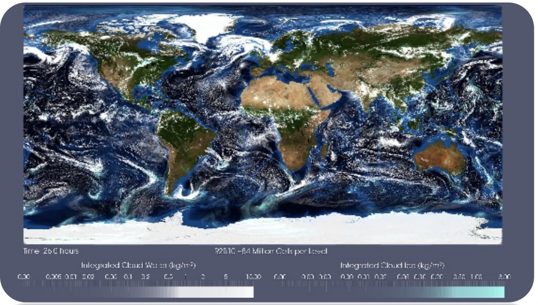
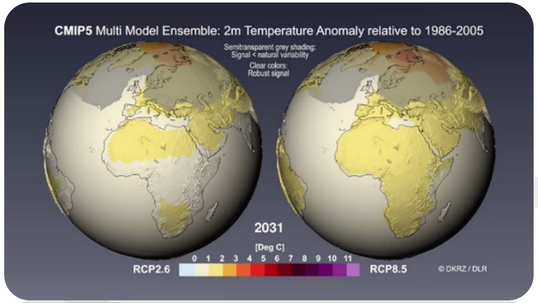
Visualization challenges imposed by HPC driven Earth system research

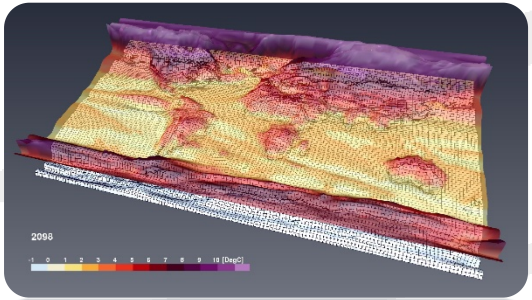
Michael Böttinger

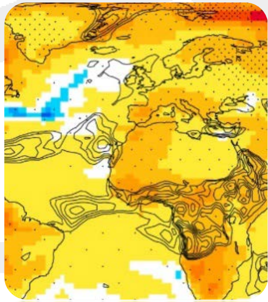
Niklas Röber, Karin Meier-Fleischer, Dela Spickermann

German Climate Computing Center (DKRZ)









Data Production at DKRZ

Deutsches Klimarechenzentrum (DKRZ) = German National HPC Facility for Earth System Research

Supercomputer &
Disk Storage

3.6 PFlops
266 TByte Mem.
Top500: #55 (06/2018)

54 PByte
Lustre File System

HPSS Archive System

8 + 1 Silos
70 Drives
up to 500 PByte

Visualization Server

- Integral part of HLRE-3 supercomputer
- FDR Infiniband Interconnect

- 12 Nodes Atos/Bull
  - 2 CPUs, 12 Core Haswell
  - 256 GB Main Memory
  - NVidia Kepler GPUs: Tesla K80
- 9 Nodes Atos/Bull
  - 2 CPUs, 18 Core Broadwell
  - 512 / 1024 GB Main Memory
  - NVidia Maxwell GPUs: M40 / M6000
- Remote 3 D-Rendering: TurboVNC/VirtualGL

Visualization Tools employed in Climate Sciences

Type	Name	URL	Properties	
Domain-specific	**NCL**	http://www.ncl.ucar.edu/	2D script-based	free
	IDV	http://www.unidata.ucar.edu/software/idv/	2D/3D interactive GUI	free
	Vapor	https://www.vapor.ucar.edu/	3D interactive GUI	free
	UV-CDAT	http://uvcdat.llnl.gov/	Collection: 2D /3D tools	free
	GrADS	http://cola.gmu.edu/grads/	2D script-based	free
	Ferret	http://www.ferret.noaa.gov/Ferret/	2D script-based	free
	GMT	http://gmt.soest.hawaii.edu/	2D script-based	free
General-purpose	**ParaView**	http://www.paraview.org/	3D interactive GUI	free
	Avizo	https://www.fei.com/software/avizo3d/	3D interactive GUI	$$ †
	IDL	http://www.harrisgeospatial.com/	2D script-based	$$
	Python / matplotlib	http://matplotlib.org/	2D script-based	free

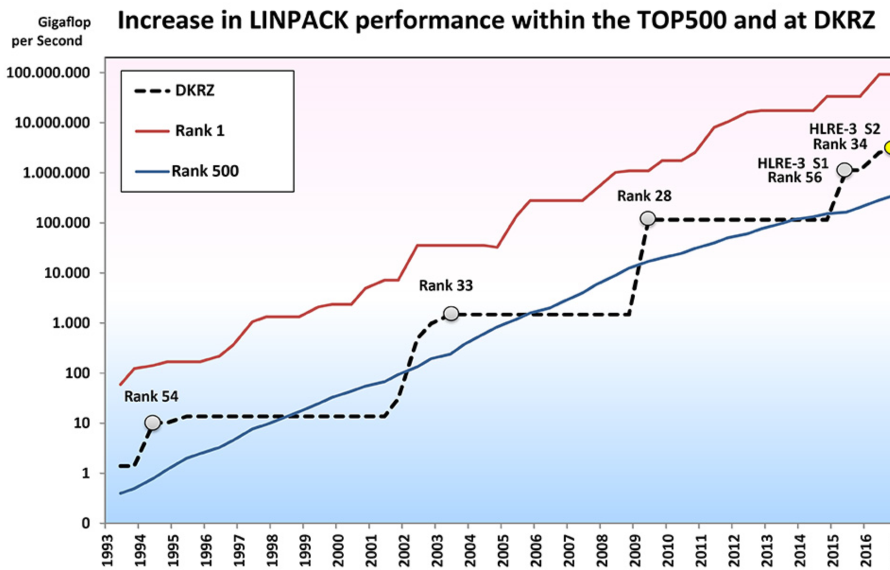
Recommended Literature:
Rautenhaus et al. (2017): Visualization in Meteorology --- A Survey of Techniques and Tools for Data Analysis Tasks, IEEE Transactions on Visualization and Computer Graphics, DOI: 10.1109/TVCG.2017.2779501

Scientific Use of Performance Increases

1 ▪ Higher spatial / temporal resolution of climate models
  - Direct simulation instead of parameterization
- Increased model complexity (more processes / variables)

2 ▪ Ensemble simulations -> uncertainty
- Very long simulations
  - Ice age cycles (110,000 years)

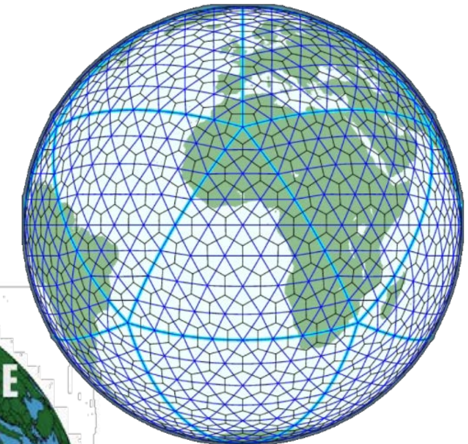
-> Challenges for visualization



Challenge 1: High Resolution

ICON, HD(CP)² and ESiWACE

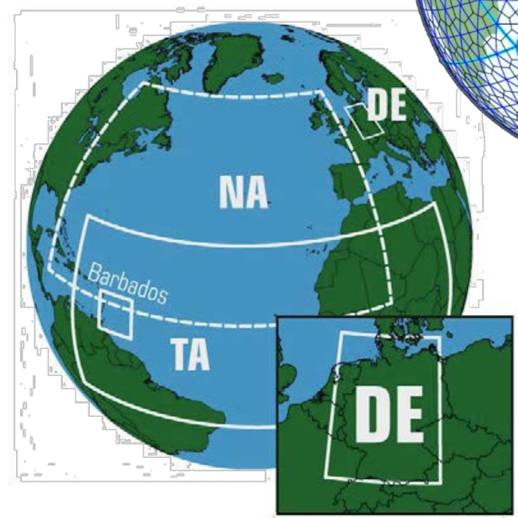
- **ICON - ICO**sahedral **N**on-hydrostatic grid
  - Joint Development by MPI-M and DWD
  - Unstructured Grid
  - ParaView U-Grid NetCDF-Reader (N. Röber)



