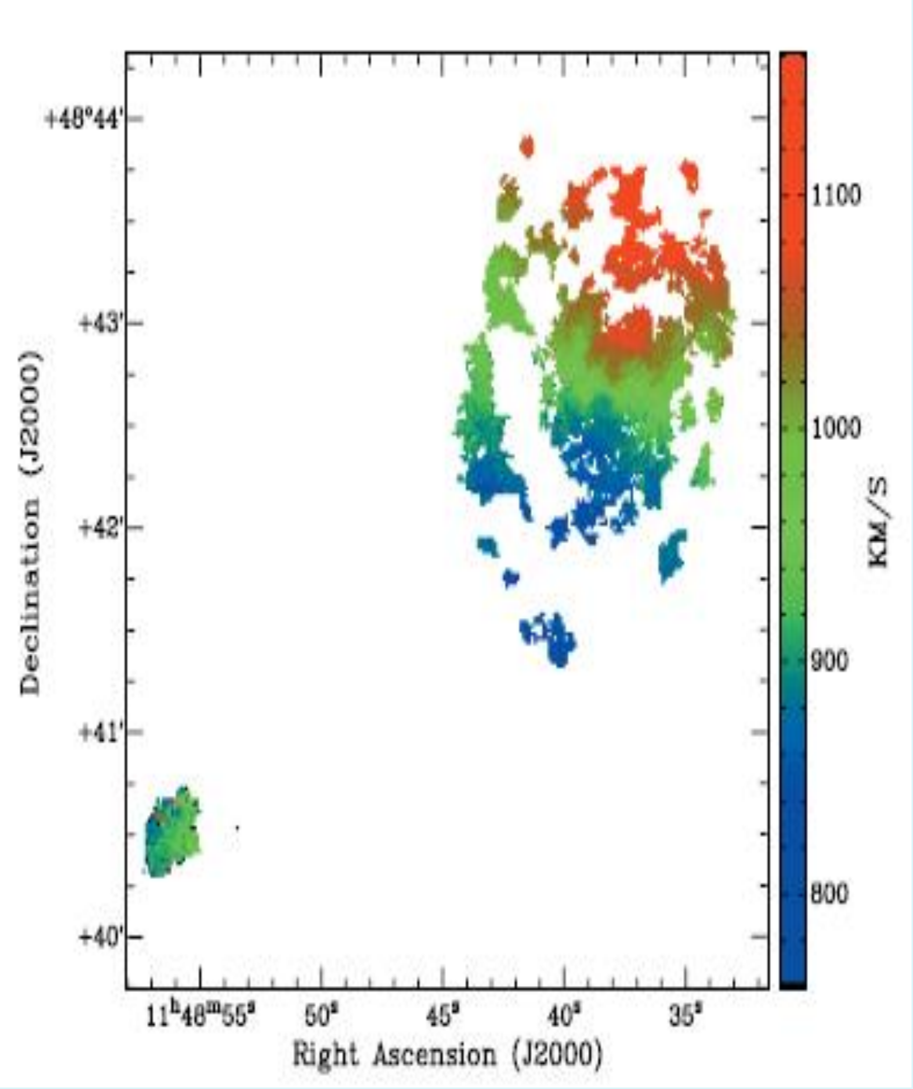
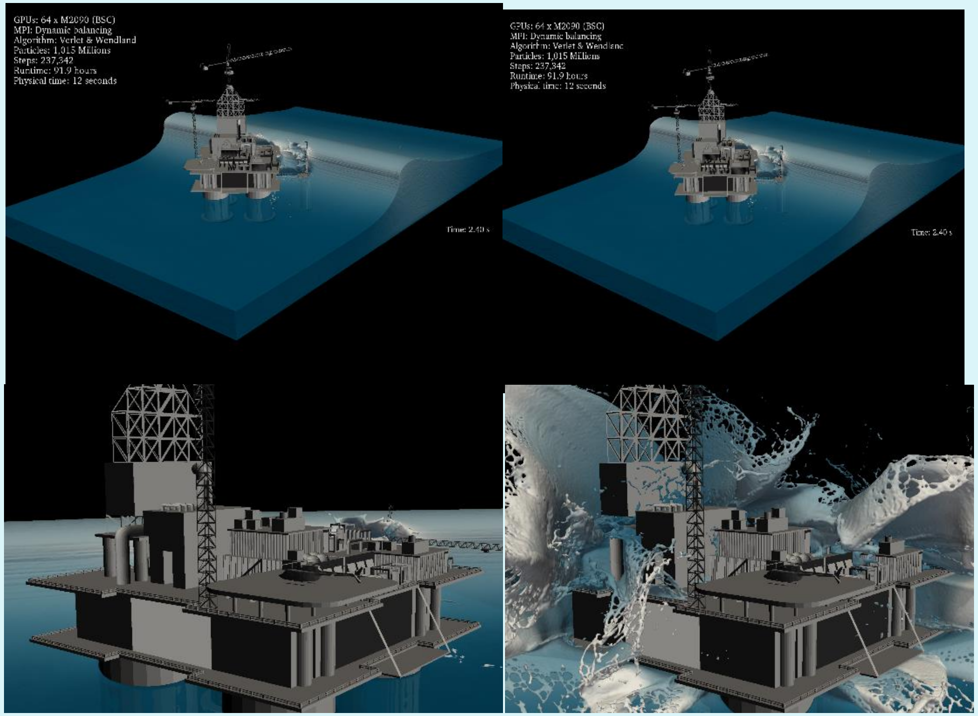


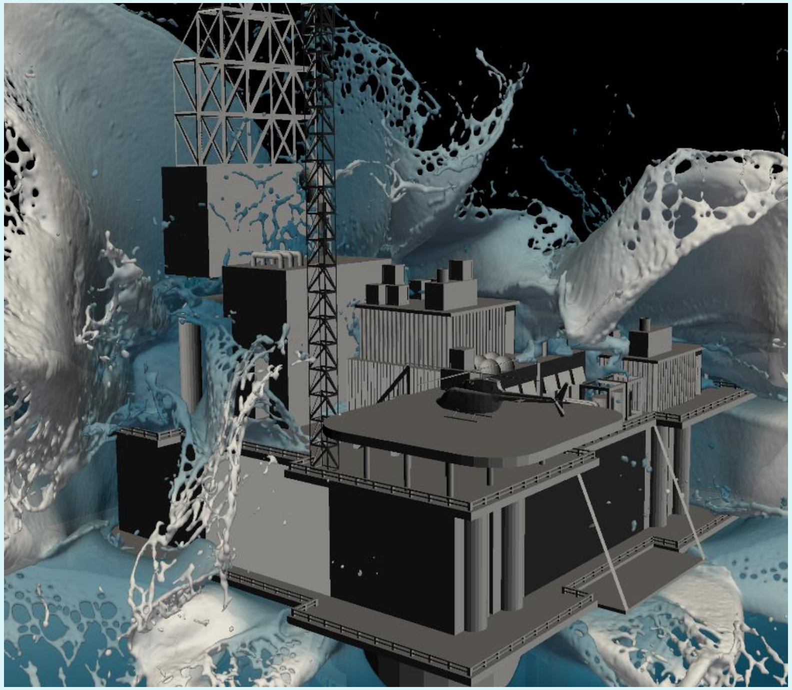
Image of the Galaxy pair KPG 302 (left) and numerical simulation of the interaction (right).

