Interactive Visual Analysis of Multi-faceted Scientific Data

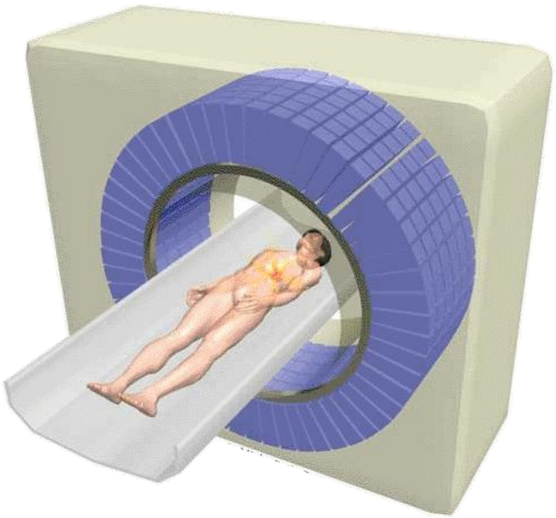
Johannes Kehrer

Visualization Group, Dept. of Informatics
University of Bergen, Norway

www.ii.UiB.no/vis

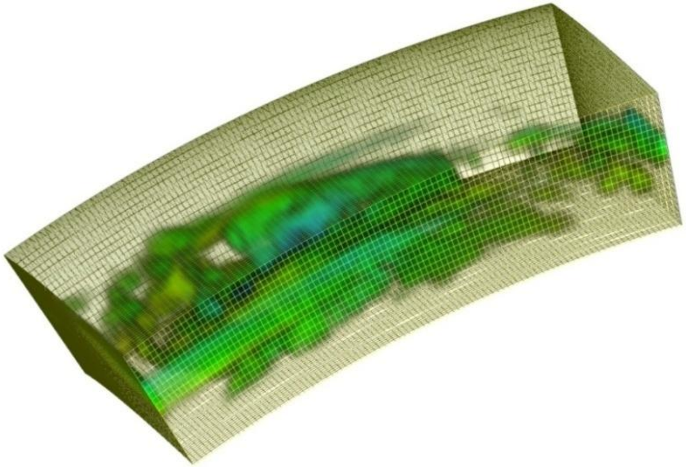
Motivation

Increasing amounts of scientific data



medical scanner

computational simulation

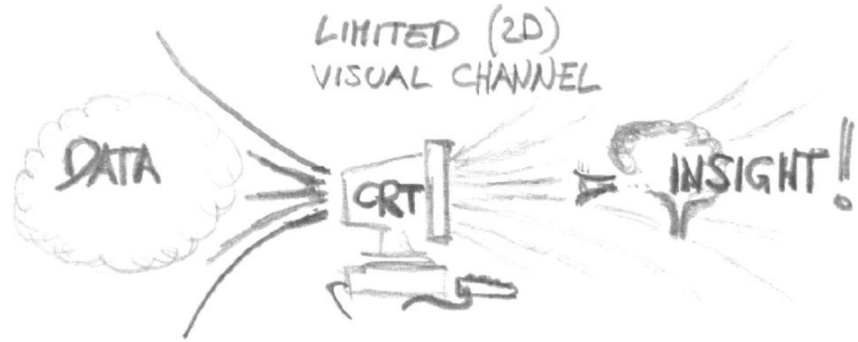


time-dependent
3D data

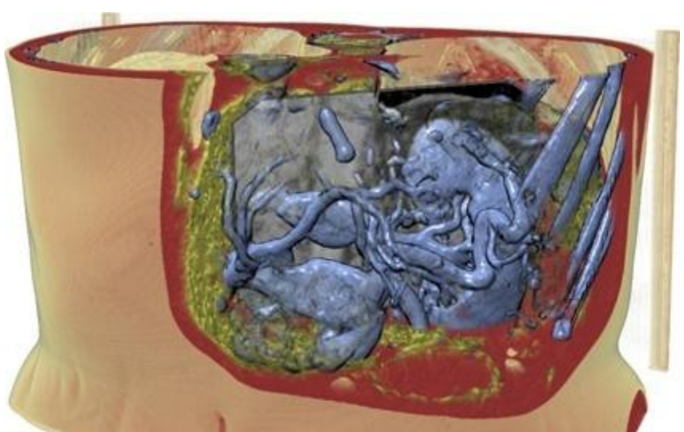
Hard to analyze and understand

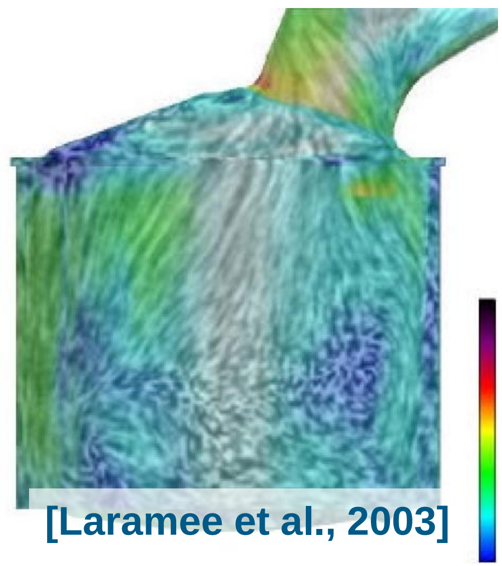
Visualization

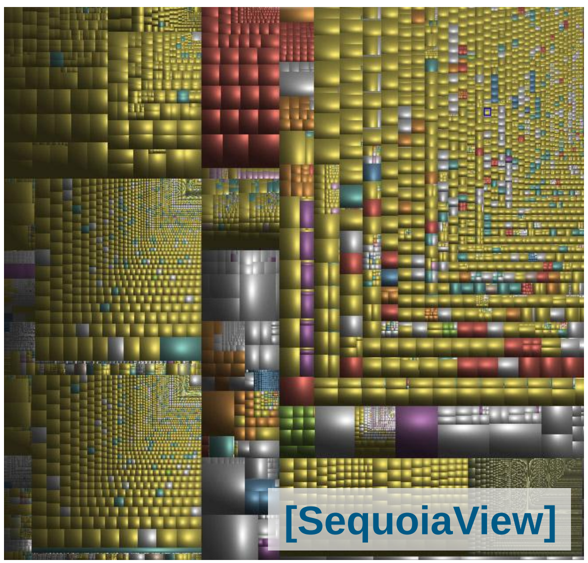
*“The purpose of visualization is **insight**, not pictures”*
[Shneiderman ’99]



Different application areas


[Burns et al., 2007]

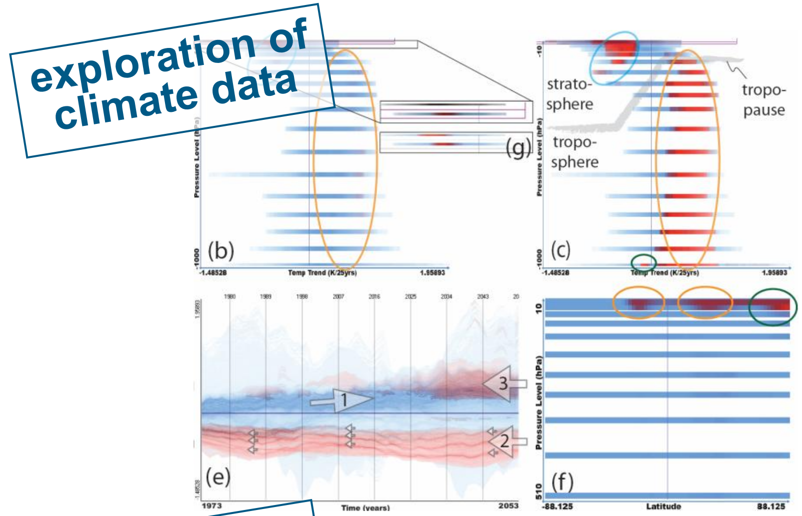

[Laramee et al., 2003]