- Global Setup
  - Ocean – 5 km resolution

    (15 million cells * 64 levels)
  - Atmosphere – 2.5 km resolution

    (84 million cells * 138 levels)



- Regional Setup (ICON-LES)
  - HD(CP)² – 120 m resolution

    (22.5 million cells * 150 levels)

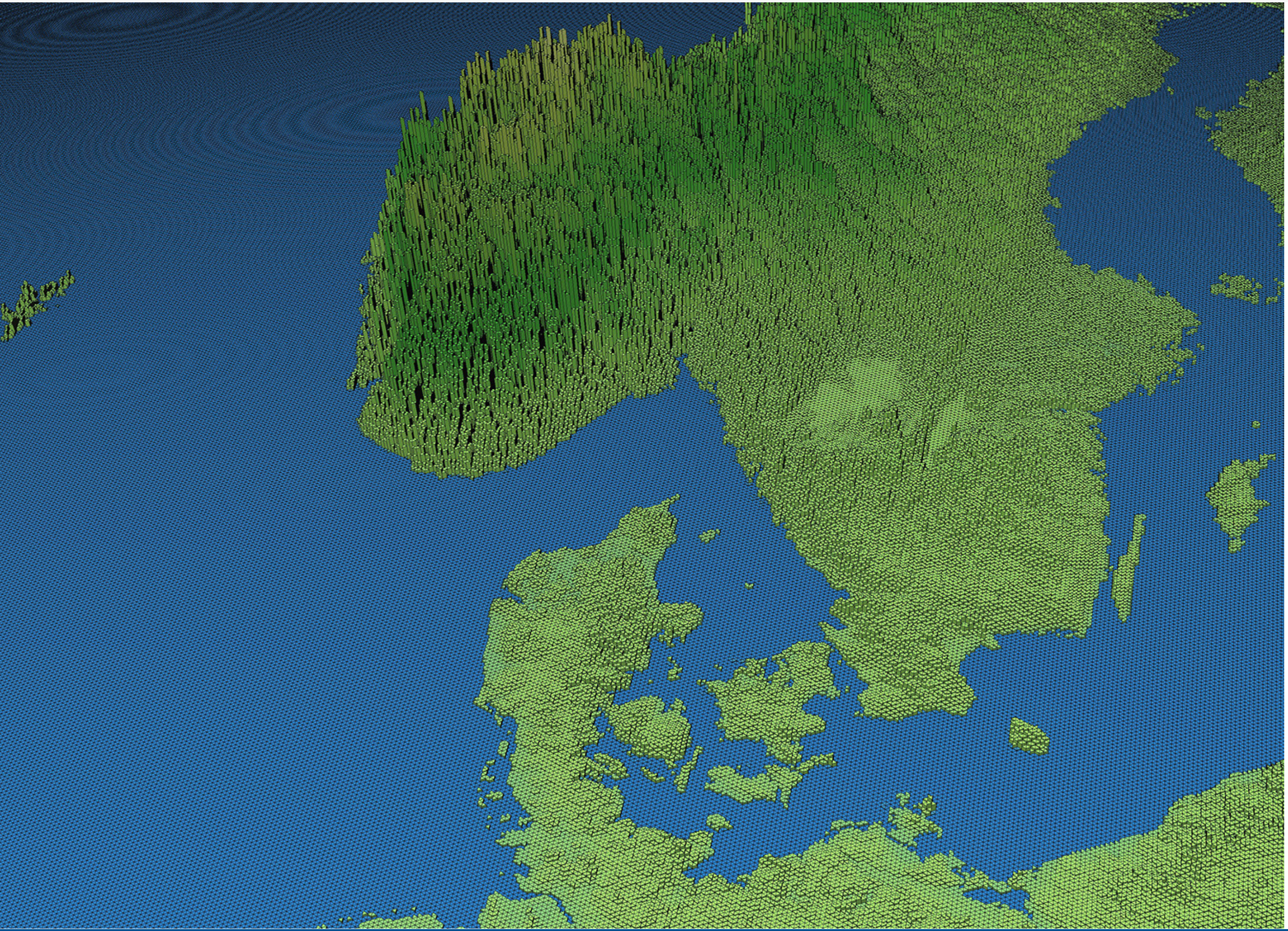
esiwace CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

HD(CP)² High definition clouds and precipitation for advancing climate prediction

Grid Resolution – Global Model

ICON R2B9
5 km
21 Mio Cells / Level
Several weeks

Grid Resolution – Global Model



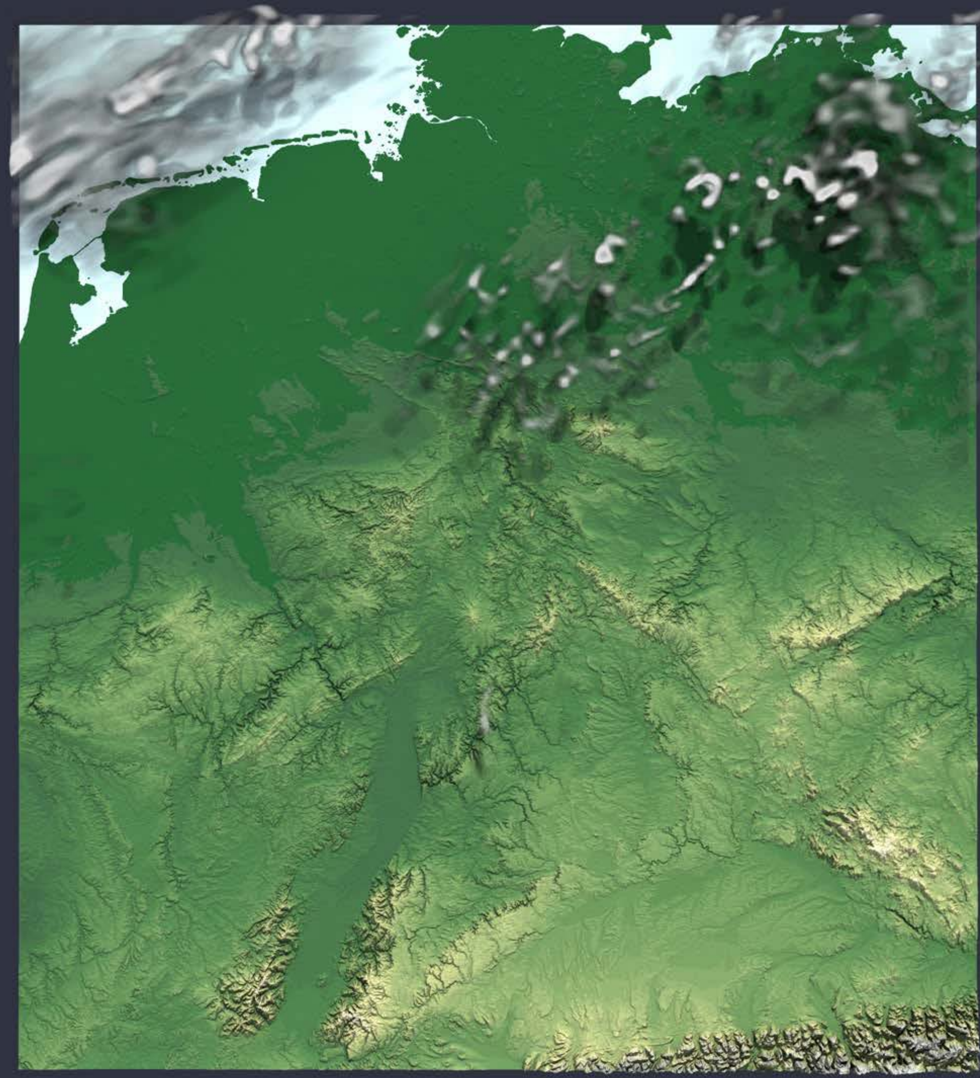
ICON R2B10
2.5 km
84 Mio Cells / Level
Several Days

ICON-LES - DE Domain

Cloud Water

Δt = 1 min (storage)