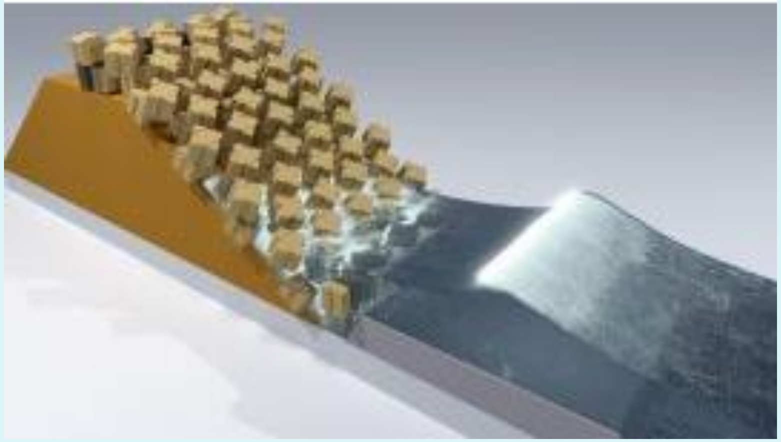




Examples of simulations

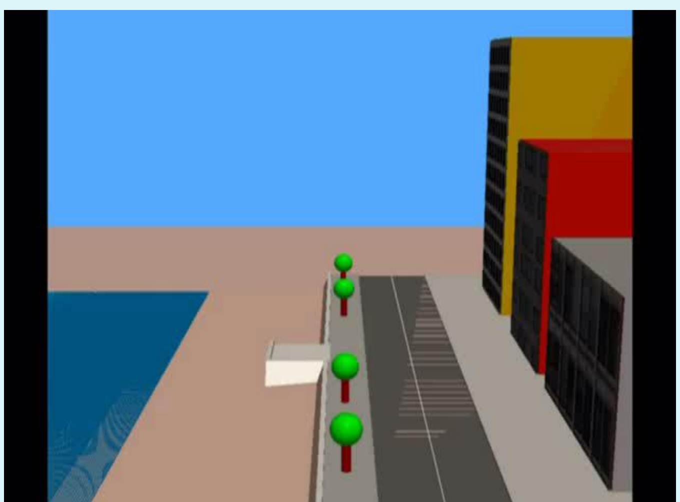


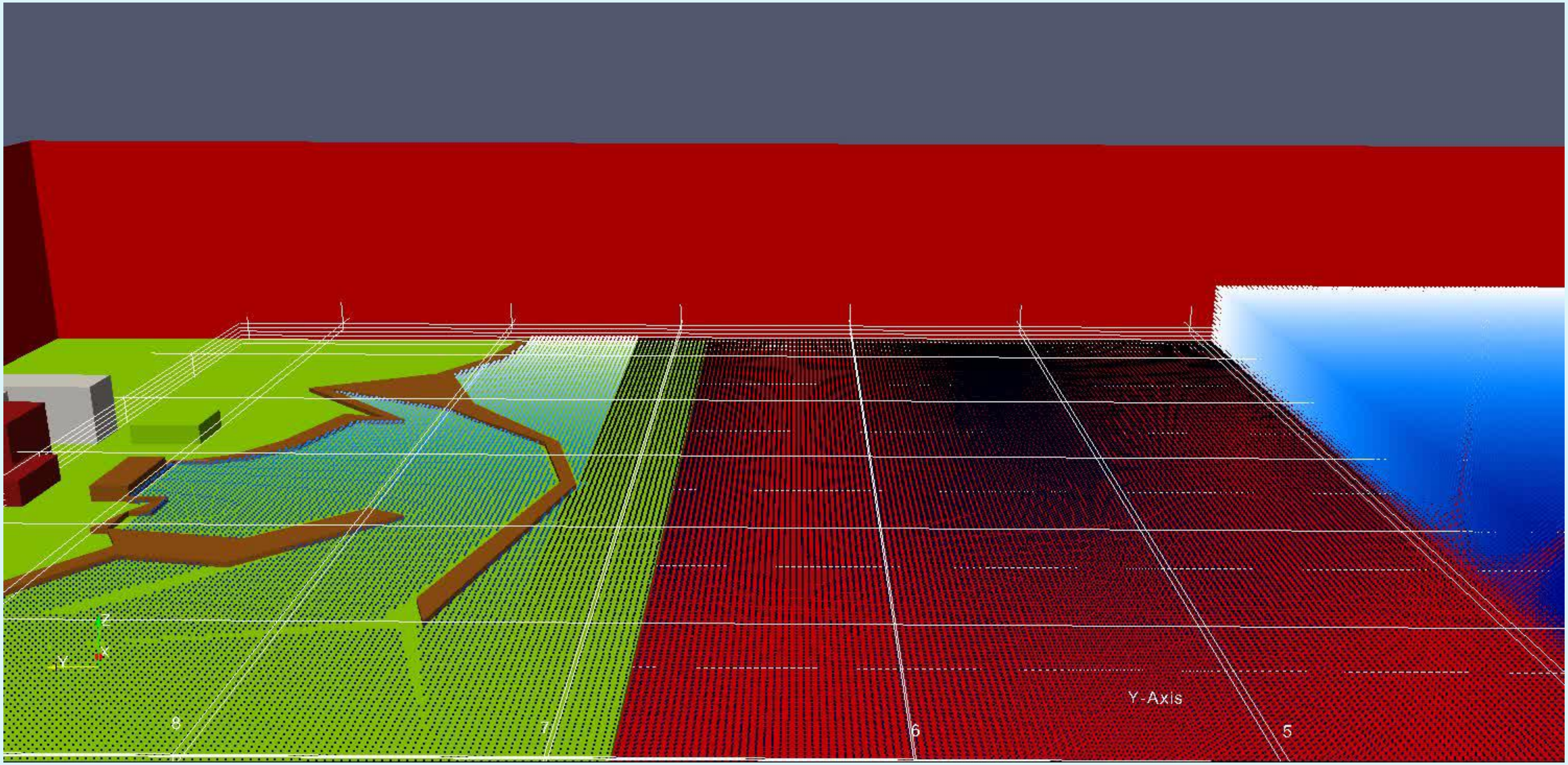




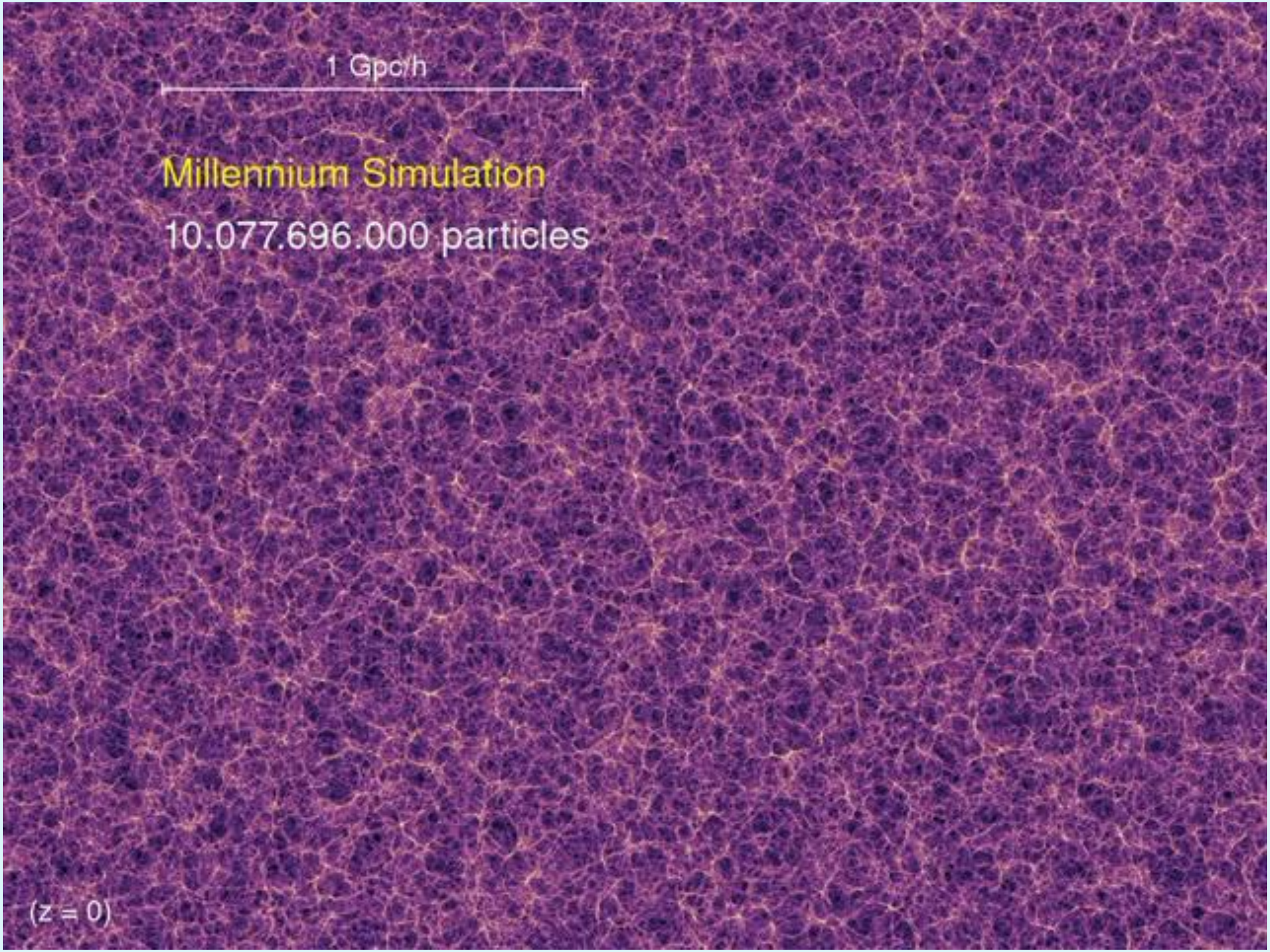
SPH simulations with GPUs of tsunamis (or large waves) vs the Mexican Laguna Verde nuclear plant in Veracruz.

Code: collaboration with U. Vigo and U. Manchester

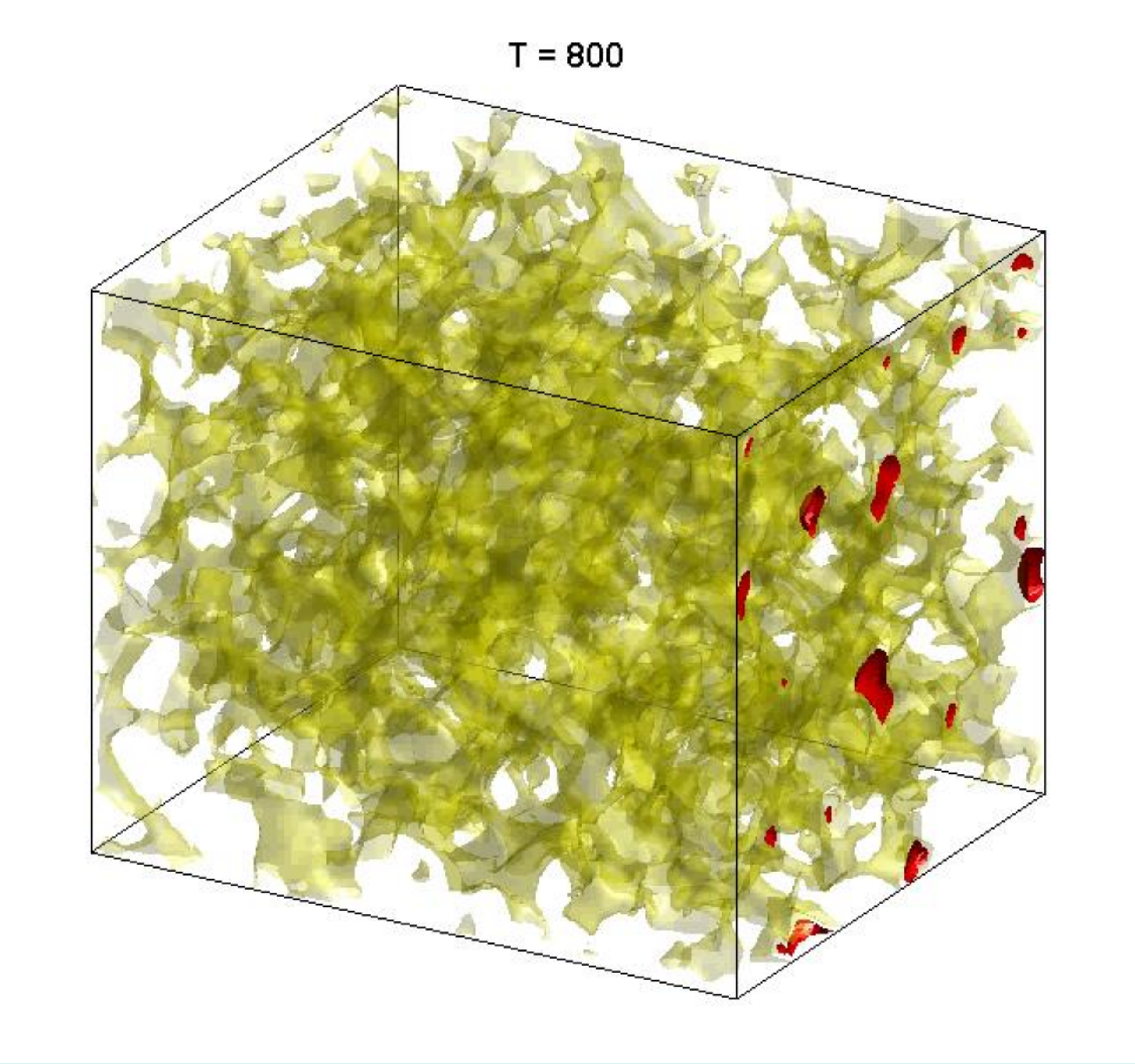



Y-Axis
8
7
6
5

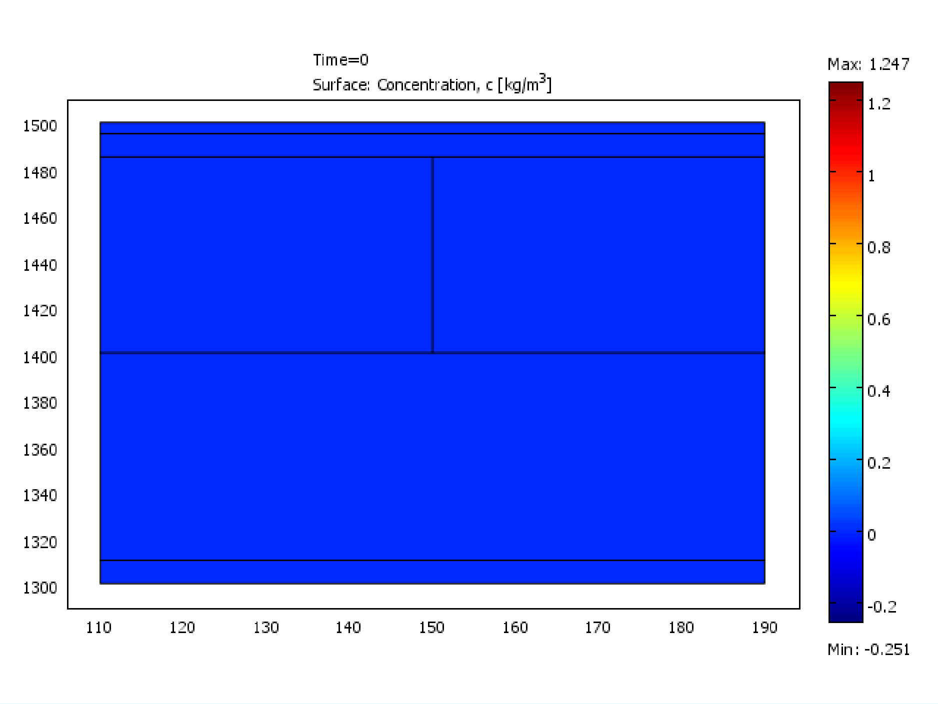
Cosmology

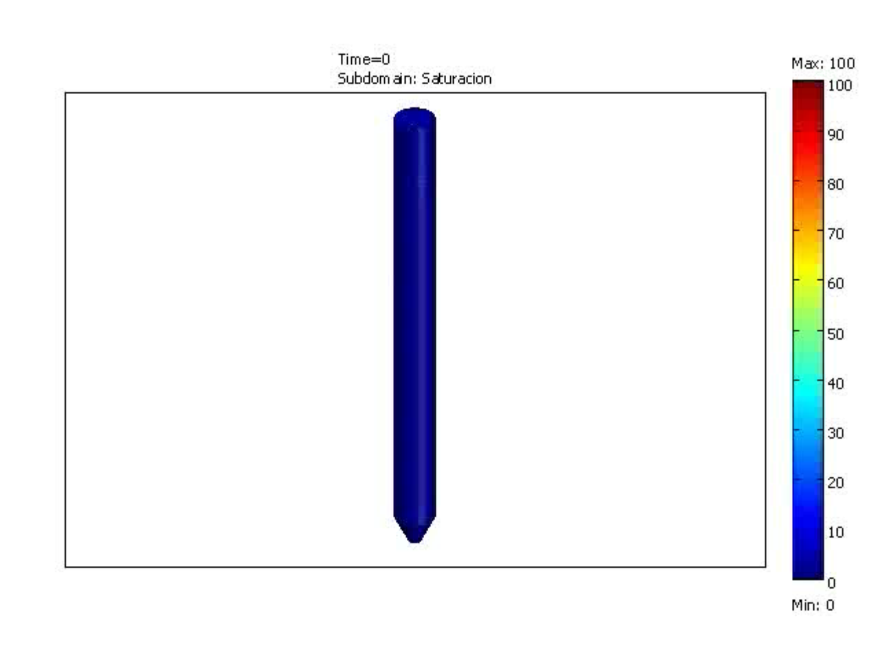


Digital petrochemistry



Multiplase flow through a porous media: Radioactive material (left) and Absorption of Arsenic and Selenium by biomass (right). Codes: COMSOL and Alya.