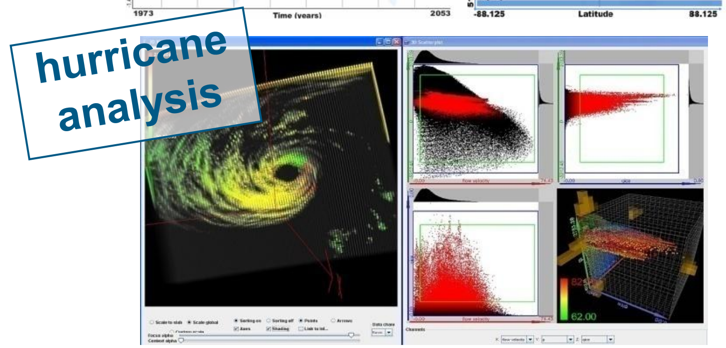

[SequoiaView]

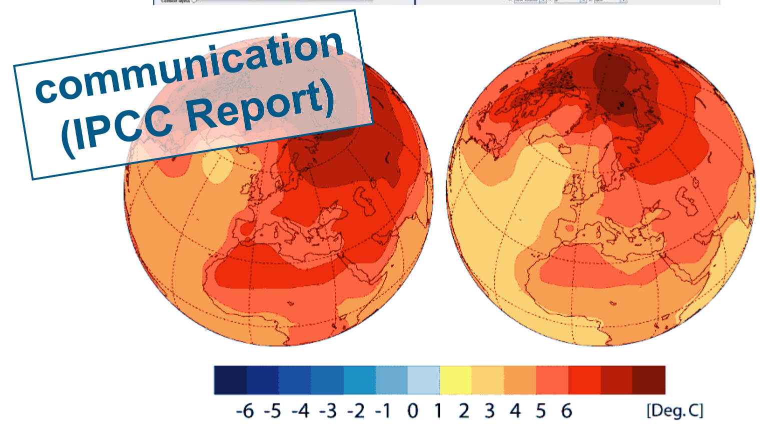
Typical Visualization Tasks

Visualization is good for

- visual **exploration**
  - find unknown/unexpected
  - generate new hypothesis
- visual **analysis** (confirmative vis.)
  - verify or reject hypotheses
  - information drill-down
- **presentation**
  - show/communicate results



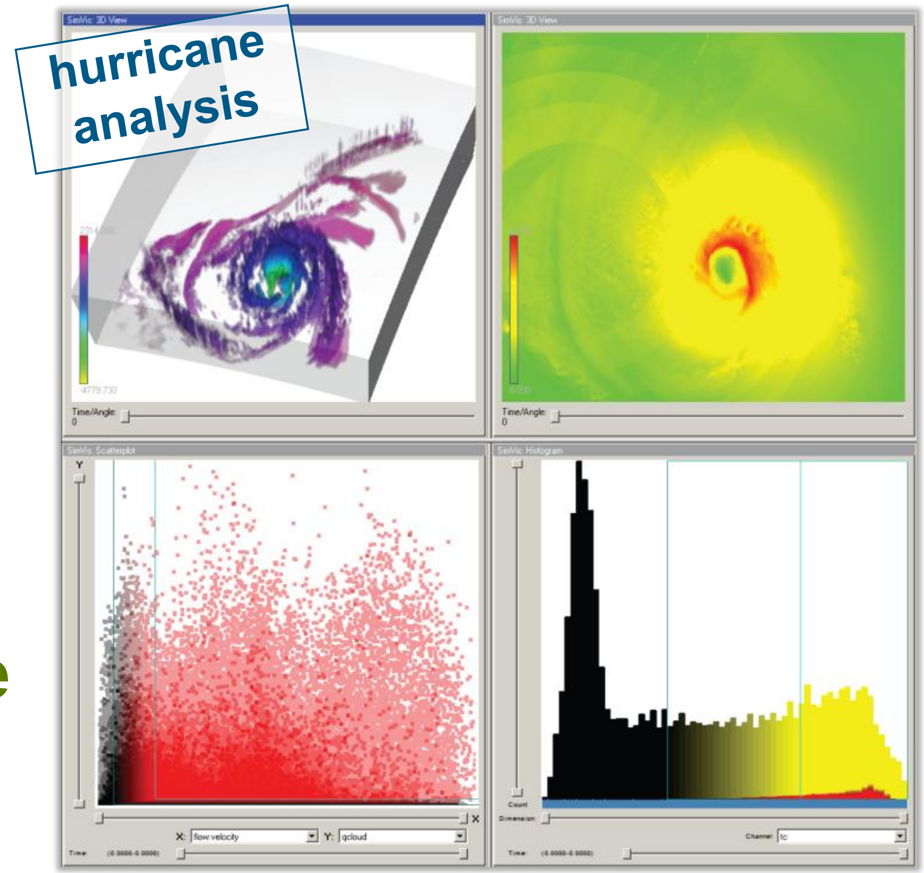




Interactive Visual Analysis (IVA)

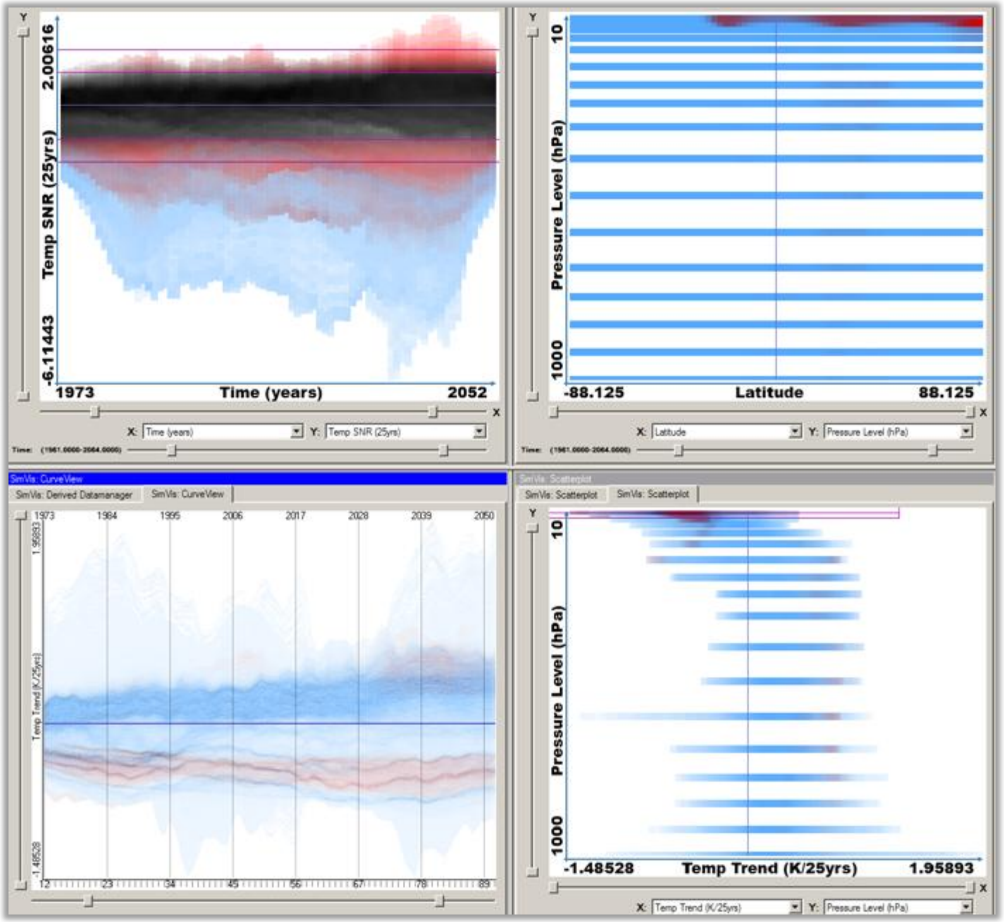
Enables visual dialogue between **user** and **data**

- **drill-down** into information ("overview first, zoom and filter, then details on demand" [Shneiderman])
- **interpret** complex data
- **find relations** ("read between the lines")
- detect **features / patterns** that are difficult to describe
- integrate **expert knowledge**



SimVis Framework for IVA

- coordinated, multiple views
- linking & brushing
- focus+context vis.
- degree-of-interest
  (DOI $\in$ [0, 1])
- on-the-fly data derivation
- interactivity, etc.