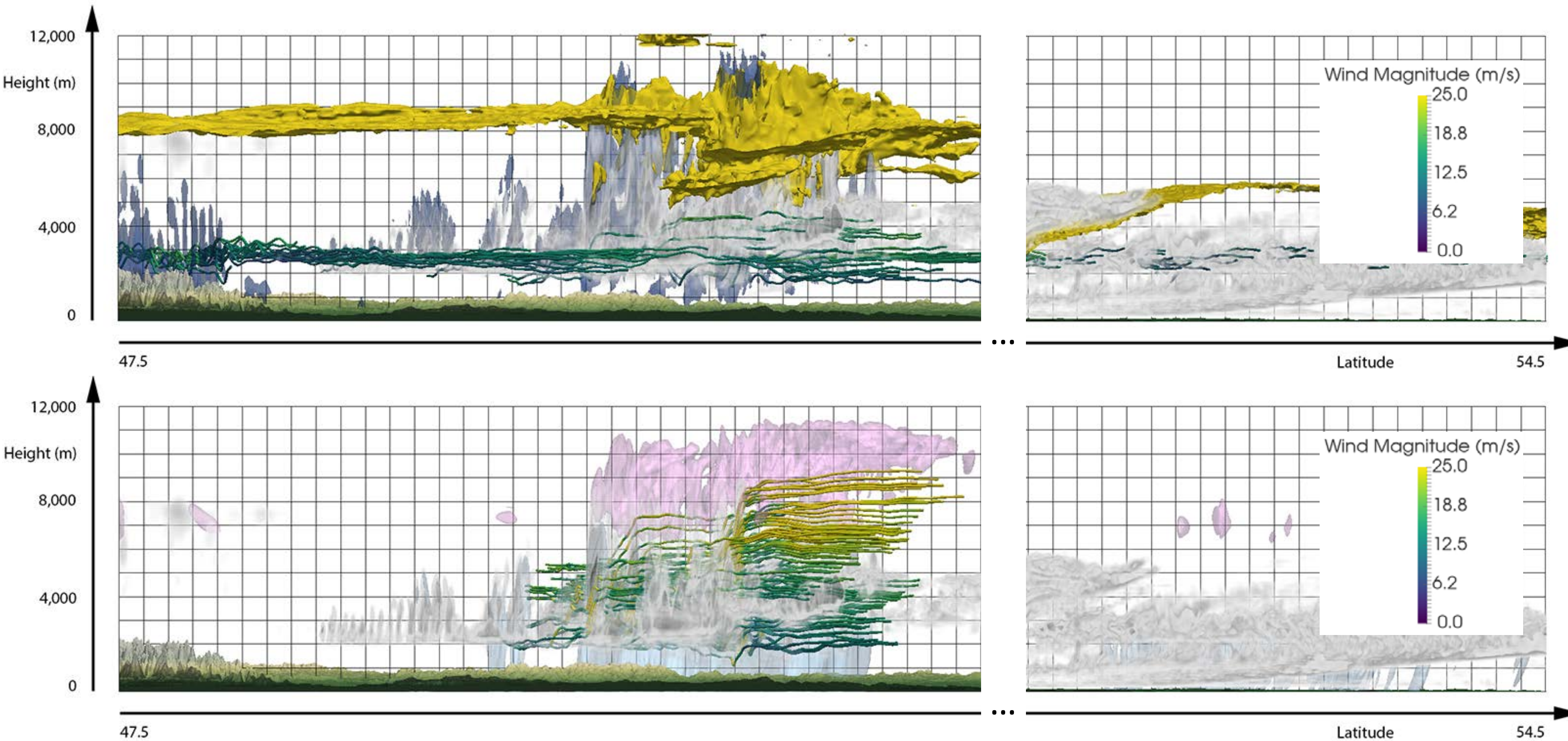
Δx = 312 m



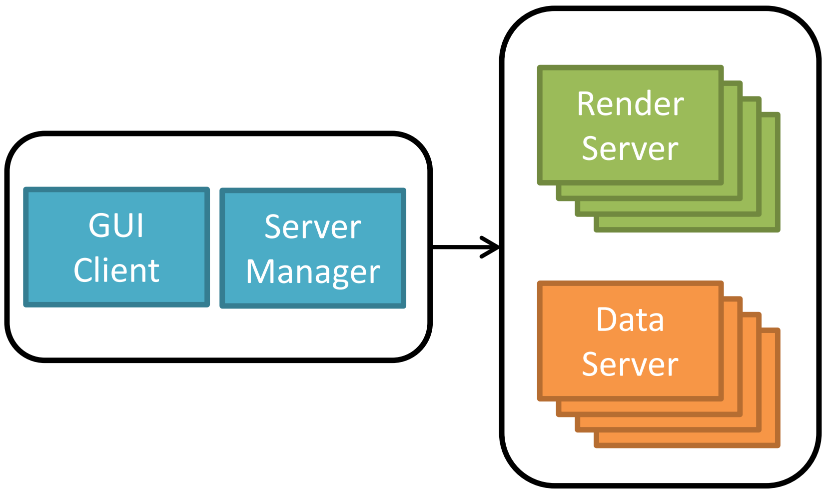
26.April 2013 0 Uhr

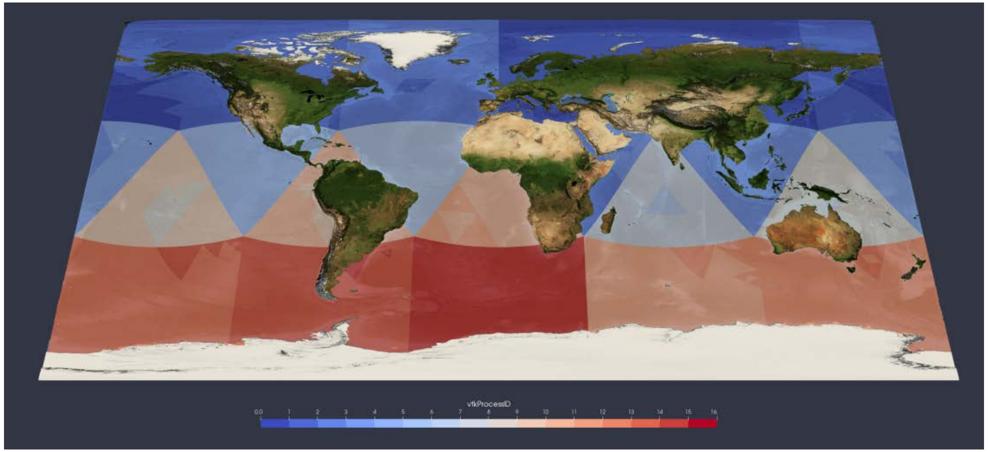
DKRZ

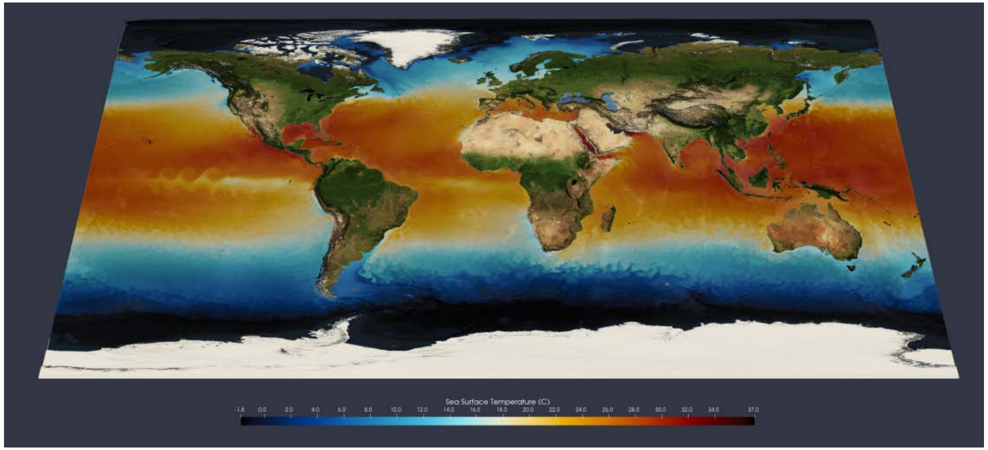
HD(CP)² Data



Parallel Processing and Visualization with ParaView

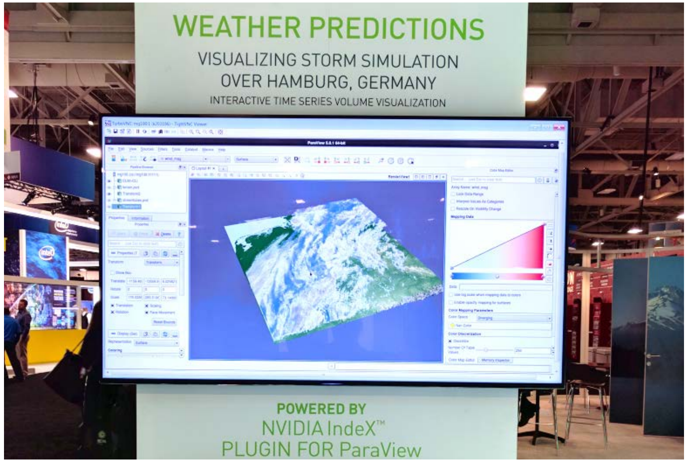




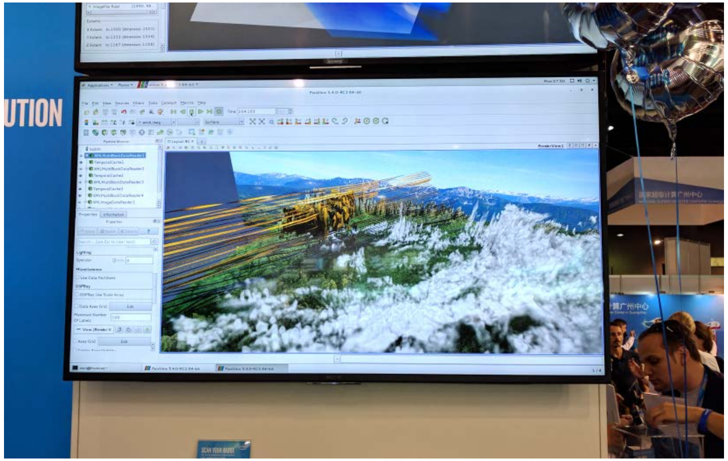


NVIDIA IndeX and INTEL OSPRay

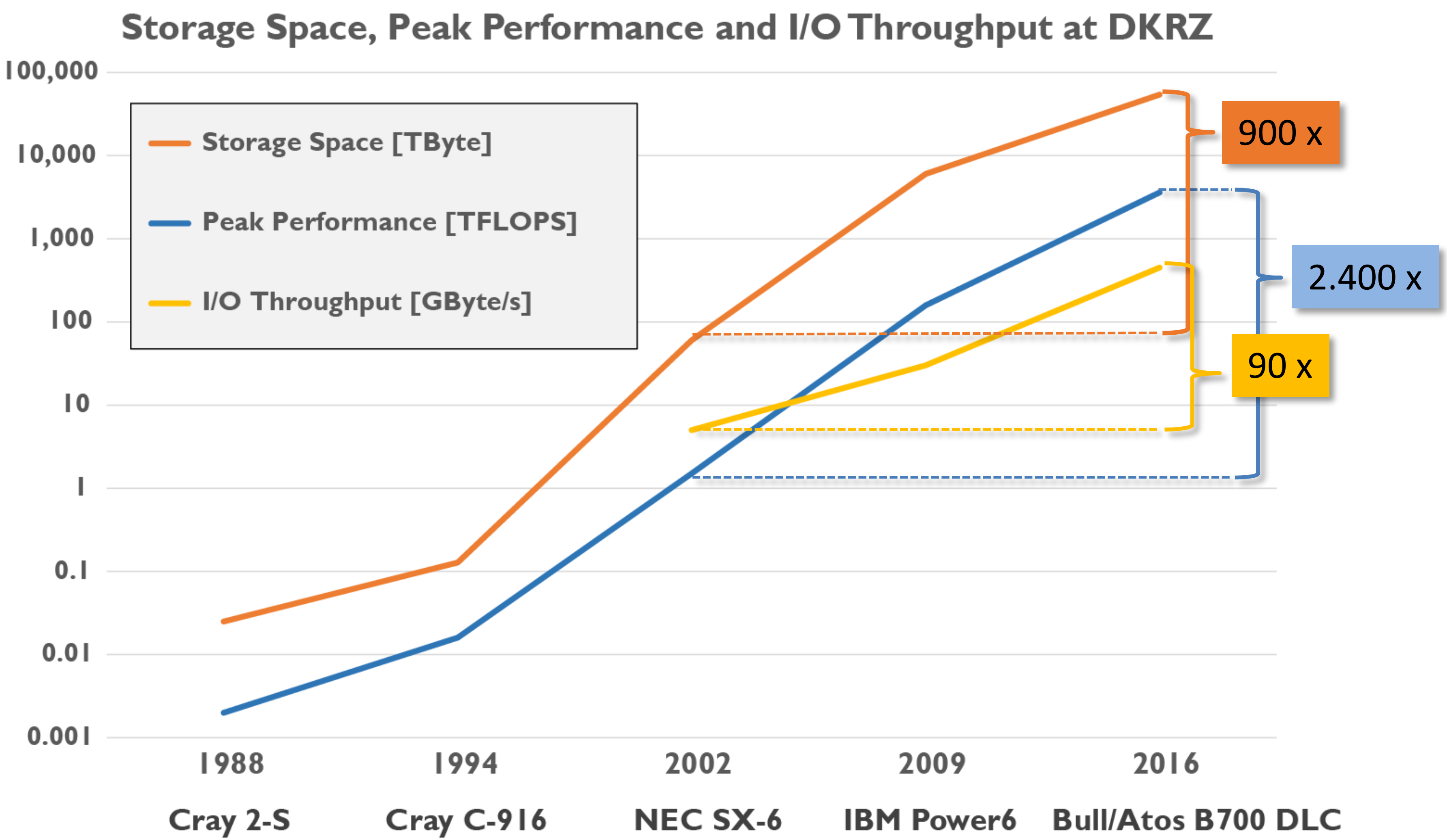
- Two different approaches for (volume) rendering of large data