Uranium transport in a contaminated site (Peña Blanca, North of México)

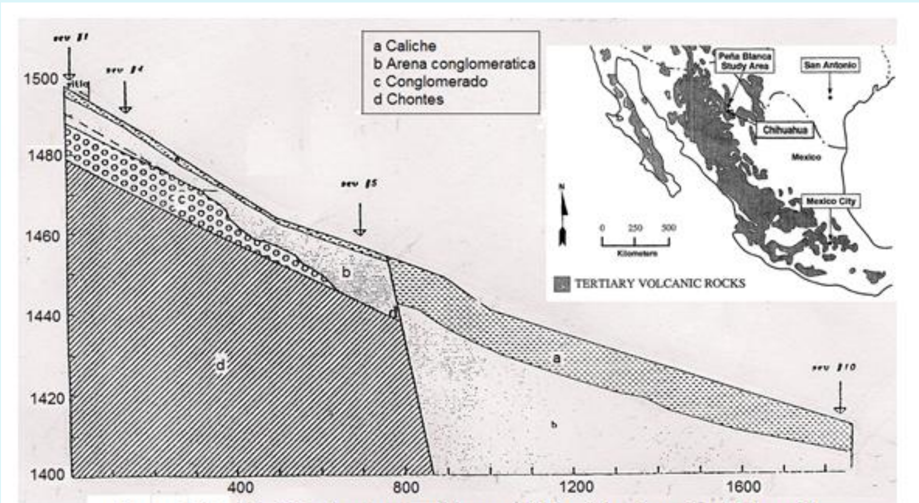
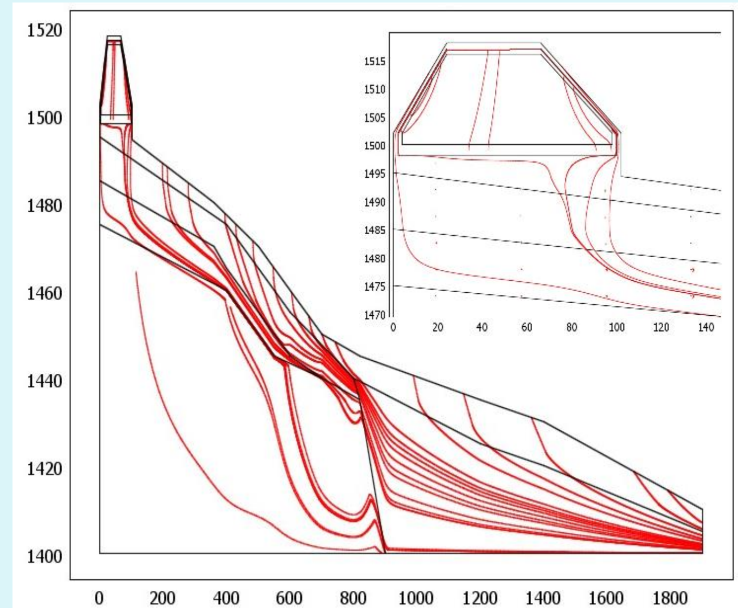


Figura 6.18a. Perfil de la distribución geológica del subsuelo de la región.
Tomada de Rojas (1996)



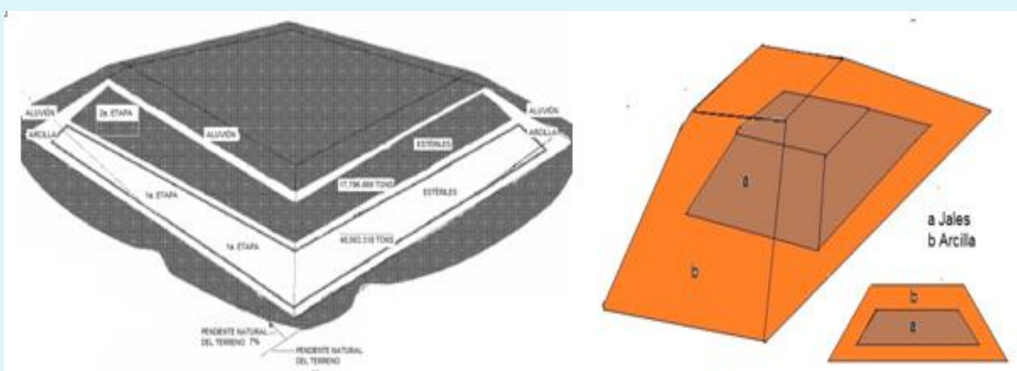
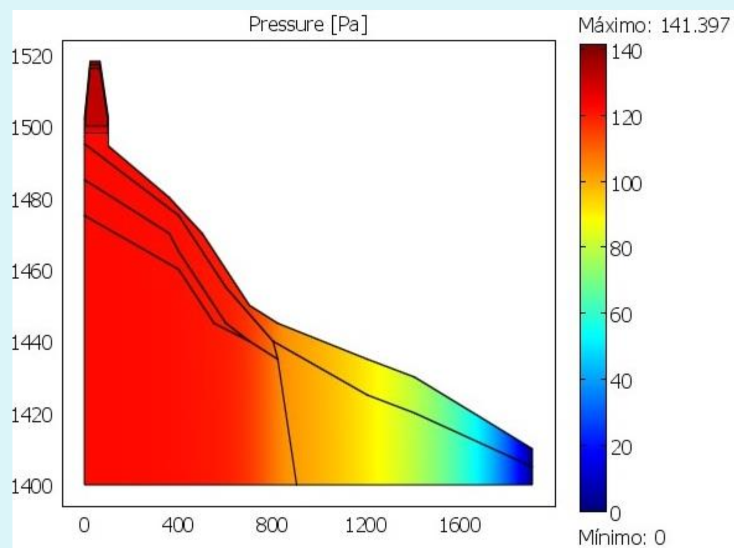
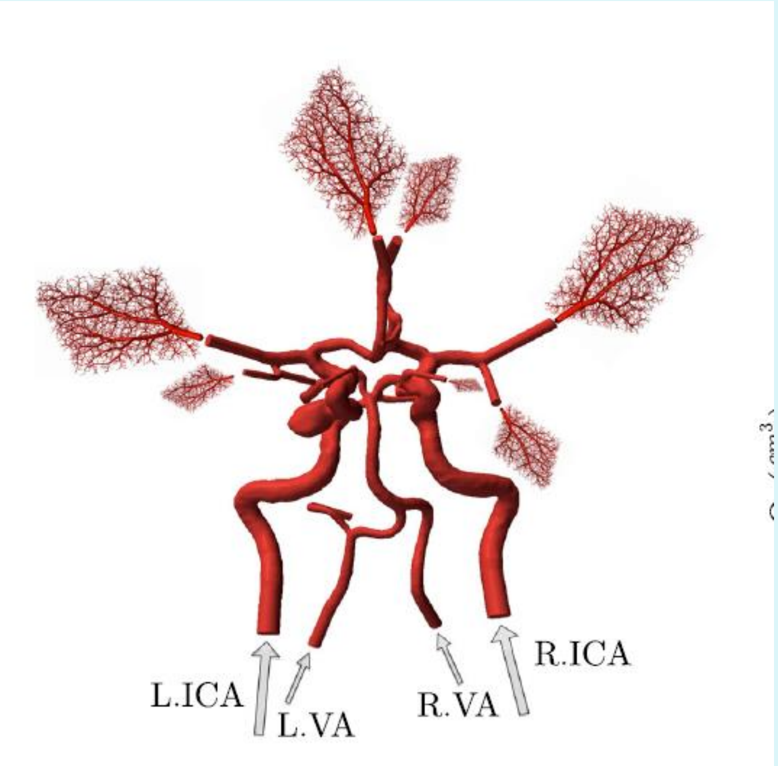
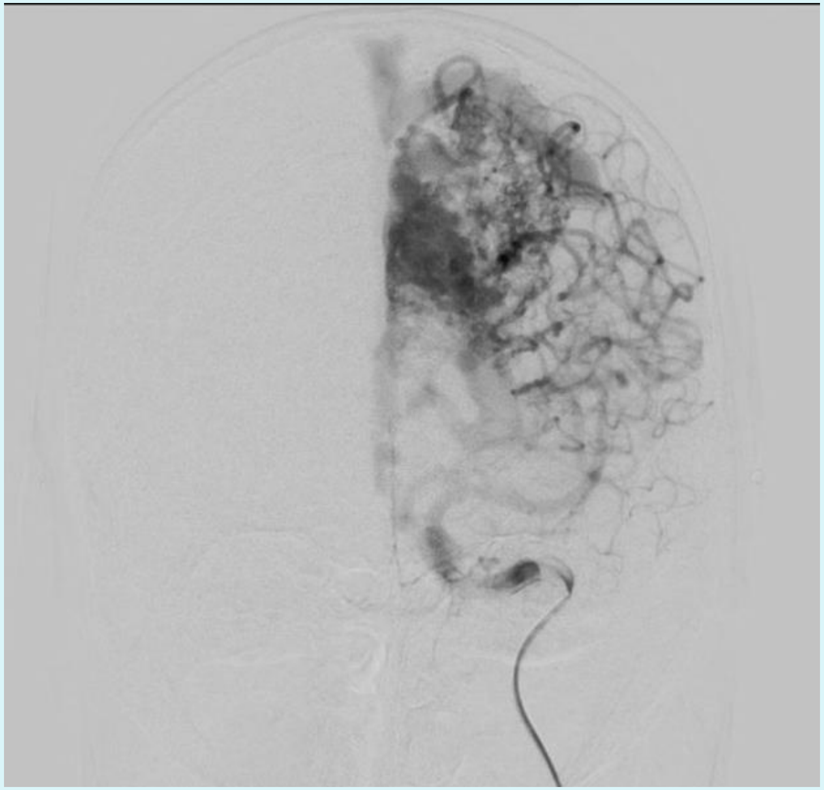


Figura 6.18b. Depósito de jales en el sitio de confinamiento.



Project: Hemodynamic simulations of brain malformations (MAVS), pre-treatment and comparison of patient variables. (Aid for helping surgeons in the treatment of Mavs, calculate risks for specific patients, better understanding the patient problem before surgery).