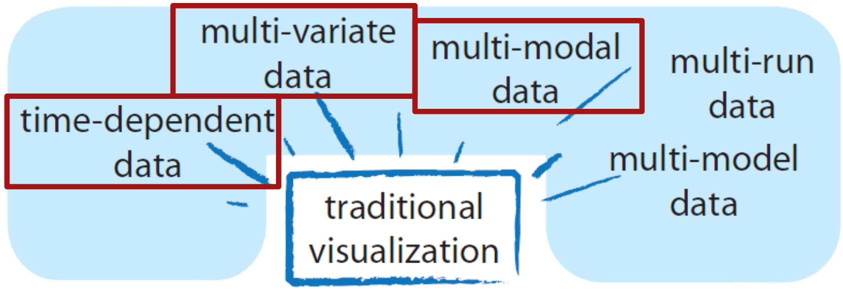


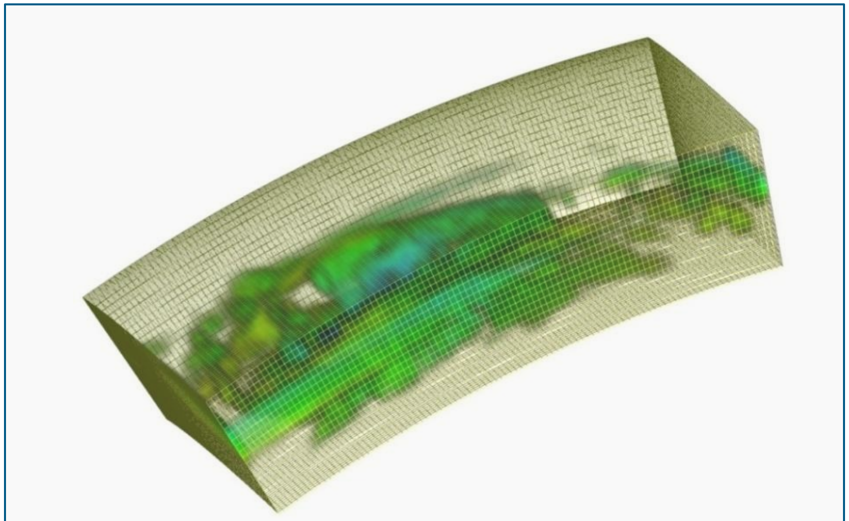
Multi-faceted Scientific Data

- Time-dependent scenarios
  (consider multiple time steps)

- Multi-variate data
  (multiple data variates, e.g., temperature, precipitation)

- Multi-modal data
  (simulation, satellite imagery, weather stations, etc.)

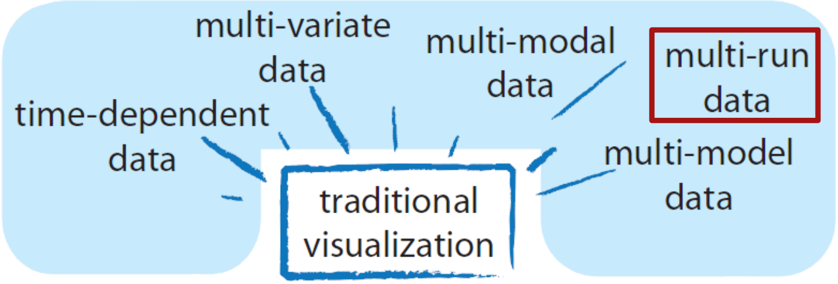


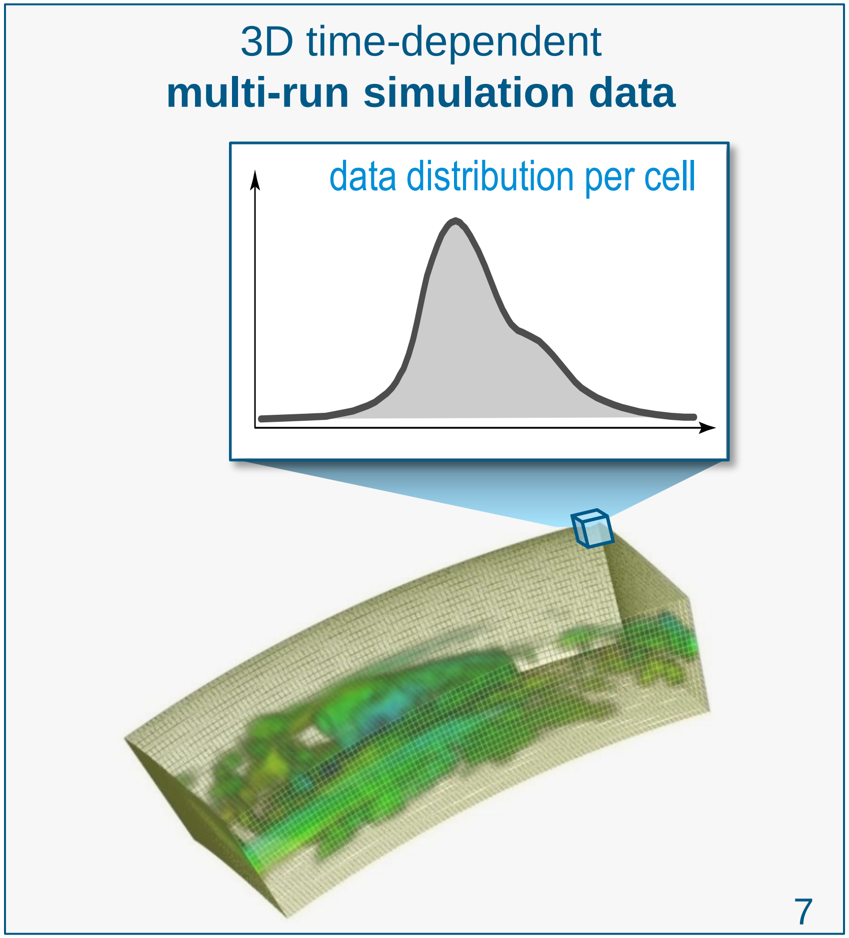


time-dependent 3D data

Multi-faceted Scientific Data

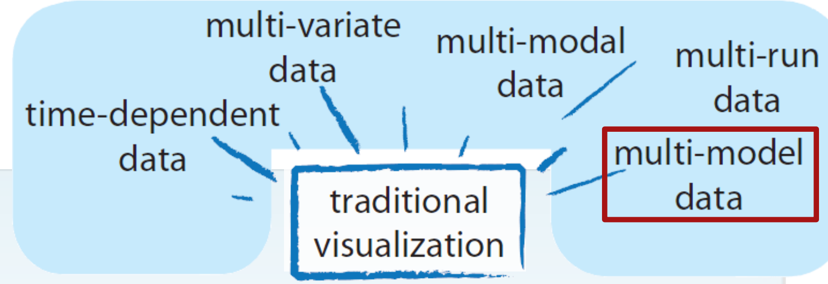
- Multi-run simulations
  (simulation repeated with varied model parameters)