  - Using GPUs and texture based volume rendering
  - Using CPUs and software based ray-tracing
- We (of course) collaborate with both ☺



SC 16



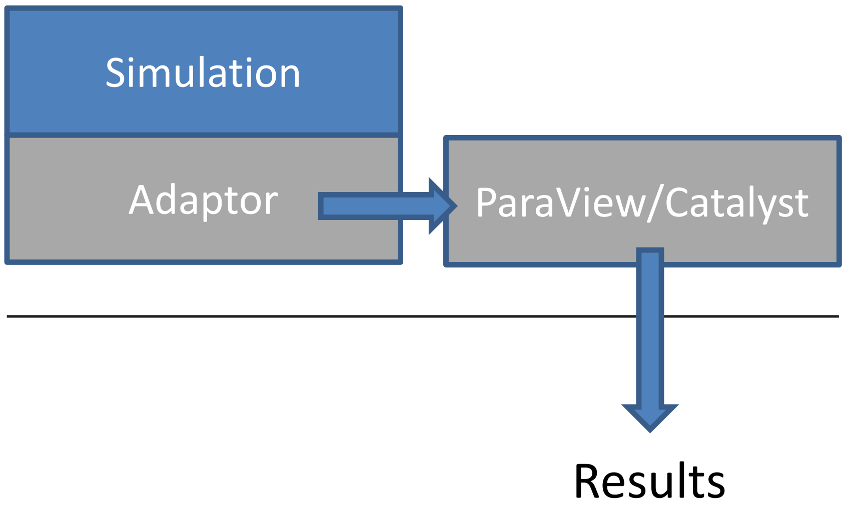
ISC 17


Storage Space, Peak Performance and I/O Throughput at DKRZ
Storage Space [TByte]
Peak Performance [TFLOPS]
I/O Throughput [GByte/s]
100,000
10,000
1,000
100
10
1
0.1
0.01
0.001
900 x
2.400 x
90 x
1988
1994
2002
2009
2016
Cray 2-S
Cray C-916
NEC SX-6
IBM Power6
Bull/Atos B700 DLC