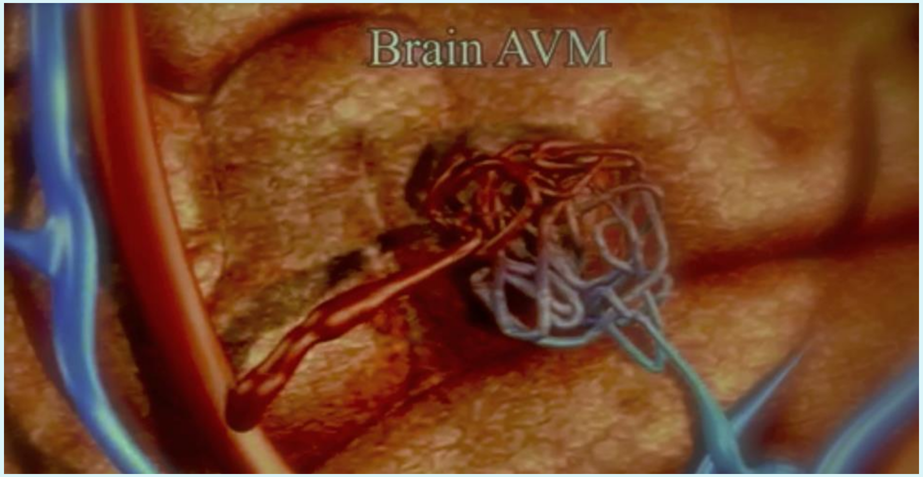


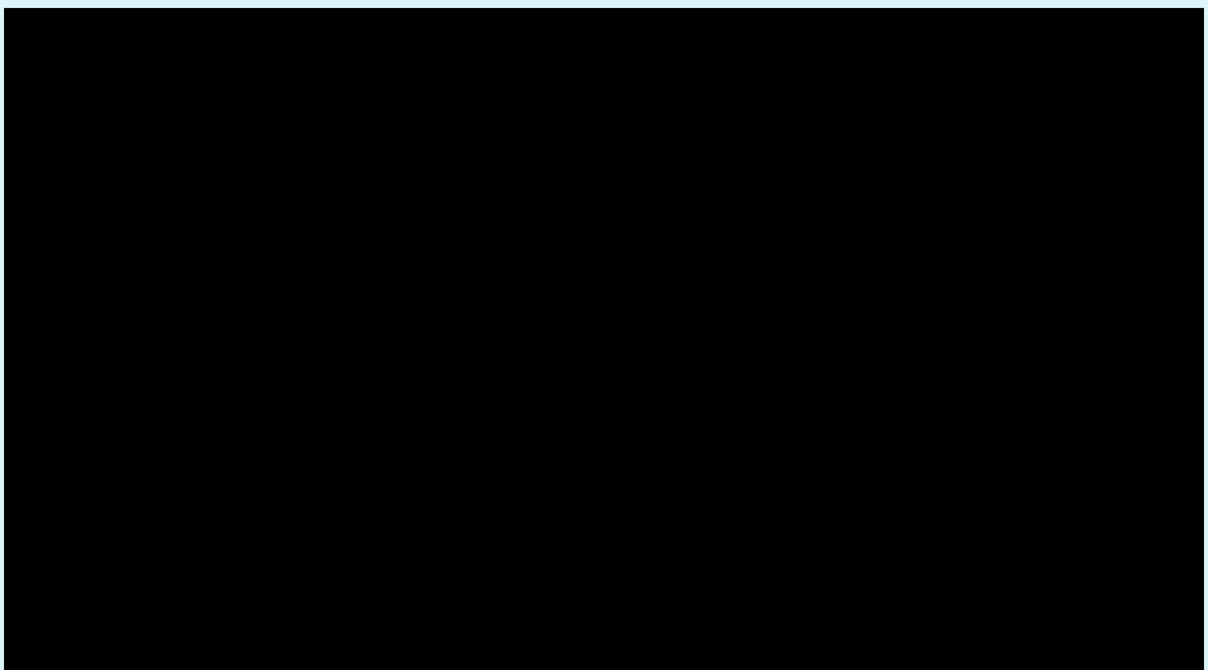


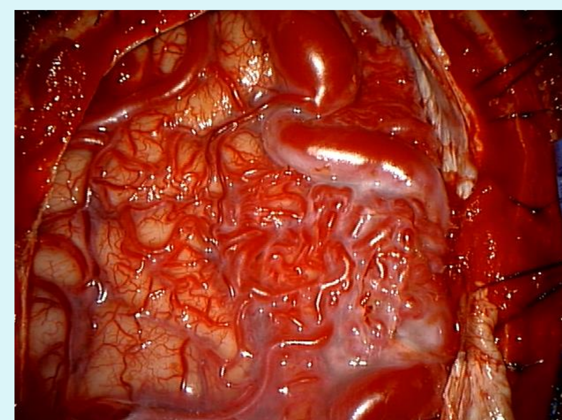
Brain arteries and arterioles system

Brain arteries and arterioles system

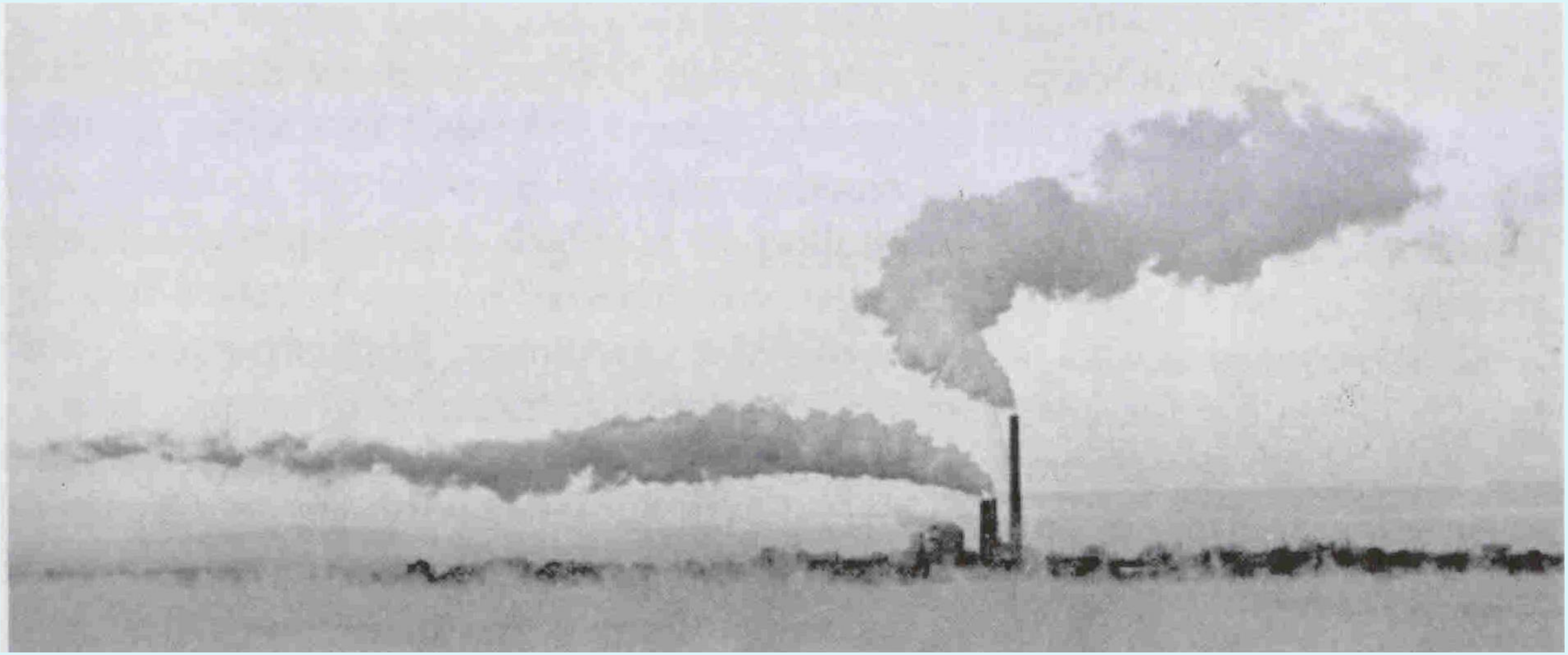
Brain malformation (MAVS)

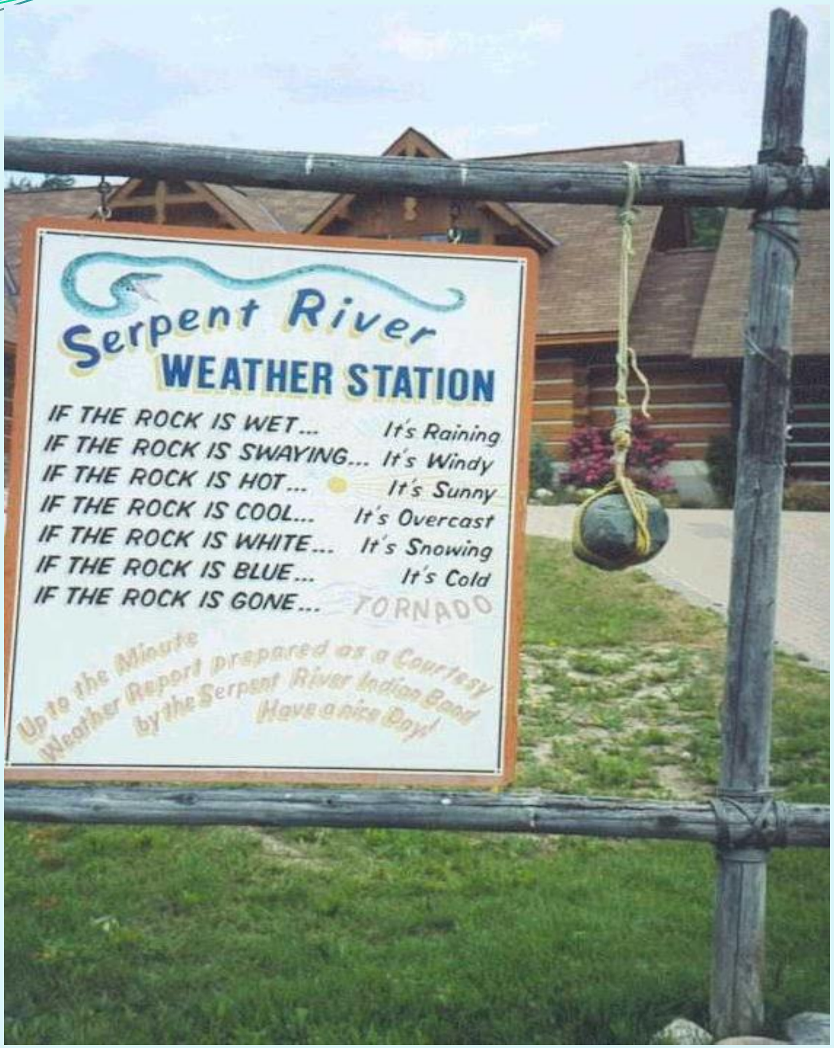


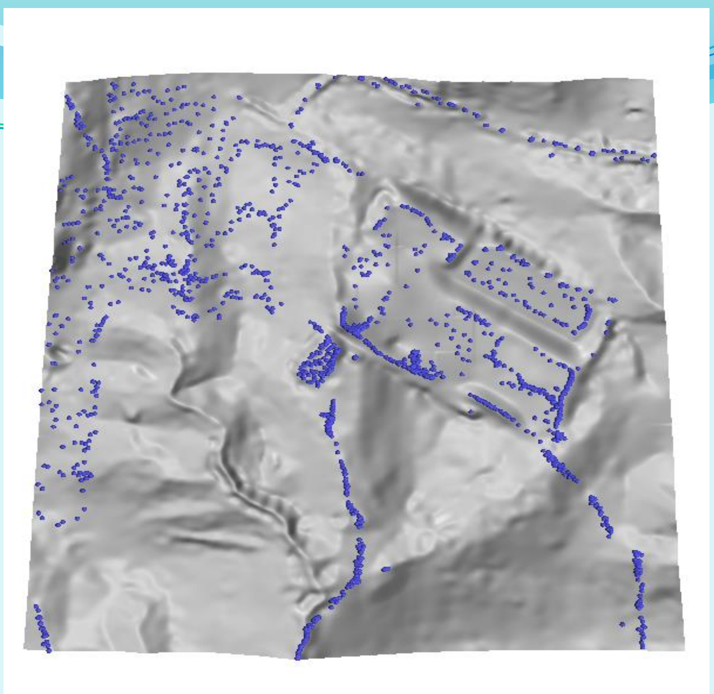


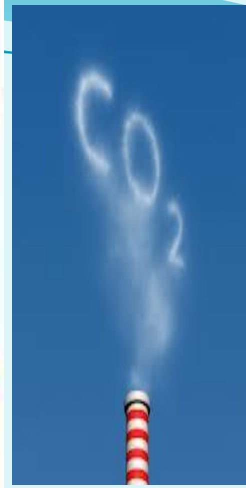


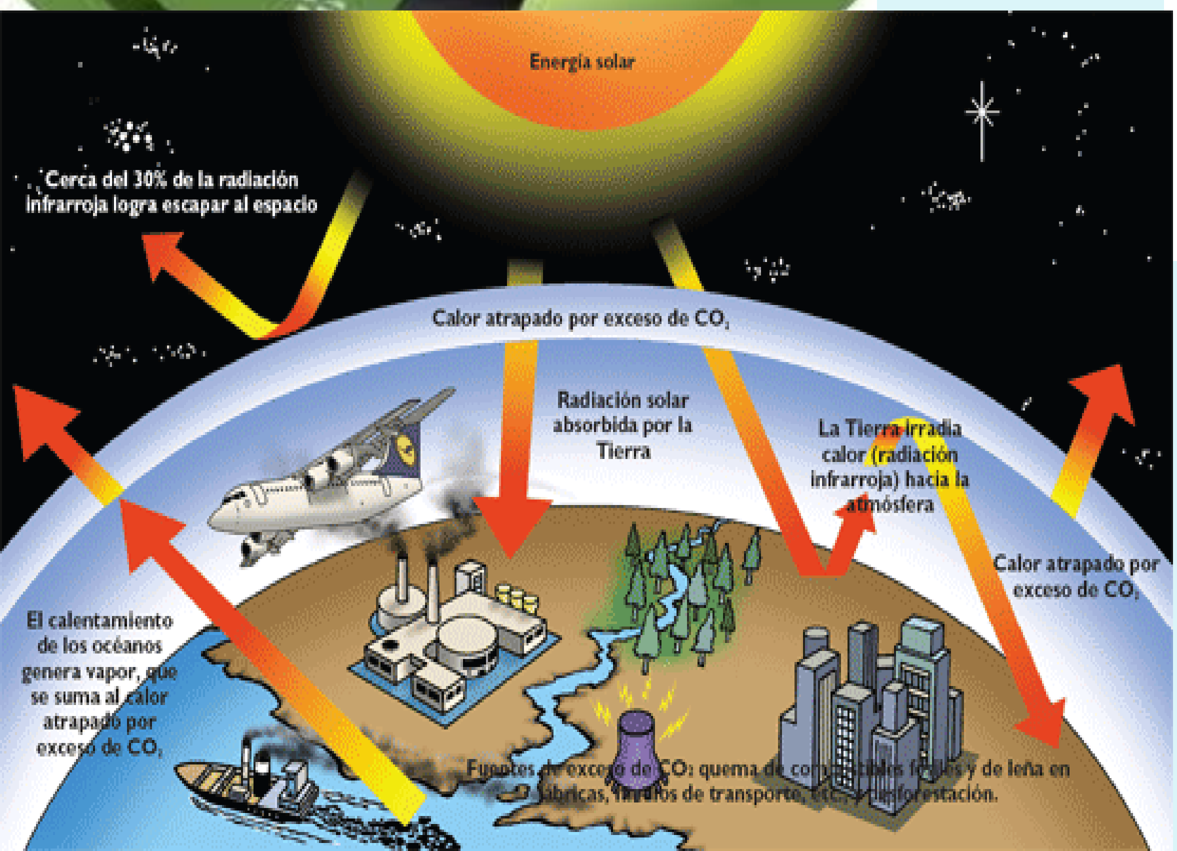
Climate change, global warming, air quality and dispersion of contaminants in air (WRF)




Serpent River
WEATHER STATION
IF THE ROCK IS WET... It's Raining
IF THE ROCK IS SWAYING... It's Windy
IF THE ROCK IS HOT... It's Sunny
IF THE ROCK IS COOL... It's Overcast
IF THE ROCK IS WHITE... It's Snowing
IF THE ROCK IS BLUE... It's Cold
IF THE ROCK IS GONE... TORNADO
Up to the Minute Weather Report prepared as a Courtesy by the Serpent River Indian Band
Have a nice Day!