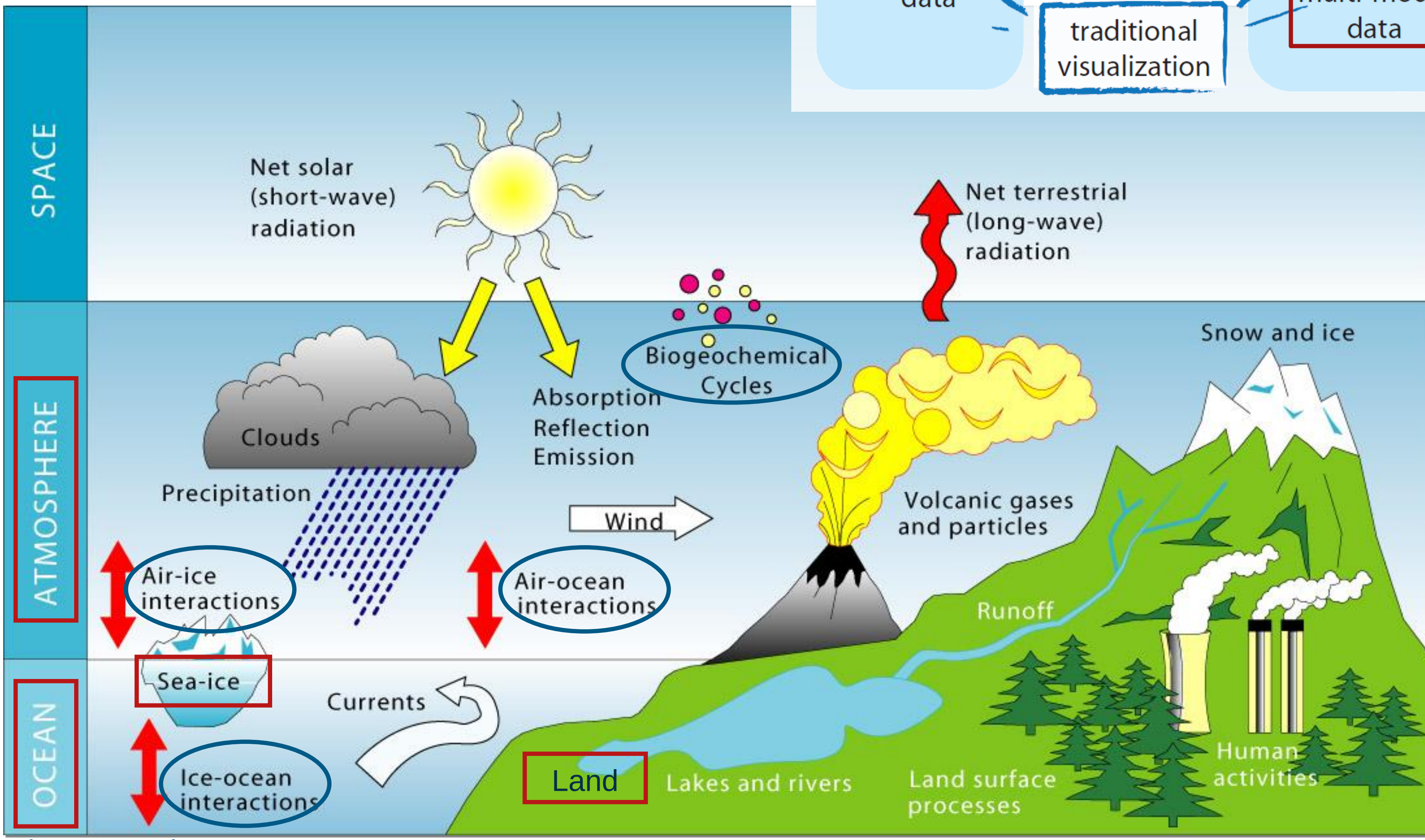




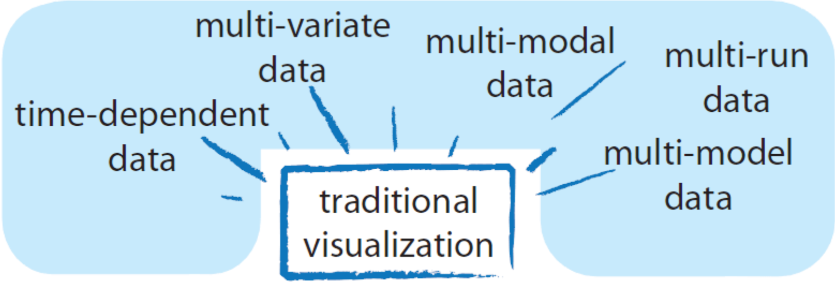
Multi-faceted Scientific Data

- Multi-model scenarios
  (e.g., coupled climate models)





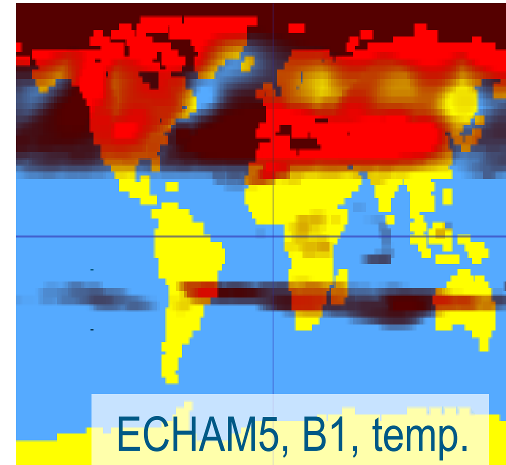
Contributions

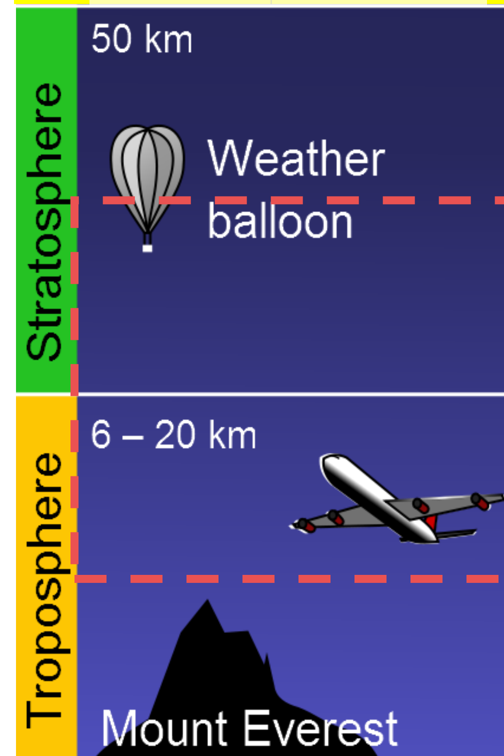


- IVA of **multi-run data**
- IVA across **2 data parts**
  (multi-model / multi-run data)
- IVA of multi-run data based on **statistical moments**
- Strategies for IVA for **hypothesis generation** in climate research
- Design guidelines for **glyph-based 3D visualization**

Hypothesis Generation

- search for potential sensitive & robust **indicators** for climate change
- characteristic climate signals that deviate from natural variability
- useful to monitor atmospheric change





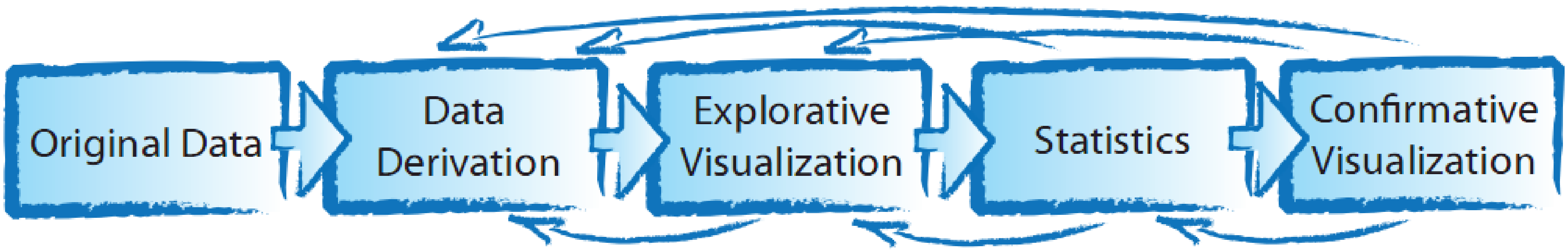
Usual Workflow

- Set research focus
- Acquire data
- Iterate
  - explore / investigate data
  - formulate particular hypothesis
  - evaluate with statistics