Two Options

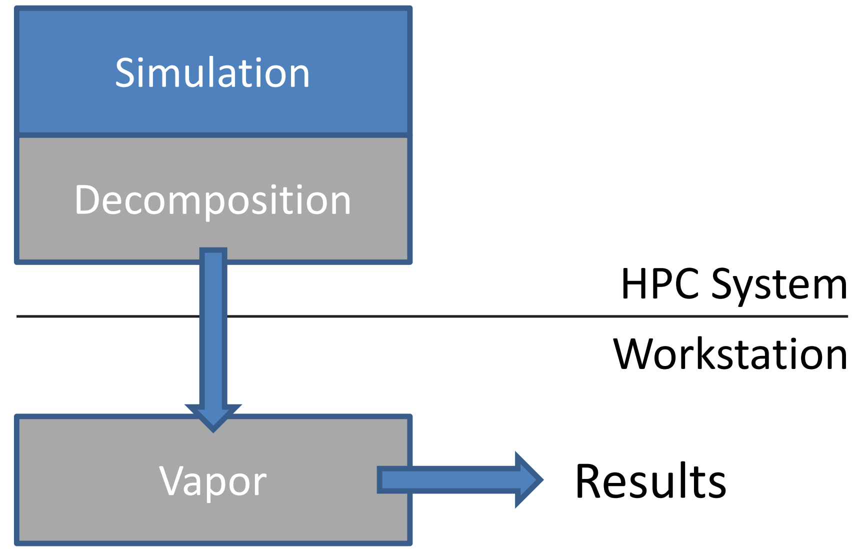
in-situ Visualization

(ParaView/Catalyst/Cinema)

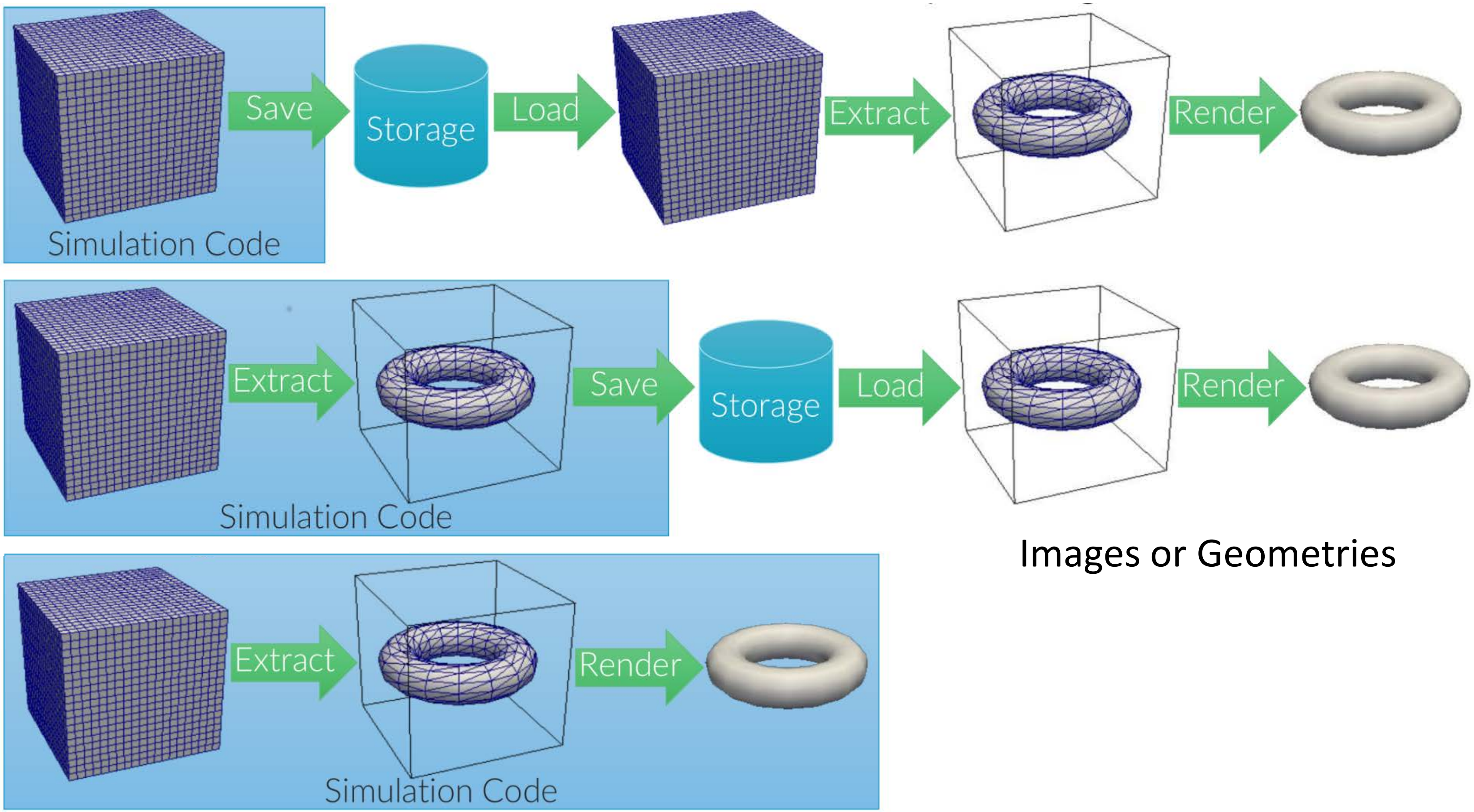


in-situ Compression

(Vapor)



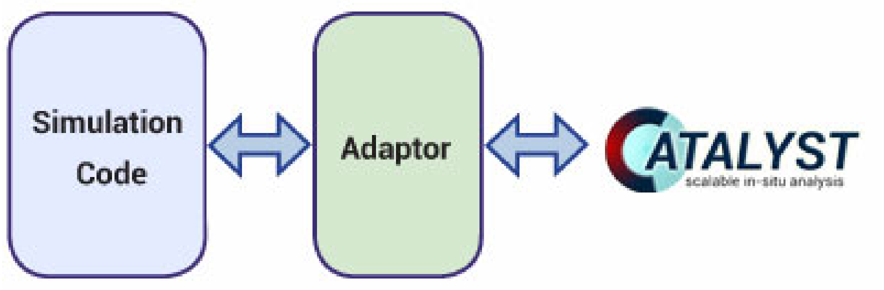
From Post Visualization to In-Situ



Images or Geometries

In-situ Visualization with ParaView

- In-situ (co)processing/visualization using Paraview/Catalyst
- Adaptor required that connects ICON (model) and Catalyst
- Two possibilities for co-processing:
  - Batch-visualization using pre-defined Python scripts
  - Live visualization within a client/server setting



- First successful tests with prototype for ICON

In-Situ Visualization - Challenges

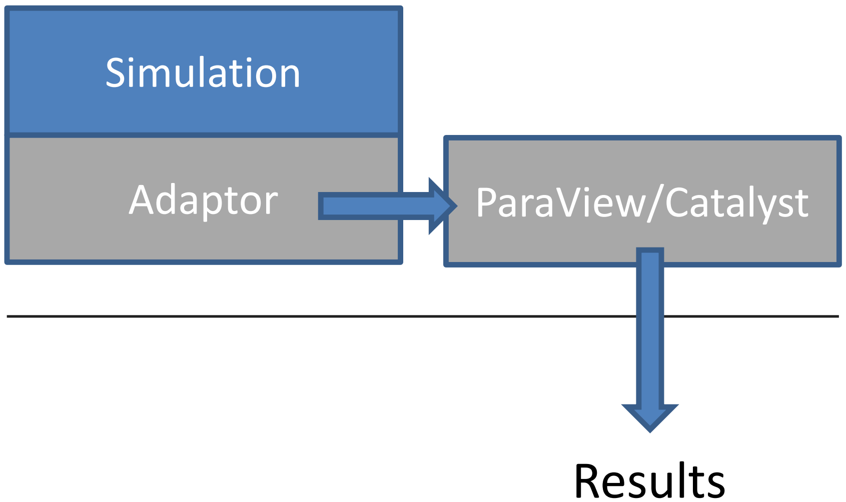
- Paradigm change in scientific and data workflow
- Research/analysis questions need to be better known ahead of the simulation
- Climate = statistics of weather. How to derive and visualize statistical information in-situ?