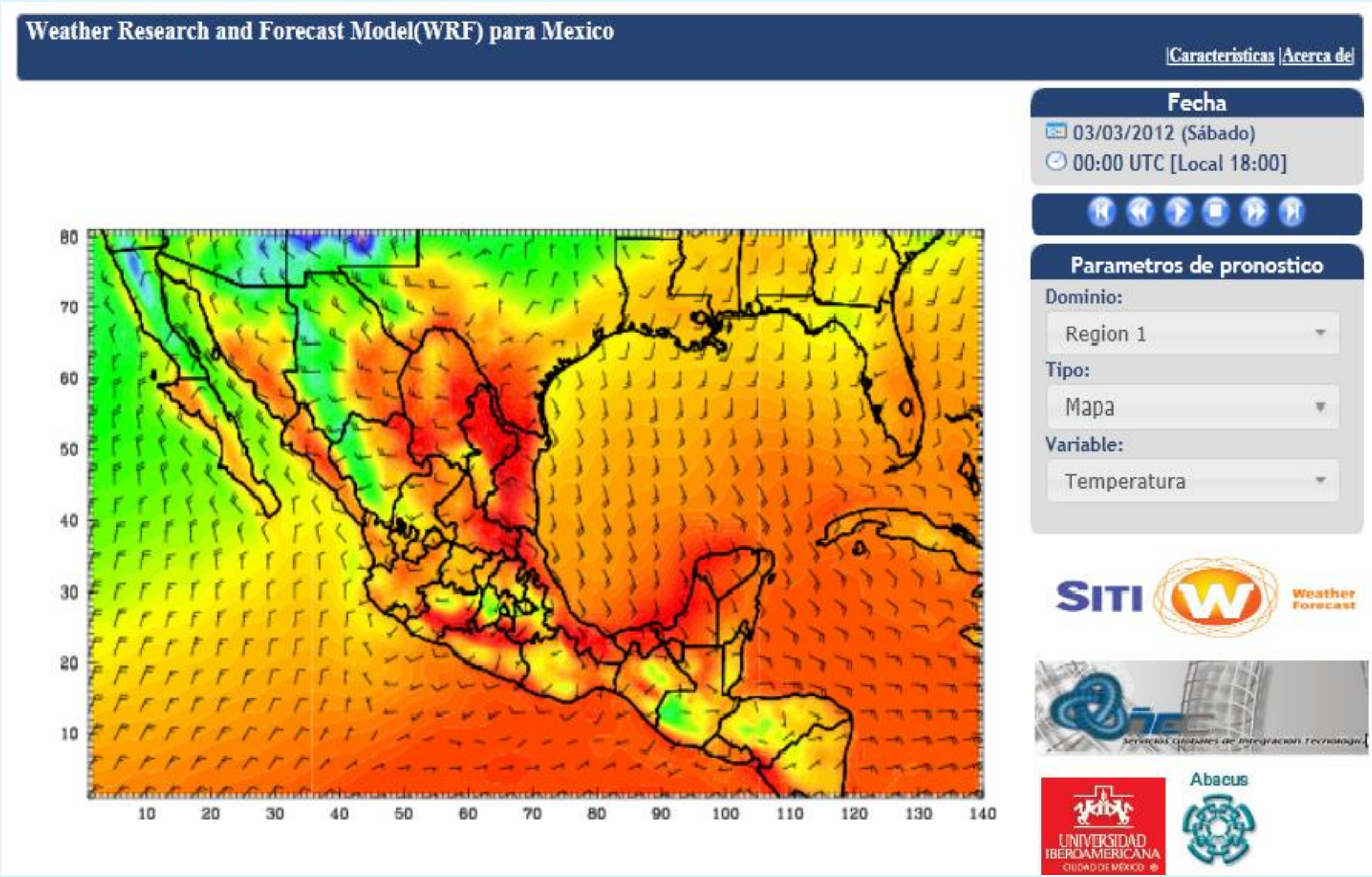






Atmospheric modelling is very complicate and computationally very expensive

Abacus forecast system



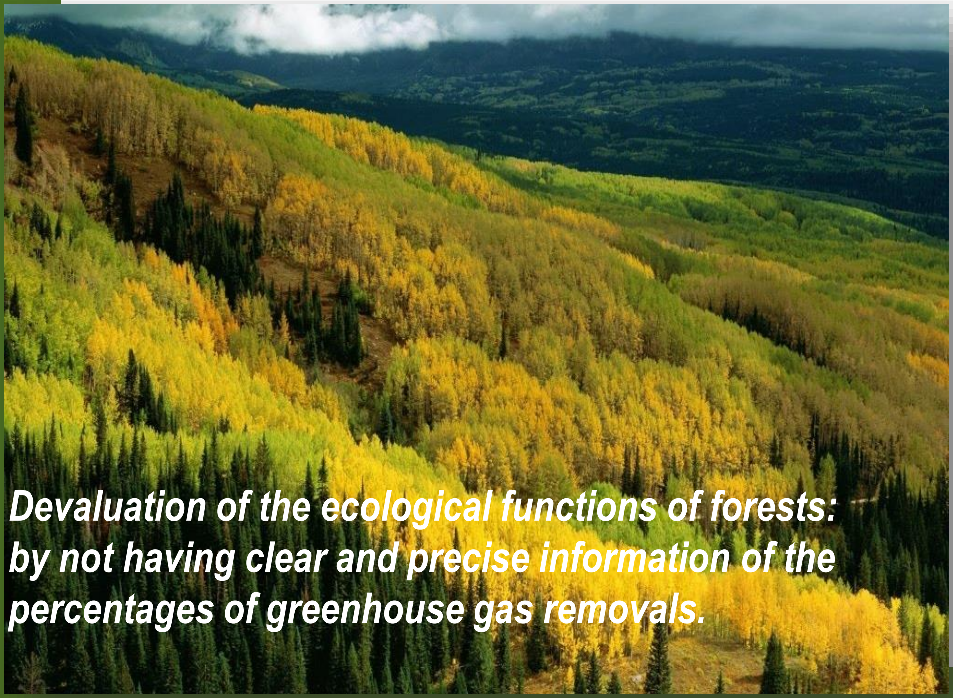
Intelligent cities

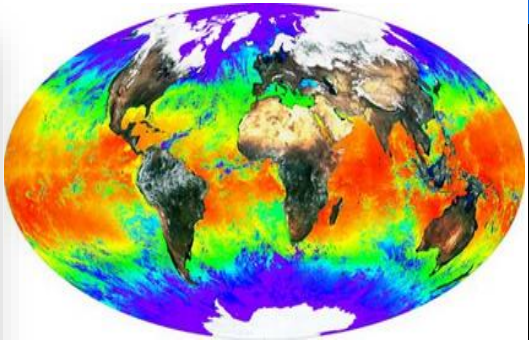
Watch the Documentary
Smarter Industries
IBM
Welcome to TheSmarterCity™
Transportation Airports & Rail
Public Safety
Social Services
Education
Healthcare
Economic Development
Transportation Traffic
Energy & Utilities

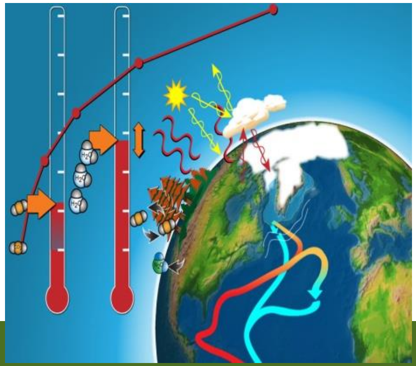
EDOMEX PROJECT: GLOBAL WARMING STUDY IN THE STATE OF MEXICO FOR UNDERSTANDING THE FOREST IMPACT ON GLOBAL WARMING AND TO HELP THE GOVERNMENT TO TAKE SPECIFIC MEASURES TO REDUCE GLOBAL WARMING

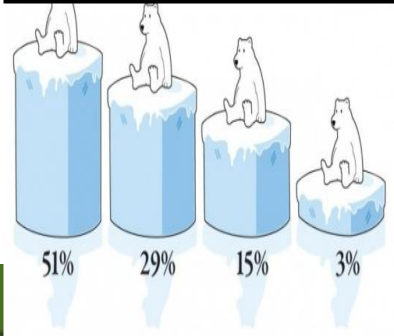
(PROJECT WITH CONAFOR: NATIONAL FOREST COMMISSION)

Forests play a role of great importance to mitigating climate change, since they have the ability to attach and absorb the carbon dioxide (CO_2), one of the main greenhouse gases, in a natural way and thereby regulate the climate.







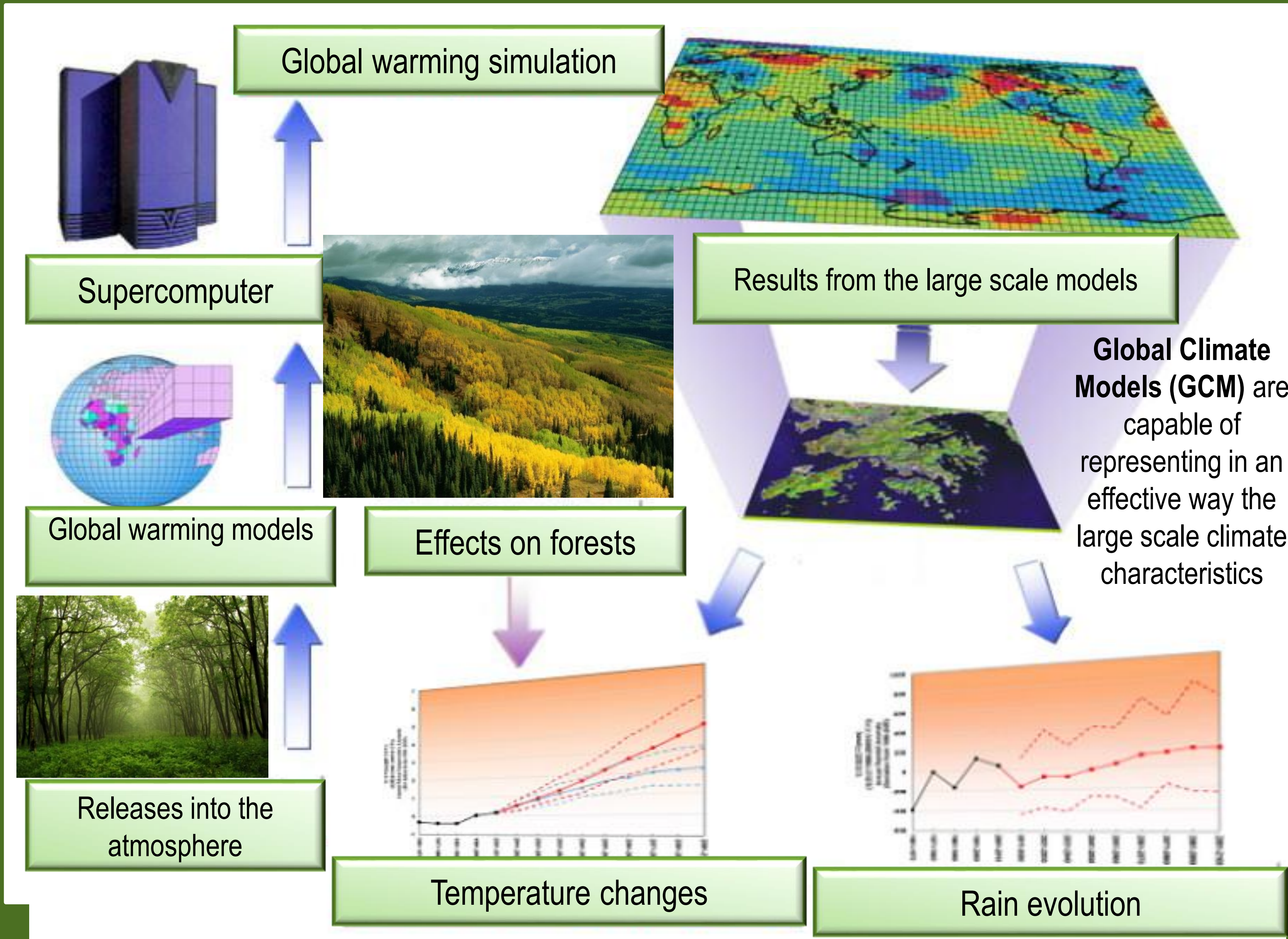


Preliminary proposal of domain definition over the state of Mexico for climate change simulations:

resolution 30 km, 360x160 cells

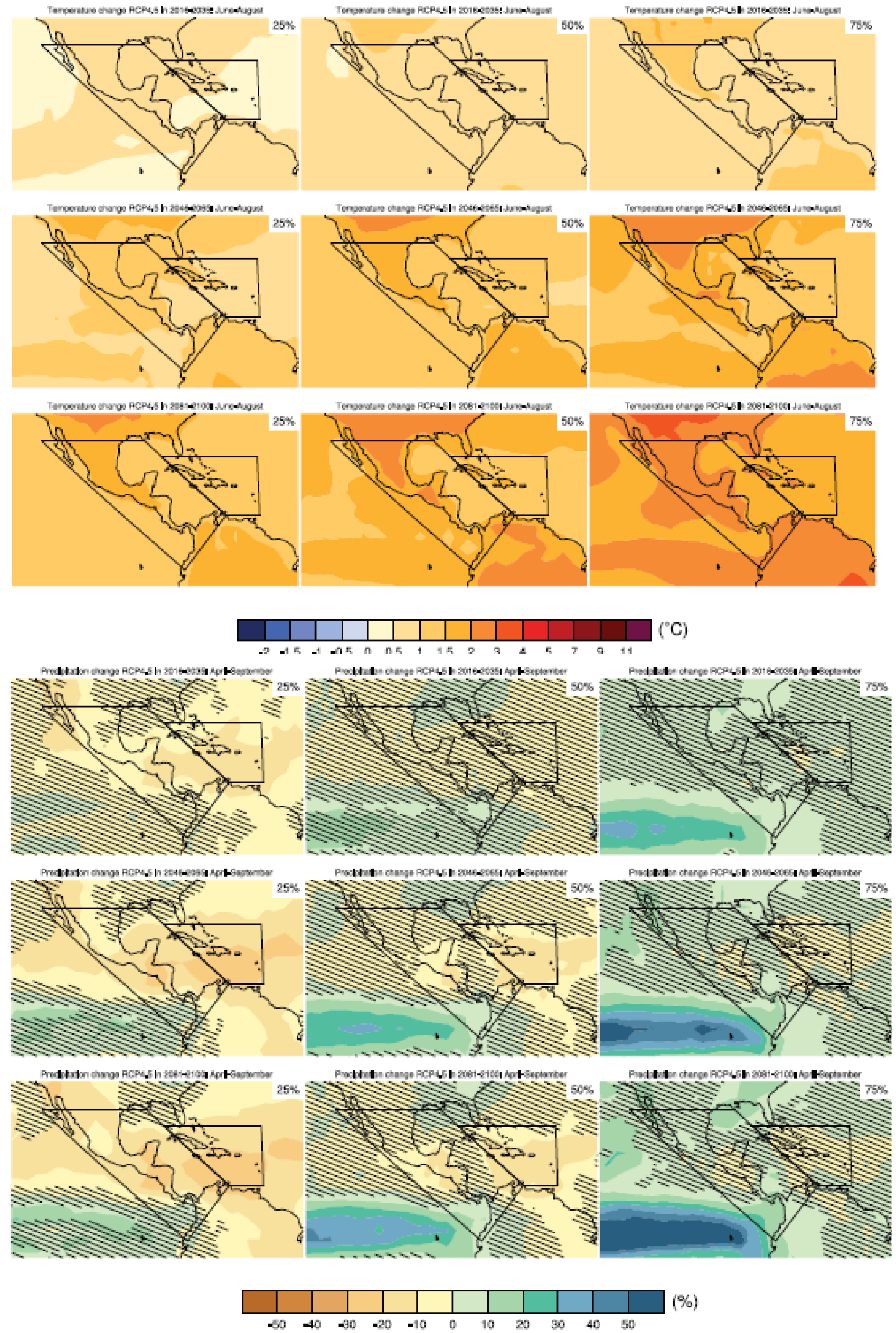
Domain 2: resolution 10 km, 226x148 cells . Both centered in 21ºN,-99ºW

Simulation of the control period using GCM data



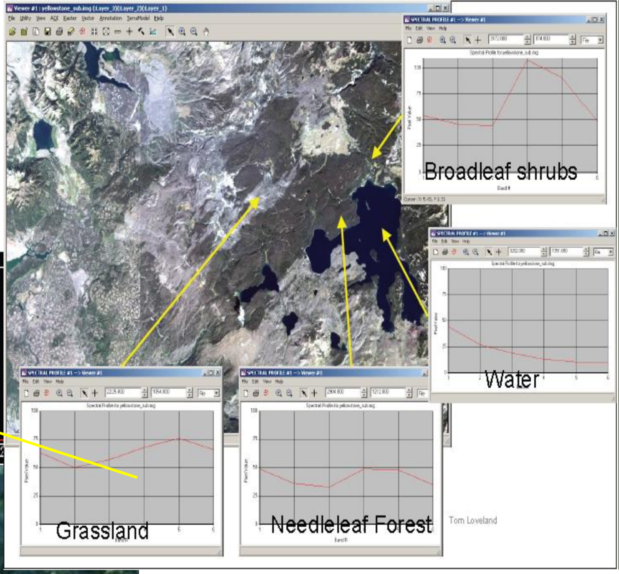
Projections of regional climate for Central America (IPCC, 2013)

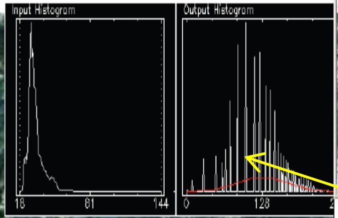
IPCC Intergovernmental Panel on Climate Change

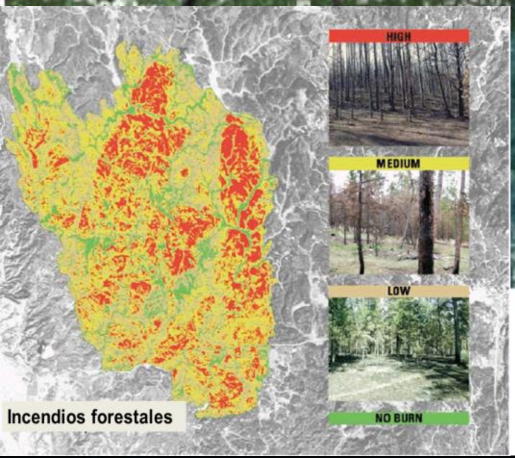


Projected changes in temperature (June to August) and precipitation (April to September) in different future periods (2016-2035 and 2046-2065, 2081-2100) regarding the current conditions (1986-2005), according to forecasts by the RCP4.5