Challenging to come up with new hypotheses

Goal: accelerate process (fast interactive visualization, more informed partner → more directed search)

Our Visual Exploration Process

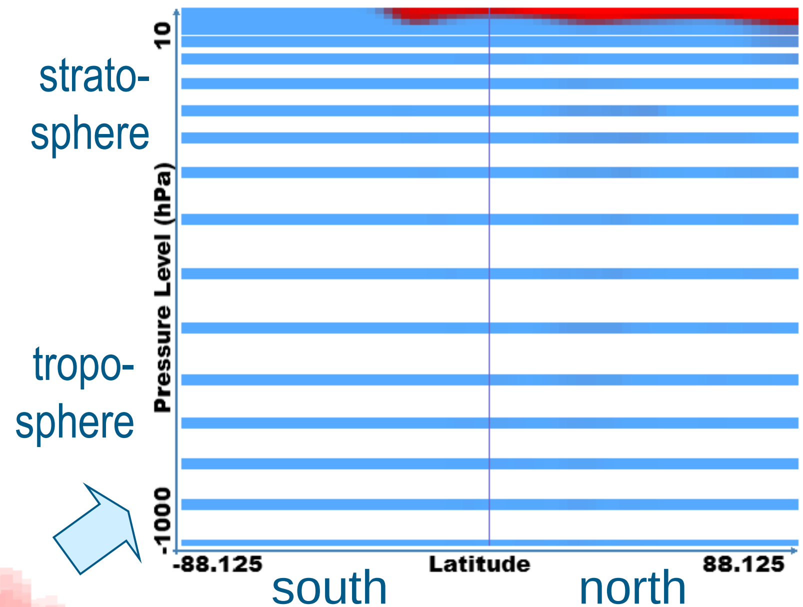


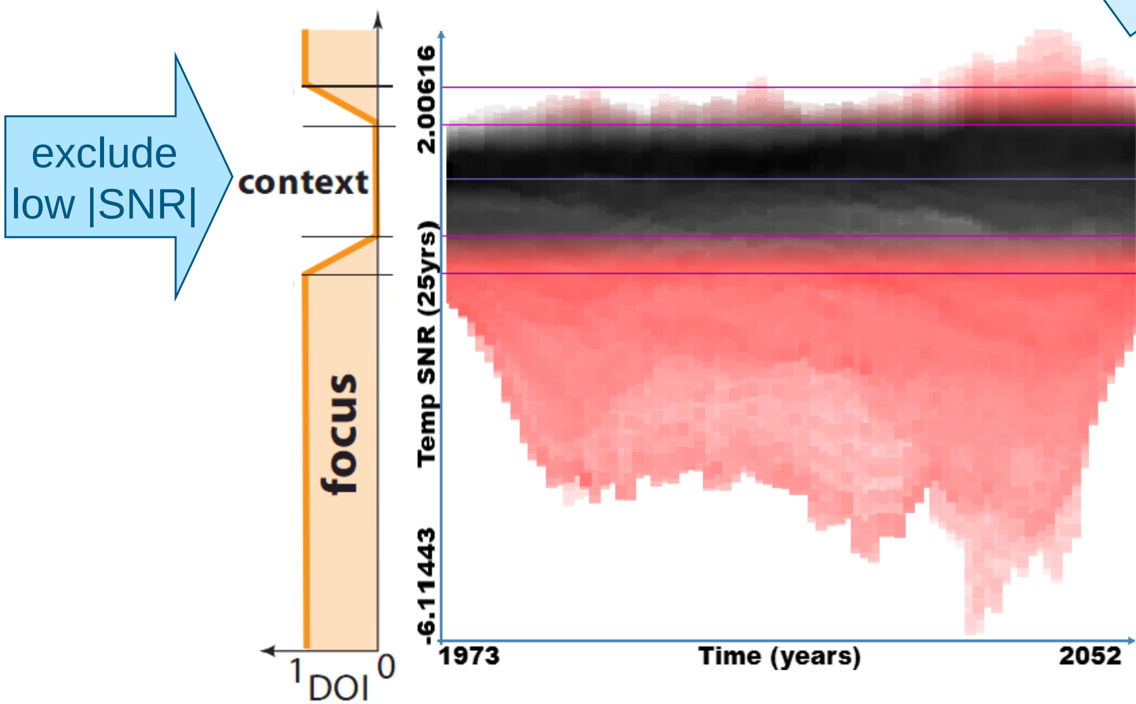
- Integrated **data derivation**
  → linear trends & signal to noise ratios (SNR)
- Interactive **visual exploration** for quick and flexible data investigation ("preview on statistics")
- Generated hypotheses evaluated using **statistics**
  → trend testing [Lackner et al. 08]
- Narrow down **parameters**