DKRZ

Large Data Visualization

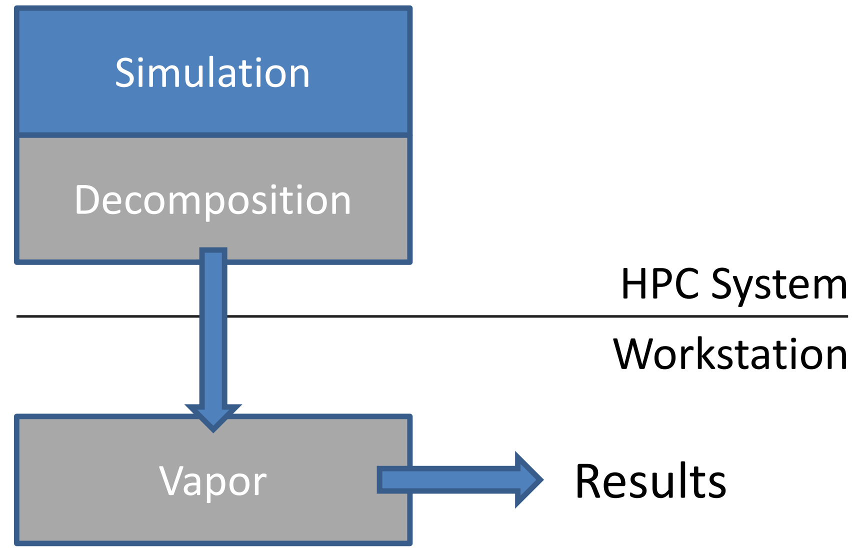
in-situ Visualization

(ParaView/Catalyst/Cinema)



in-situ Compression

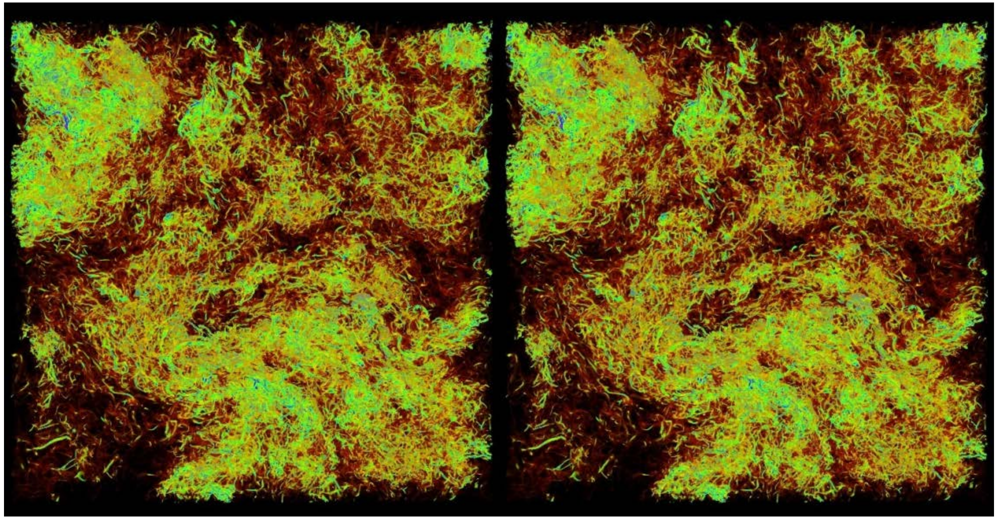
(Vapor)



VAPOR and Compression

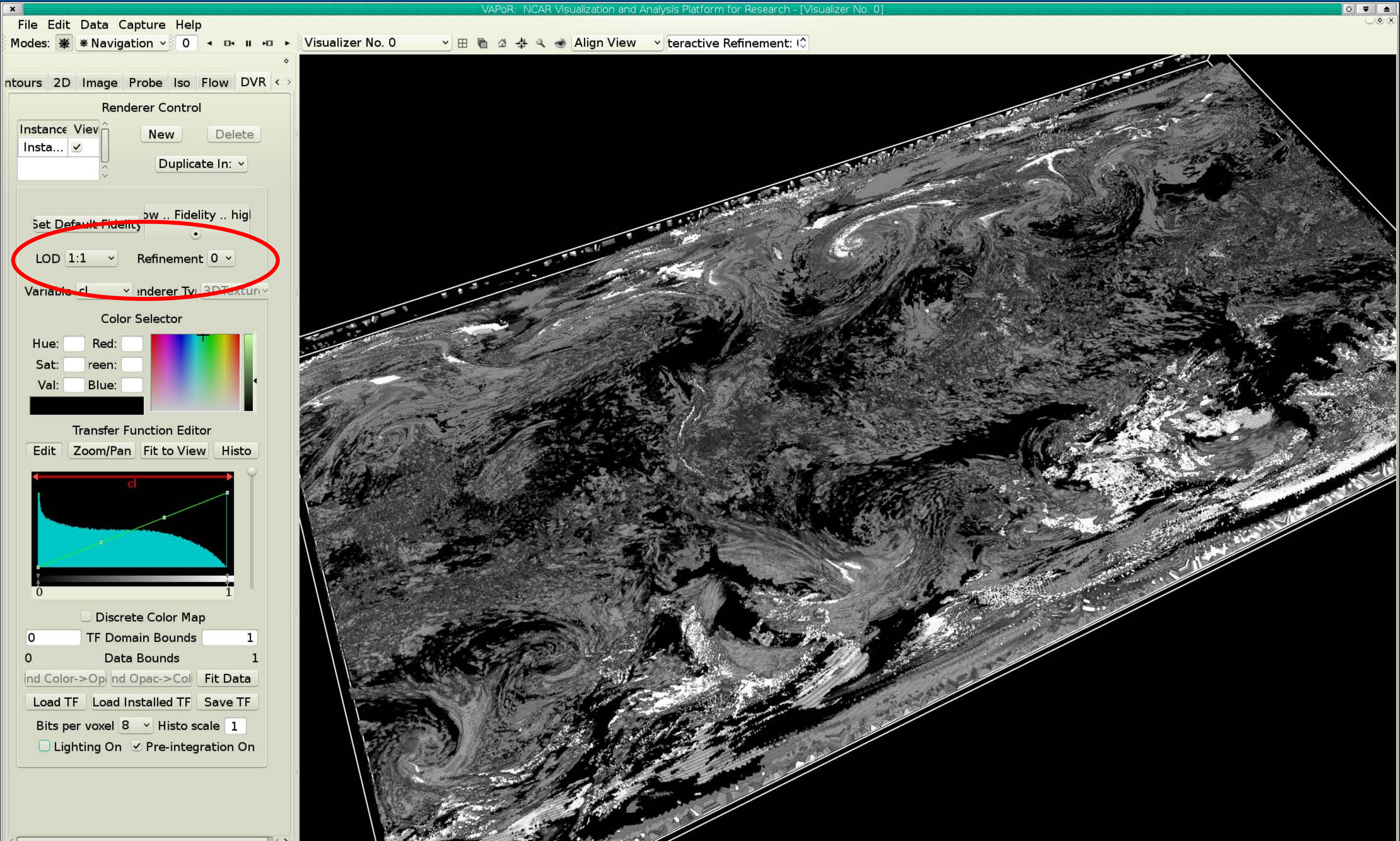
- Wavelet-based compression and data storage
- Progressive data access with multiresolution rendering
- Coefficients are sorted and prioritized