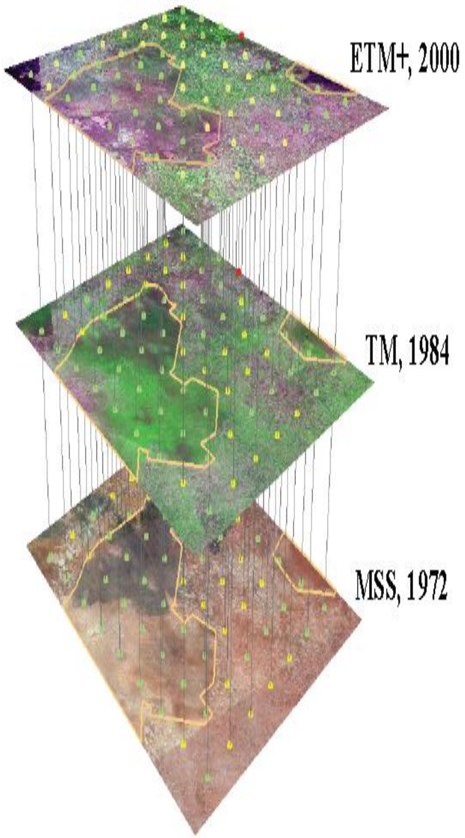
Vegetation cover change detection

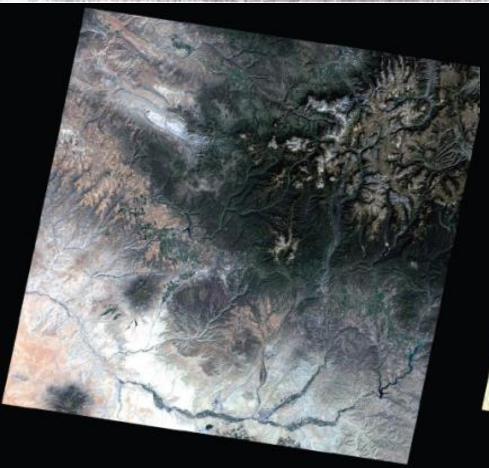


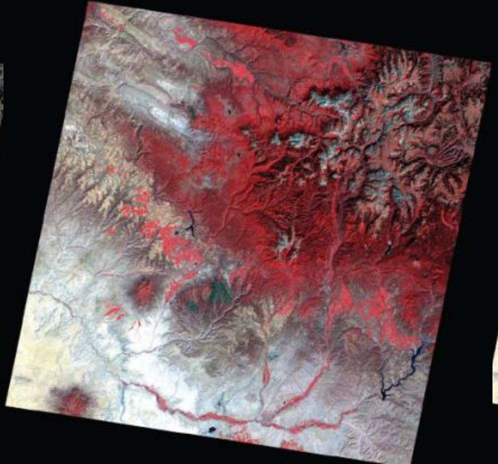


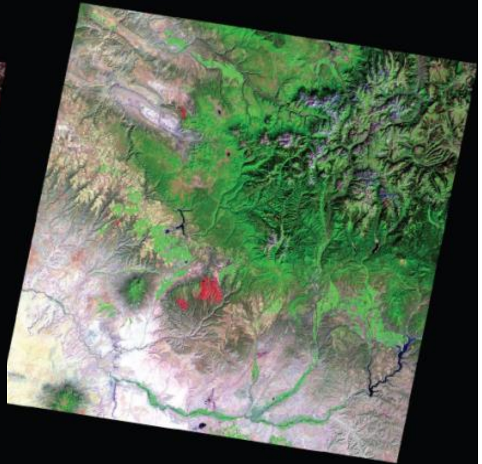


ERDAS, ENVI, IDRISI, ILWIS, MULTISPEC, GVSIG, ARCGIS





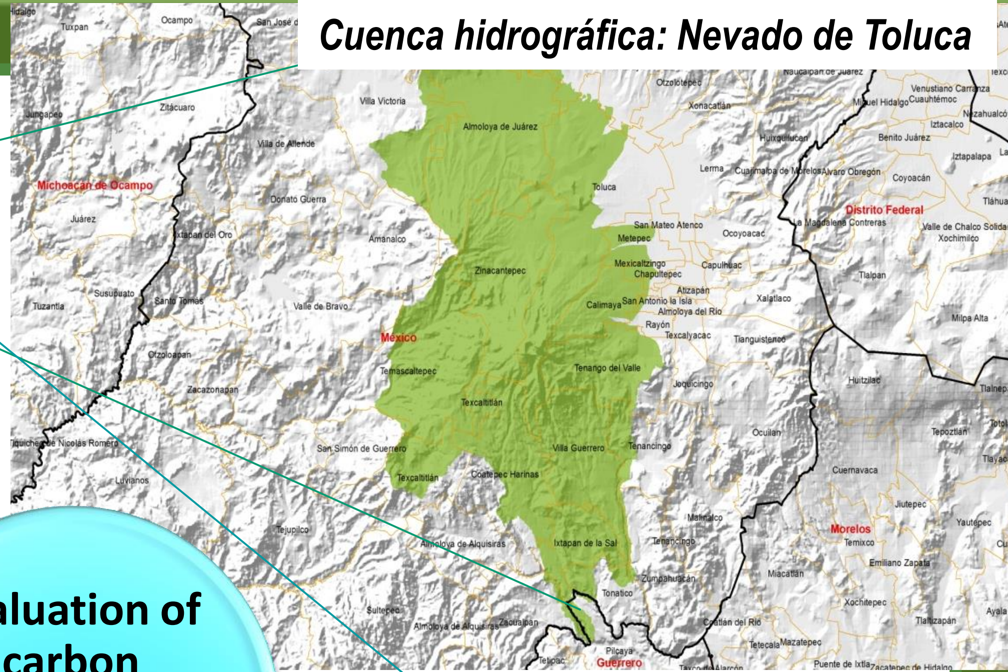




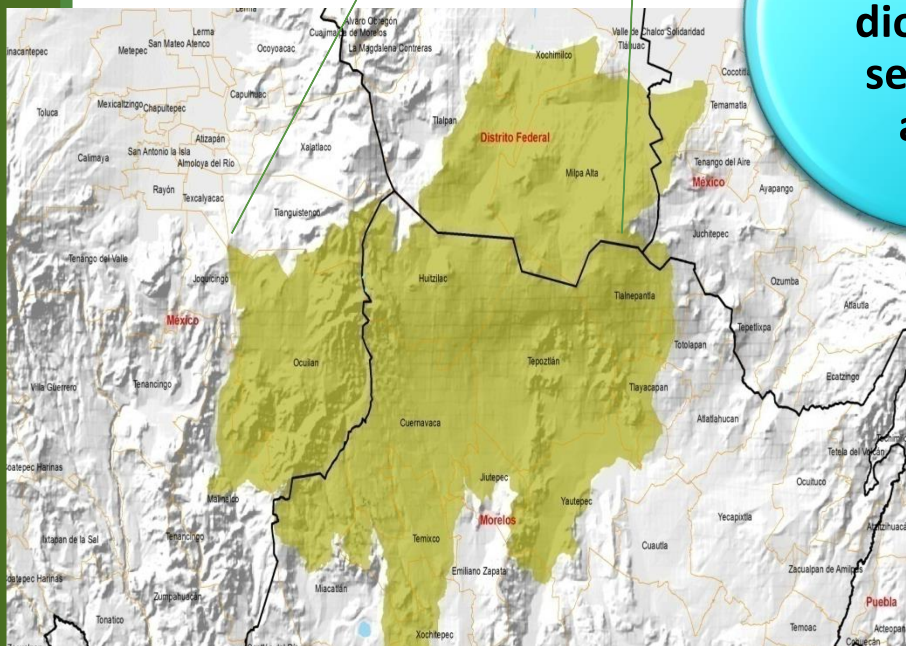
Stage 4

CONAFOR

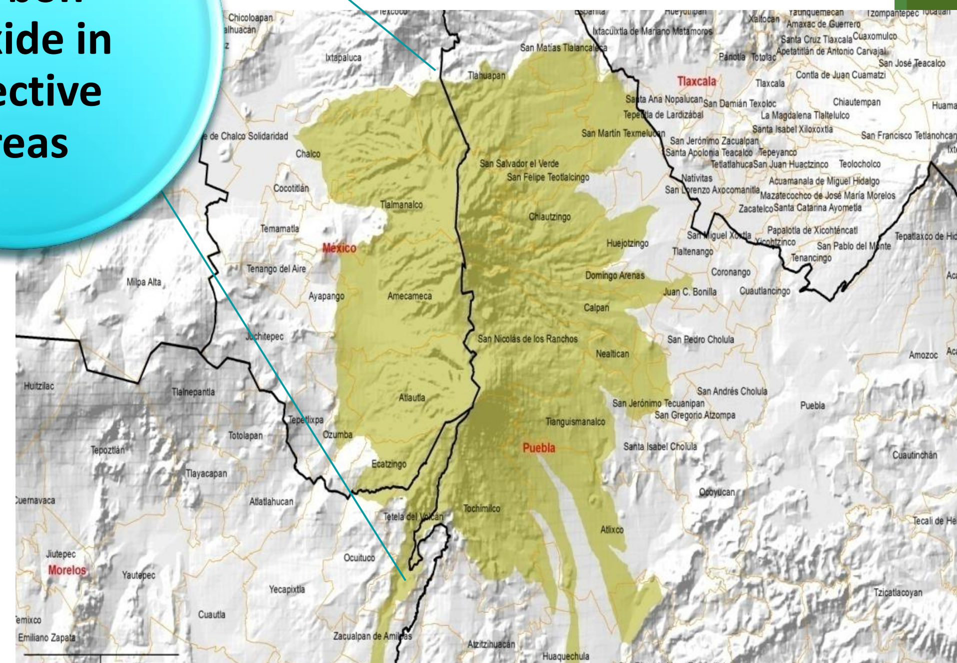
Cuenca hidrográfica: Nevado de Toluca



Evaluation of carbon dioxide in selective areas

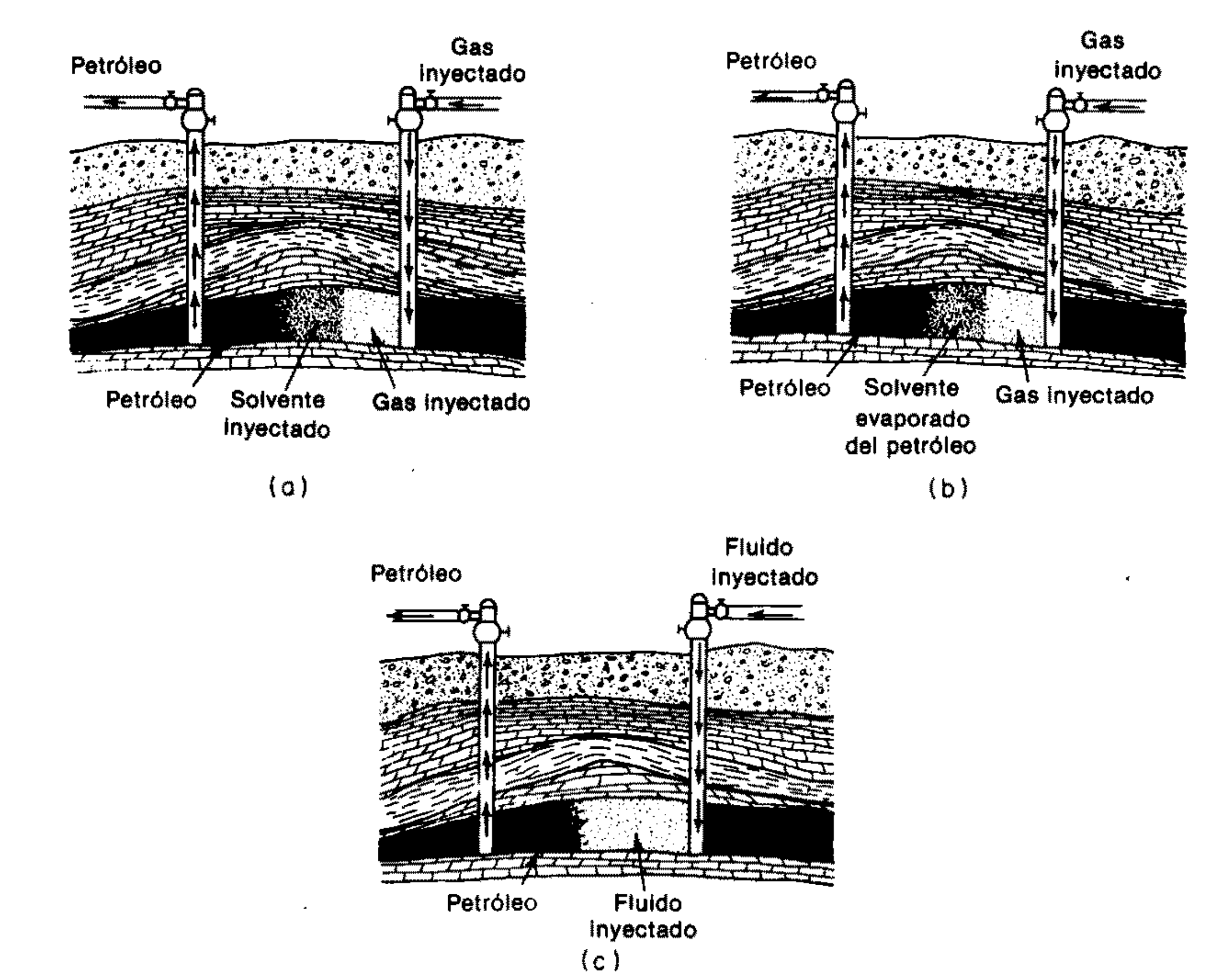


Cuenca hidrográfica: Chichinautzin

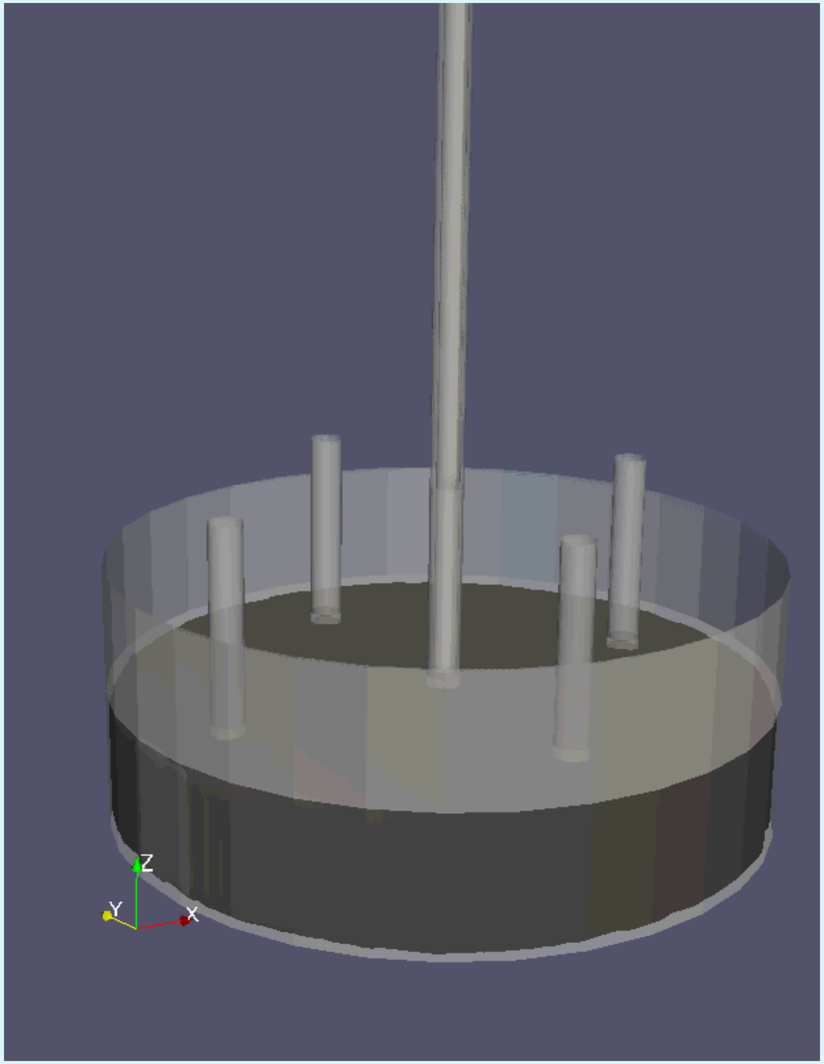


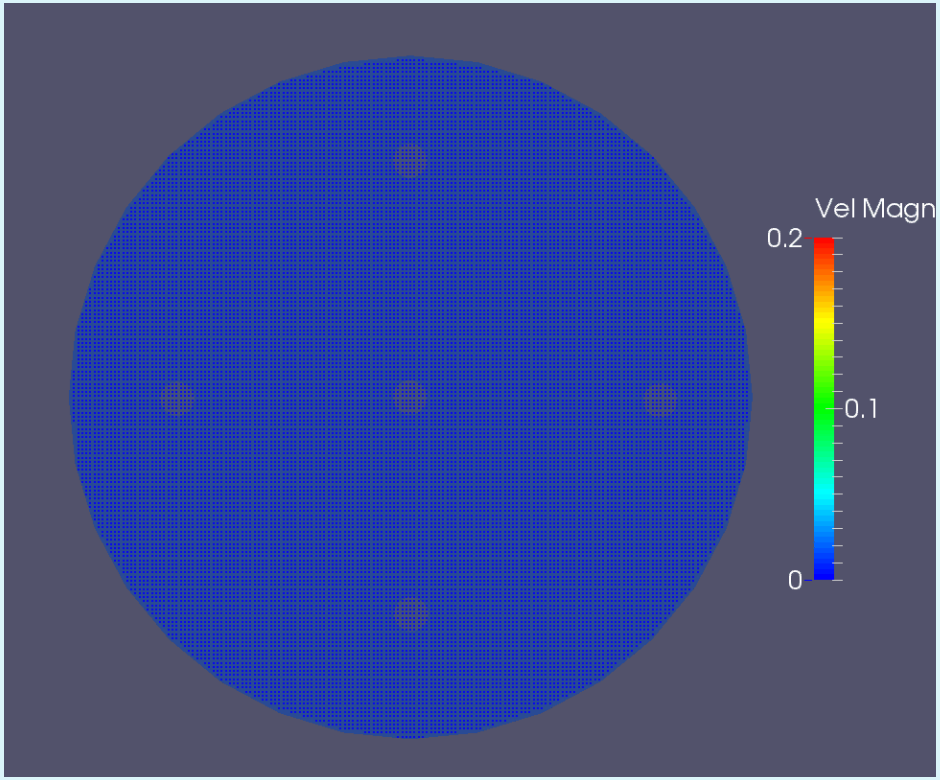
Cuenca hidrográfica: Iztaccíhuatl–Popocatépetl