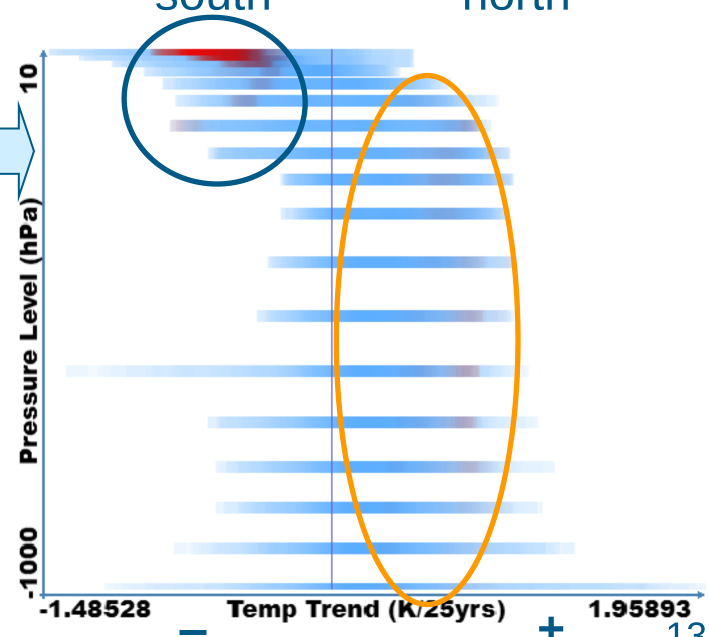
Focus on Expressive Data

Localize robust indicators

- areas with high significance
- smooth specification

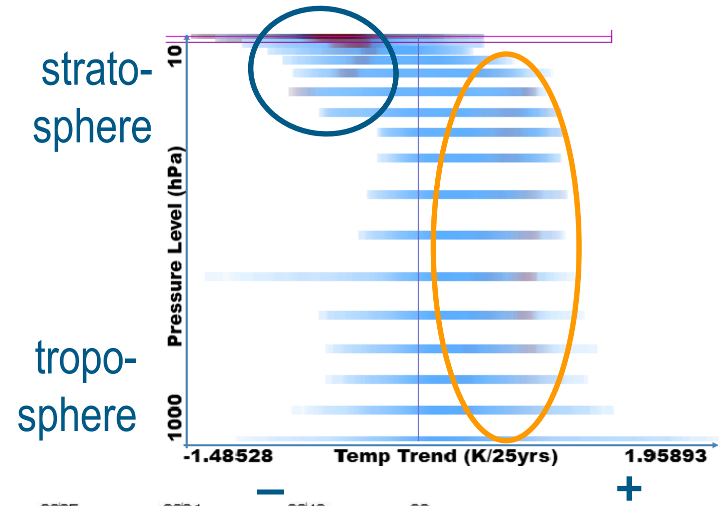


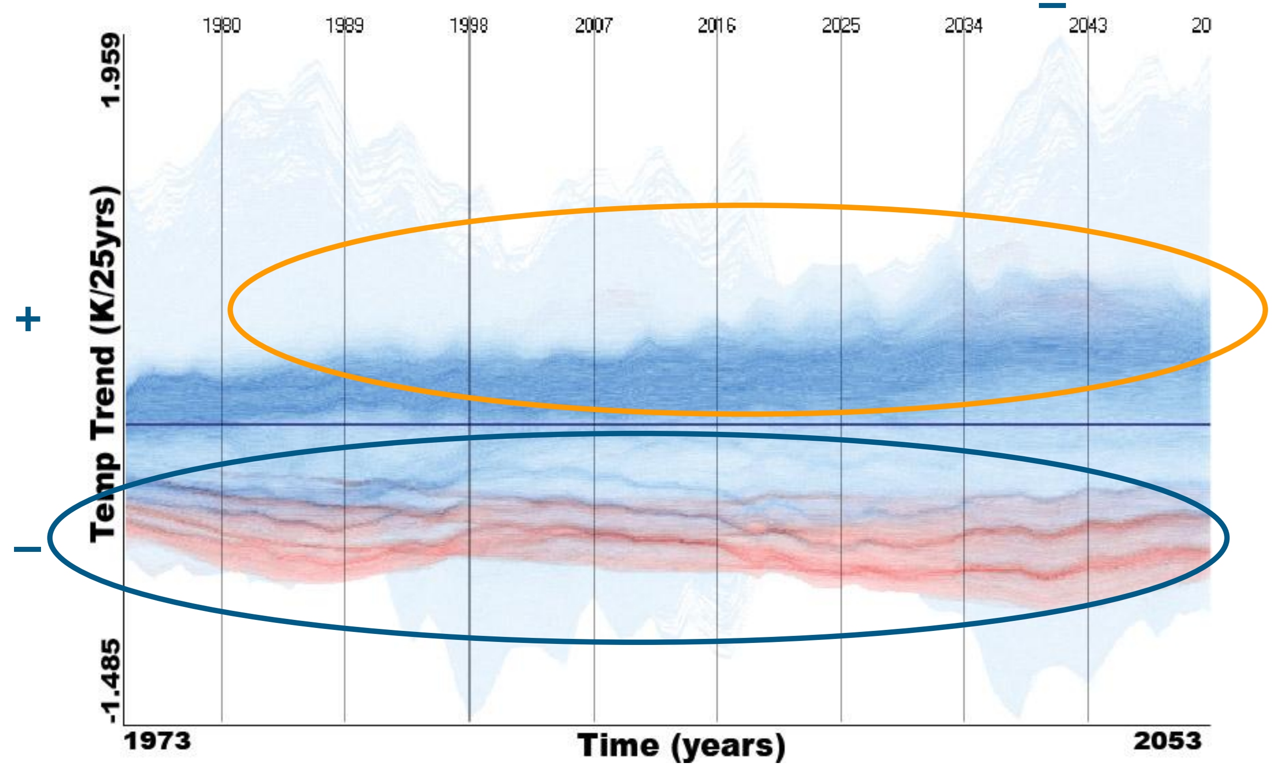




Explore Trend Variation over Time

Robust cooling trends

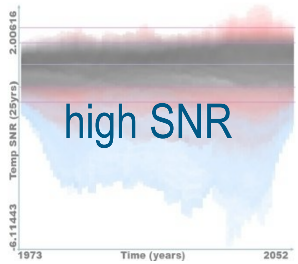


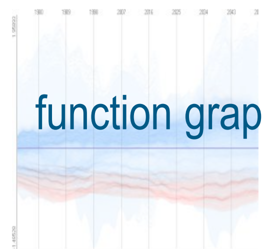


Analyze Relations between Dimensions

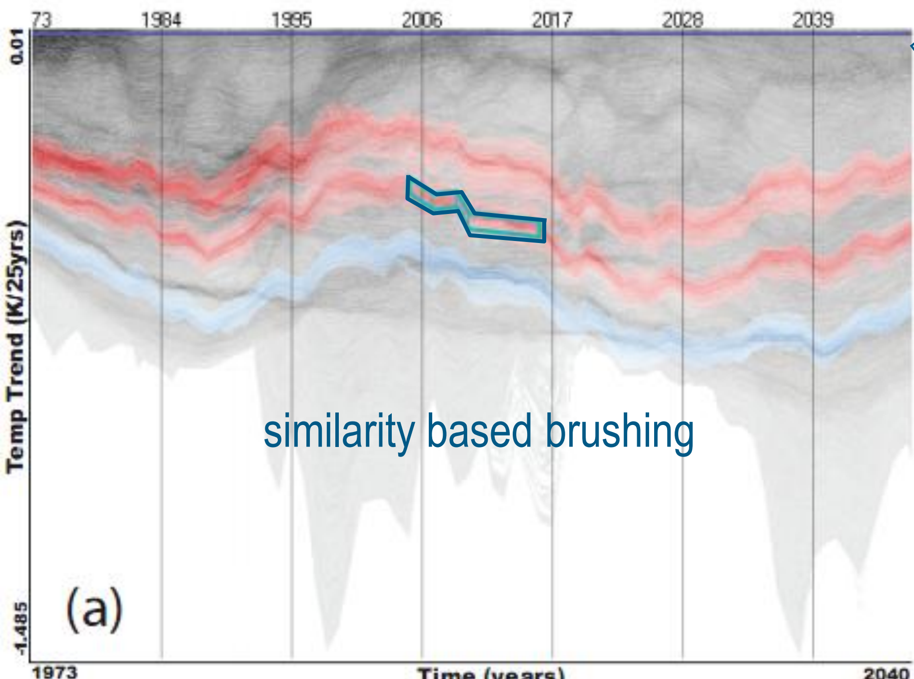
Up to now:

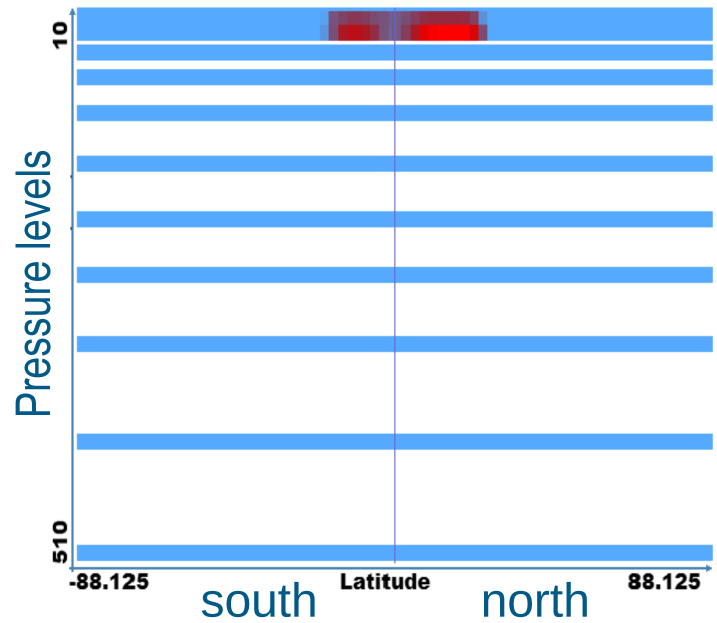
→ investigation in one direction

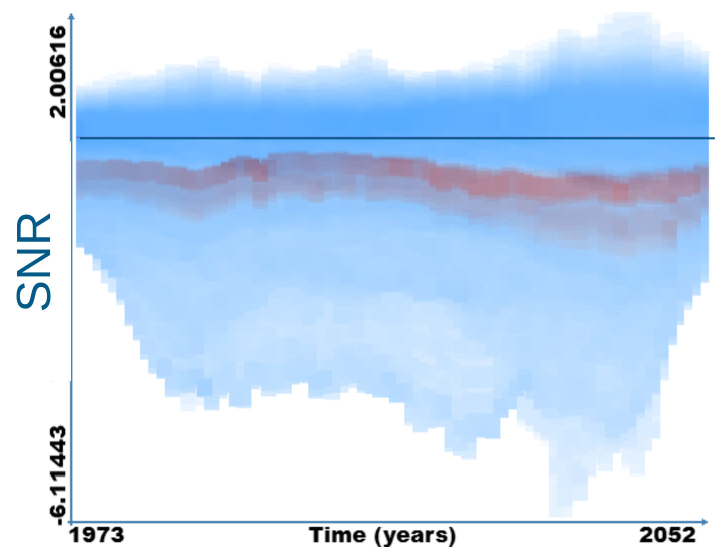




→ check relation in other direction

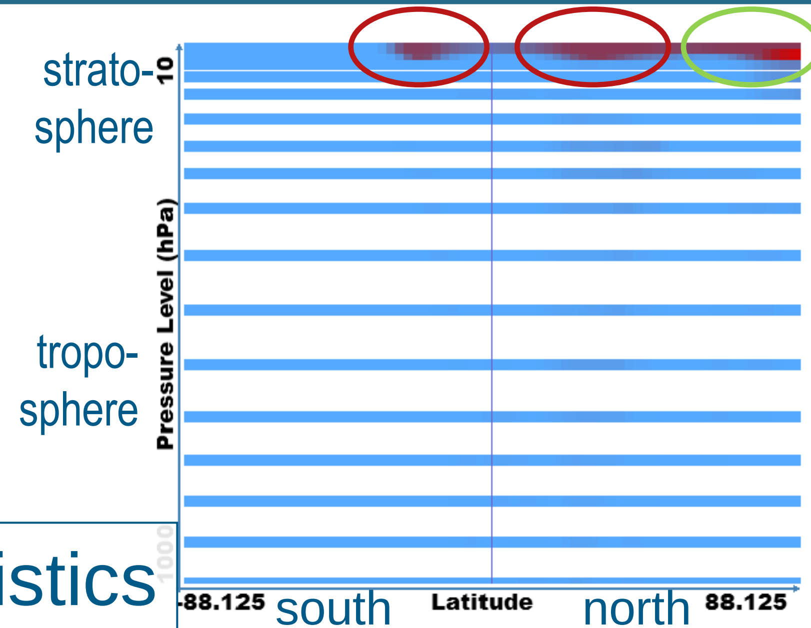




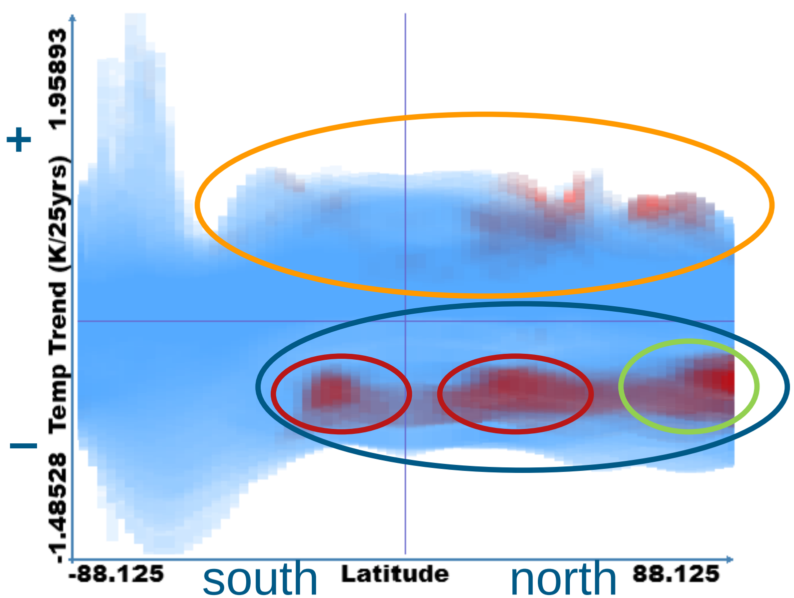


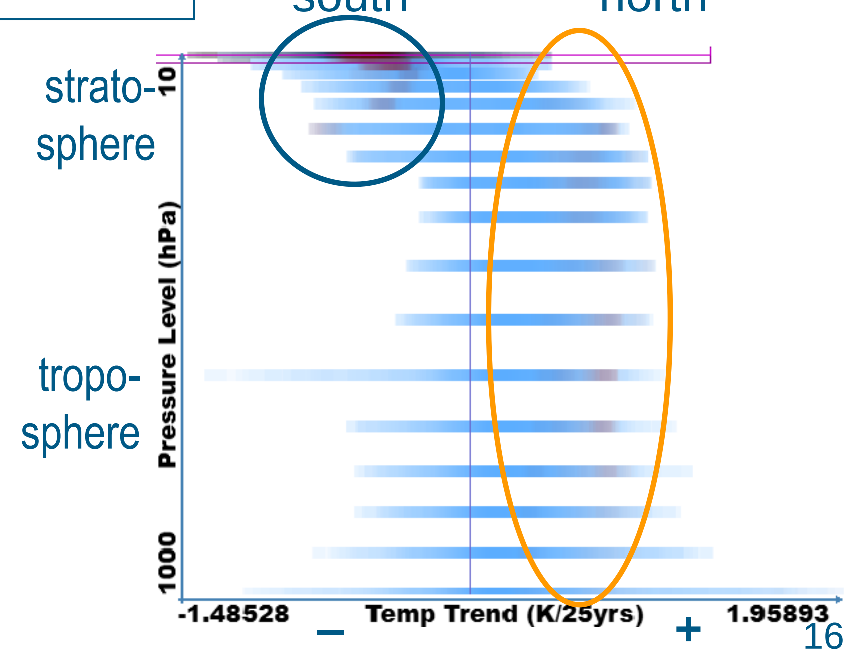
Generated Hypothesis / ECHAM5 temp.

Promising indicator region in **lower stratosphere** at **northern latitudes** & **tropics**. **Cooling trend** considered robust over investigated time span.