Original
275 GBs / 3D field

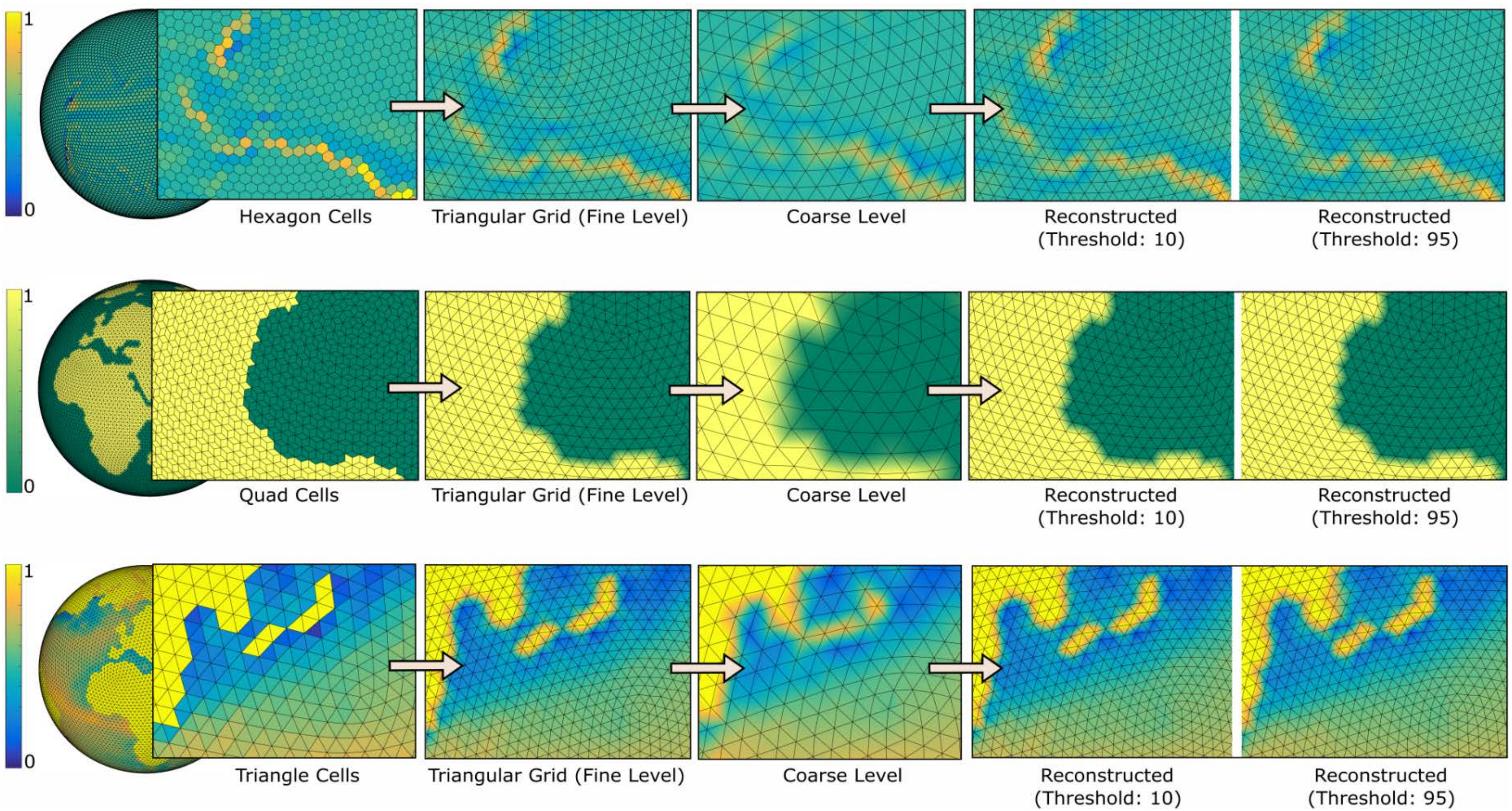


800:1
0.34 GBs / 3D field

(c) John Clyne, NCAR

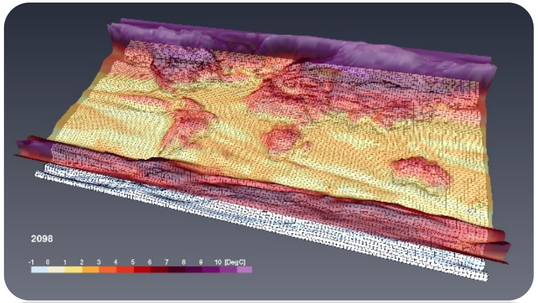

VAPoR: NCAR Visualization and Analysis Platform for Research - [Visualizer No. 0]
File Edit Data Capture Help
Modes: Navigation
Visualizer No. 0
Align View
teractive Refinement:
ntours 2D Image Probe Iso Flow DVR
Renderer Control
Instance View
Insta...
New
Delete
Duplicate In:
Set Default Fidelity
ow .. Fidelity .. higl
LOD 1:1
Refinement 0
Variable cl
enderer Ty 3DTexture
Color Selector
Hue:
Red:
Sat:
reen:
Val:
Blue:
Transfer Function Editor
Edit
Zoom/Pan
Fit to View
Histo
cl
0
1
Discrete Color Map
0
TF Domain Bounds
1
0
Data Bounds
1
ind Color->Op
nd Opac->Col
Fit Data
Load TF
Load Installed TF
Save TF
Bits per voxel 8
Histo scale 1
Lighting On
Pre-integration On

Multiresolution with Icosahedral Maps



[1] Jubair et.al. *"Icosahedral Maps for a Multiresolution Representation of Earth Data"*, VMV 2016

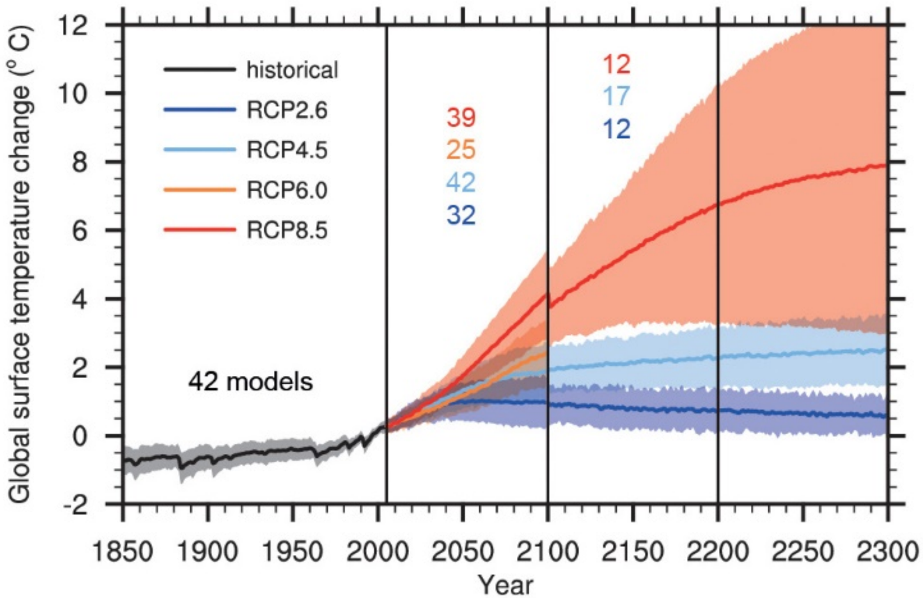
Challenge 2: Ensembles and Uncertainty



Uncertainty or Robustness?