Simulator of naturally fractured reservoirs

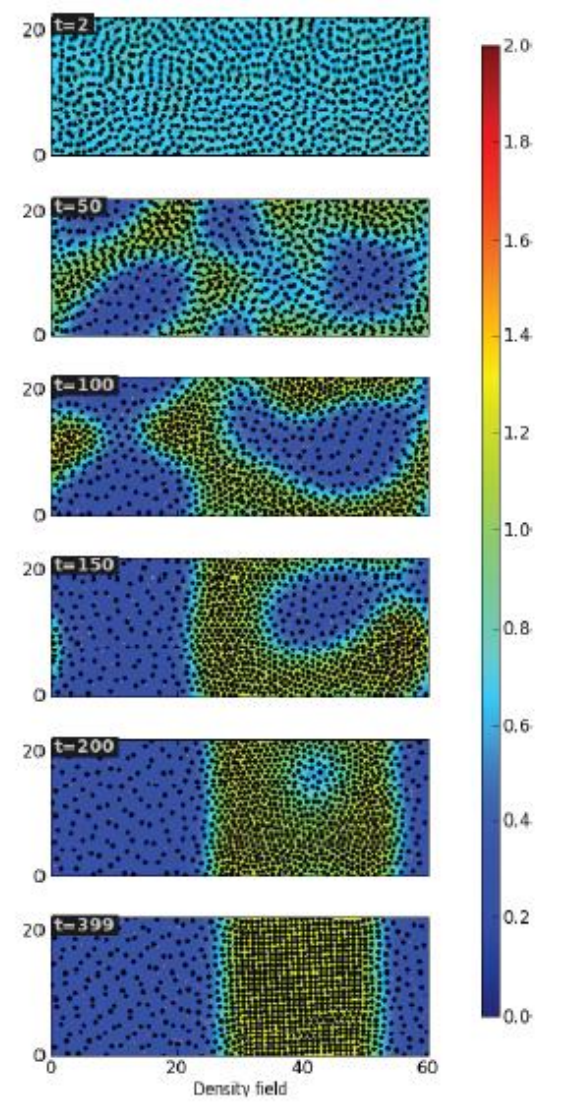


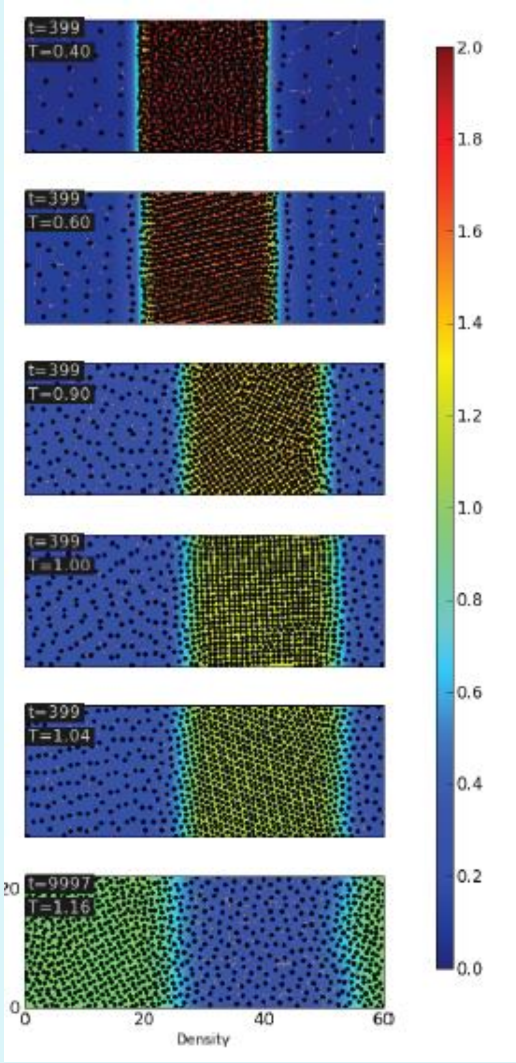
Injection of gas and liquids for oil extraction

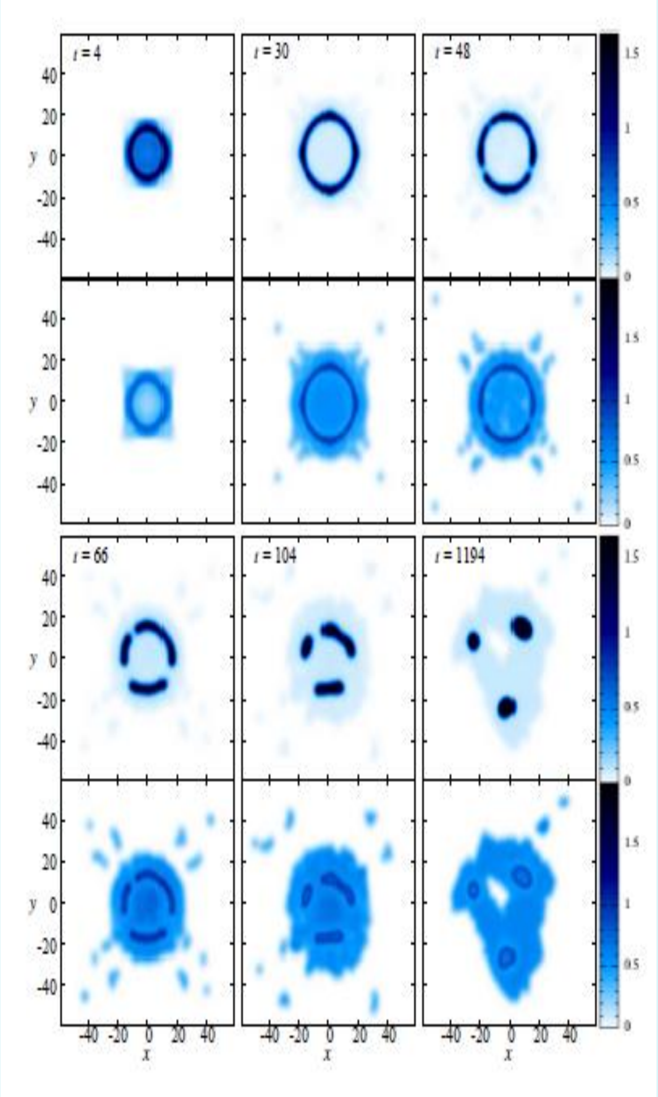


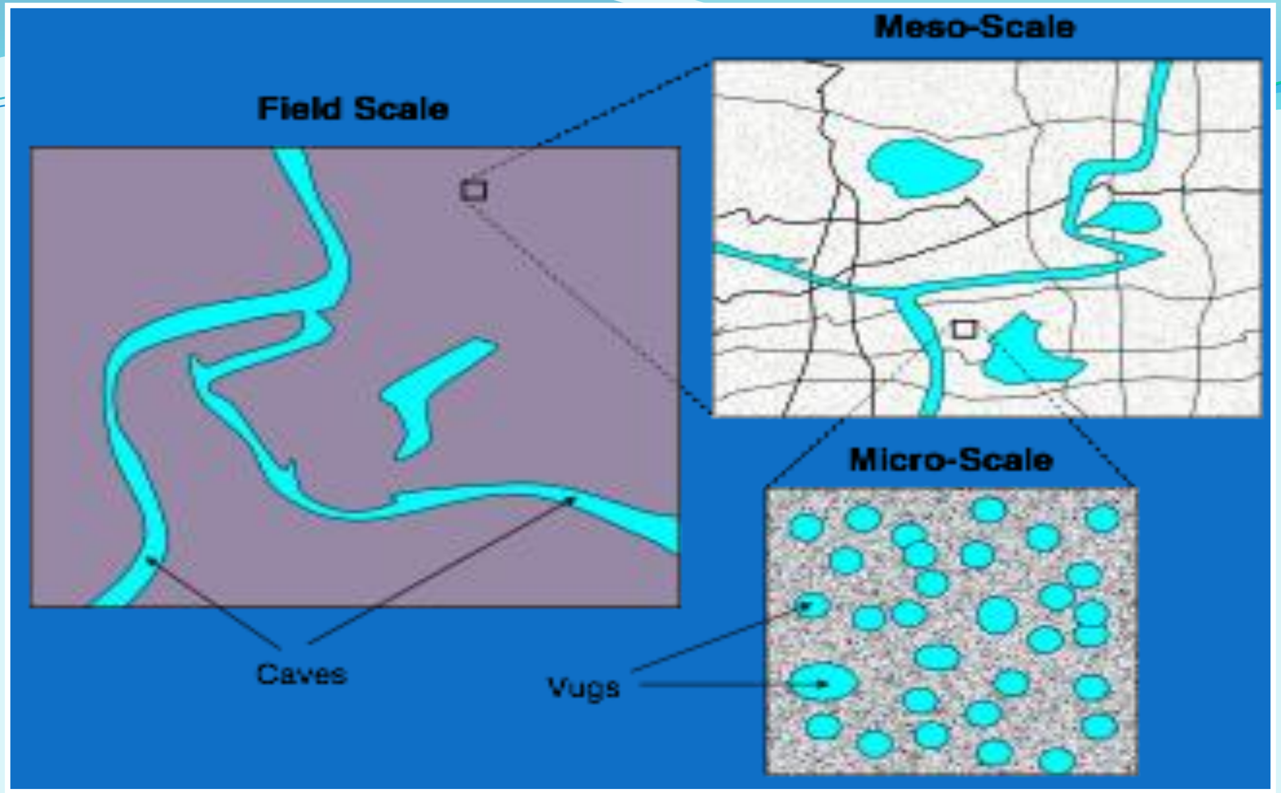


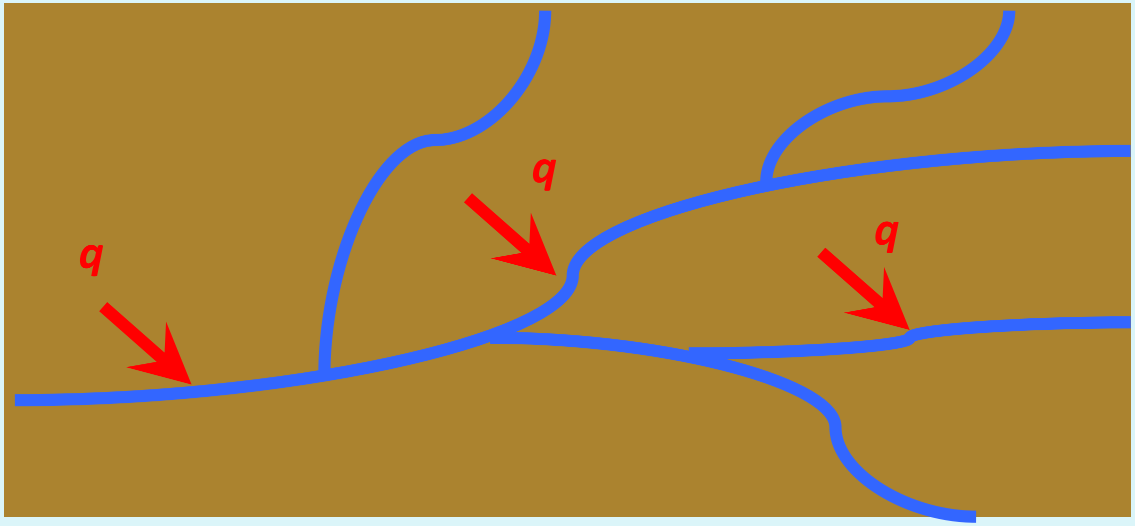
Modeling of 2-phase coexistence, surface tension, capillarity and porosity.








Meso-Scale
Field Scale
Micro-Scale
Caves
Vugs


q
q
q

Conclusions

We have presented some of our Computational Fluid Dynamics Projects.

The Abacus Project will produce a large amount of data, perhaps 1 PB per year.

We have to solve the problem of moving the data to the final user.

I have only described the Computational Fluid Dynamic applications in Abacus, we have many more applications: biology, chemistry, energy, medicine, engineering, economy, health, etc.