- Uncertainty
  - What we don‘t know
  - Communicate some result plus some uncertainty range (model spread)
- Robustness
  - What we know (= positive)
  - Determine which part of the data (i.e. projected signals) is statistically significant

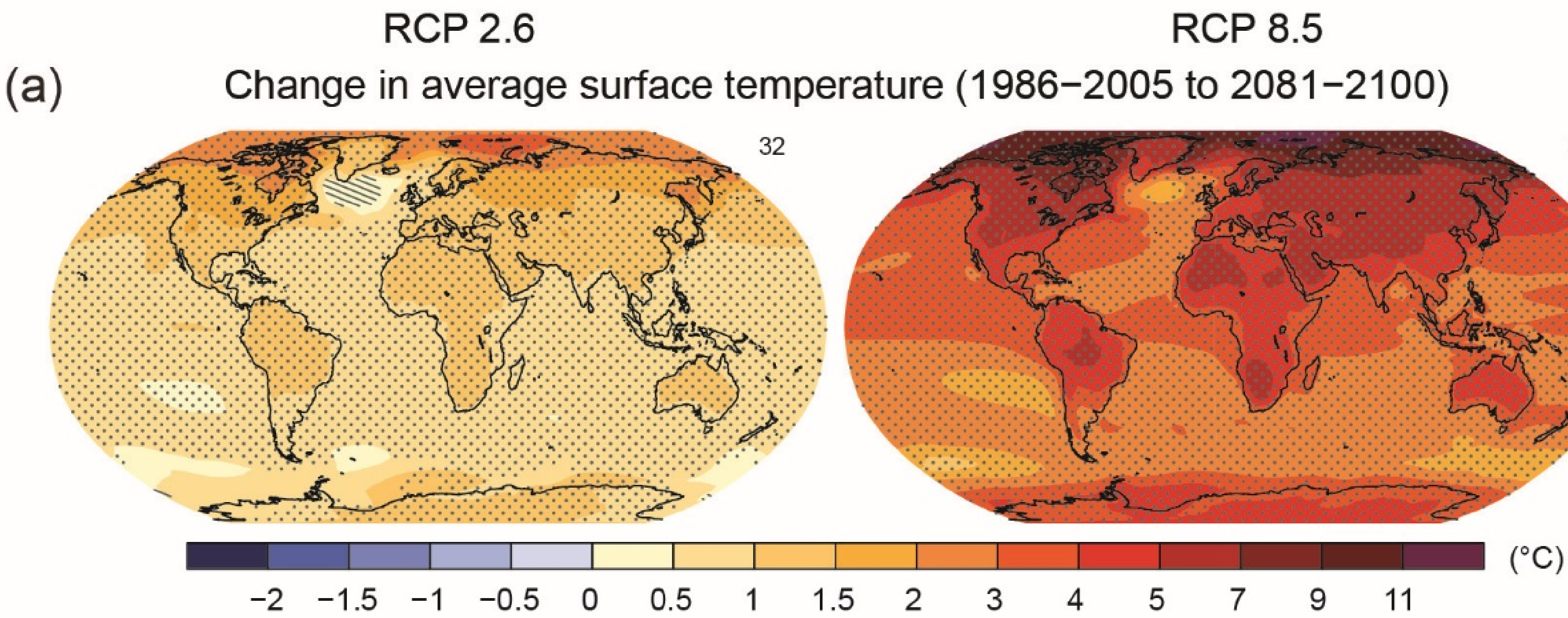
CMIP5 - Projected Changes in the 2m-Temperature



Source: IPCC AR5, TS Figure 15

Source: IPCC AR5, Figure 12-09

Overlayed Stippling and Hatching indicating Robustness



Stippling indicates multi-model mean is more than two standard deviations of natural internal variability in 20-yr means; hatching indicates multi-model mean is less than one standard deviation of natural internal variability in 20-yr means, and where at least 90% of models agree on the sign of change.

Source: IPCC AR5, SPM Figure 8

Sources of uncertainty

- Internal climate variability
- Model uncertainty:
  - representation of processes
  - climate sensitivity
- Scenario uncertainty (following example: only RCP4.5)

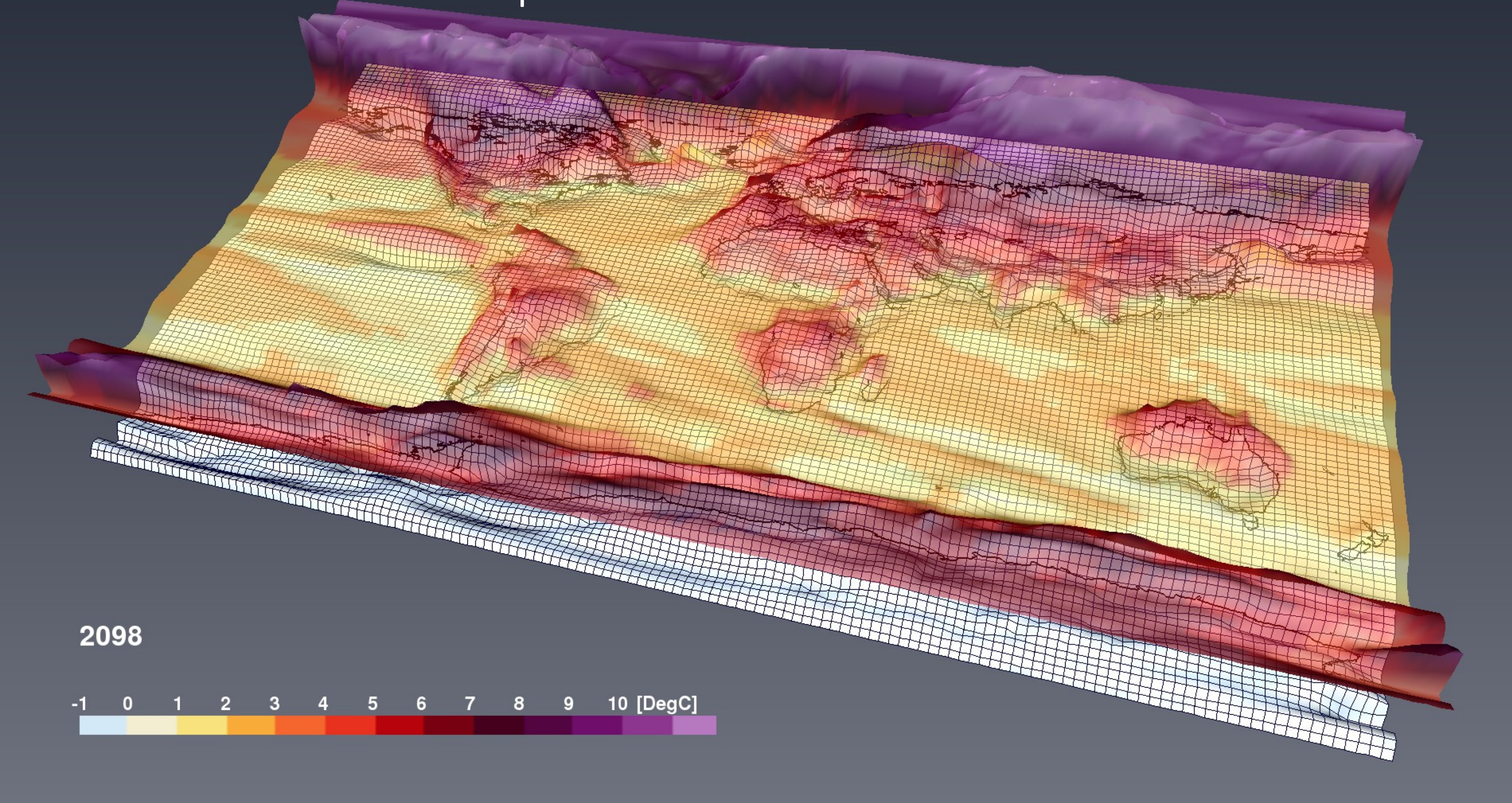
- Unknown volcanic activity (space, time, magnitude)

  Model and experiments described in

  Ingo Bethke, Stephen Outten, Odd Helge Otterå, Ed Hawkins, Sebastian Wagner, Michael Sigl & Peter Thorne: Potential volcanic impacts on future climate variability, Nature Climate Change volume 7, pages 799–805 (2017), doi:10.1038/nclimate3394

Uncertainty range of temperature anomaly

Possible Values: space between ensemble minimum and maximum



Development of ensemble mean, minimum and maximum

Conclusions

- Even with moderatly parallel vis resources, visualization of current ESM simulation data sizes is scratching the technical limits
- In-situ visualization, compression and analysis are promising technical solutions
  - Require paradigm and workflow changes in scientific practice
- Ensemble Simulations add further dimension
  - Robustness versus uncertainty
  - Stippling/hatching versus semi-transparent shading
  - Combined heightfield displays of ensemble miniumum and maximum visualize the uncertainty space

The CMIP5 visualization shown in Challenge 2 was used in a recent video produced by the WCRP (World Climate Research Progamme) to promote the CMIP project.

https://www.youtube.com/watch?v=WdRiYPJLt4o

Questions?

boettinger@dkrz.de

www.dkrz.de