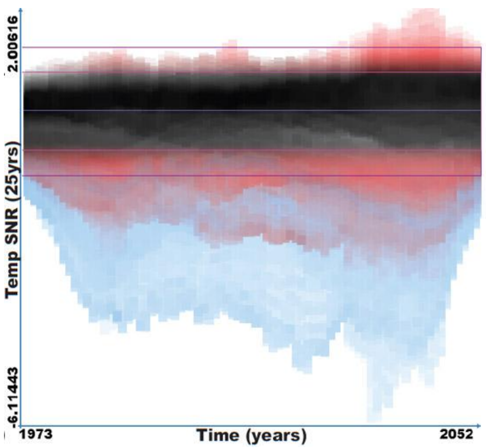
hypothesis handed over to statistics

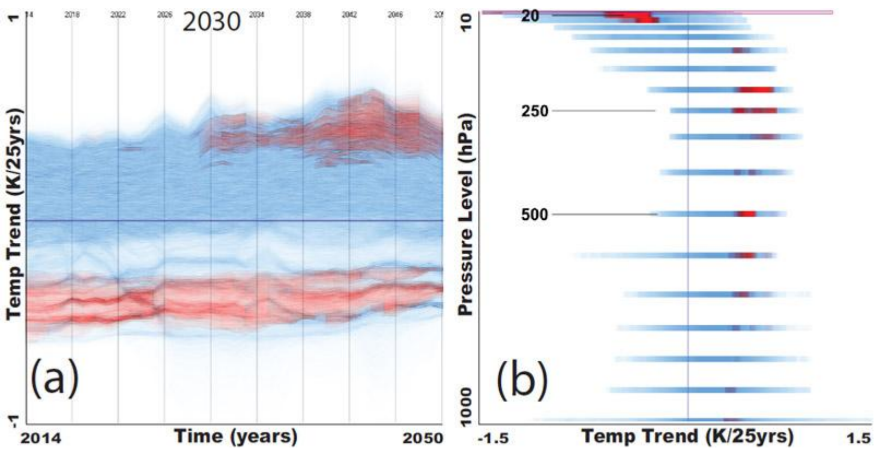


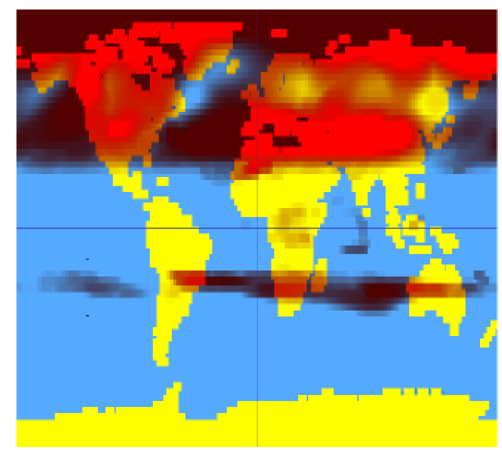


Hypothesis Generation with Visual Exploration

- Kehrer et al. **Hypothesis generation in climate research with interactive visual data exploration.** *IEEE TVCG, 14(6):1579–1586, 2008.*
- Ladstädter et al. **SimVis: an interactive visual field exploration tool applied to climate research**. In *New Horizons in Occultation Research,* pages 235–245. Springer, 2009.
- Ladstädter et al. **Exploration of climate data using interactive visualization**. *Journal of Atmospheric and Oceanic Technology, 27(4):667–679, 2010.*

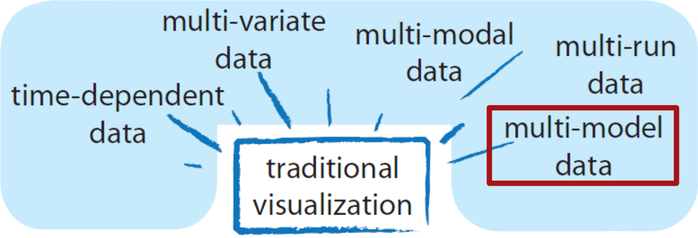




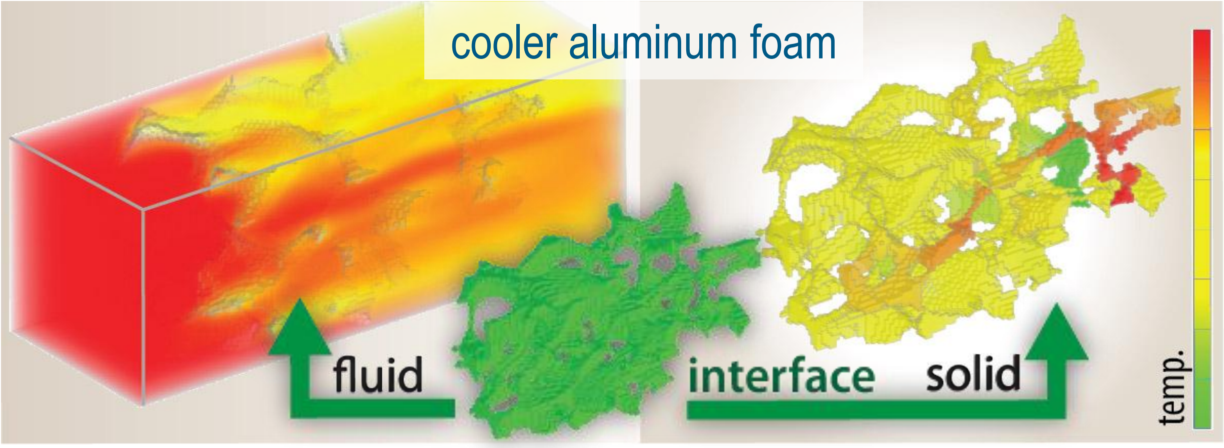


IVA across two Parts of Scientific Data

Multi-part scenarios



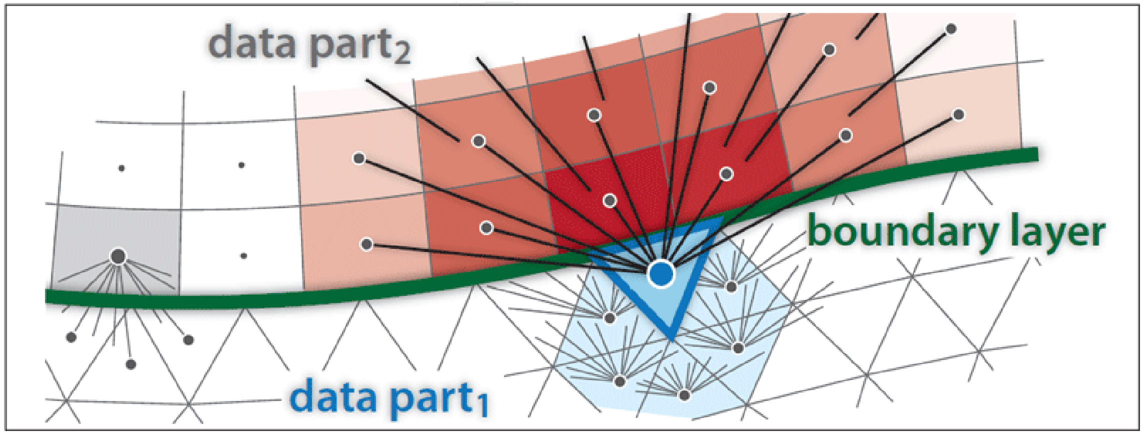
- Coupled atmosphere-ocean model
- Fluid-structure interactions (FSIs)



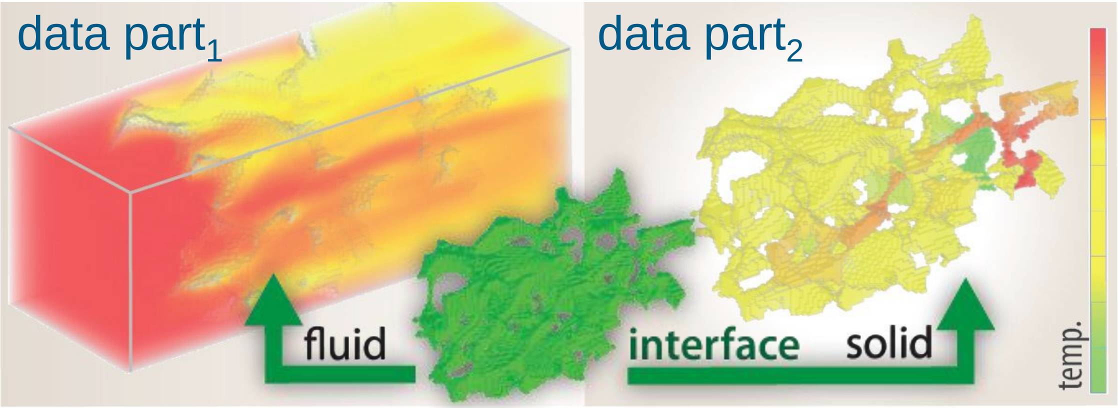
How to relate features across different data parts?

IVA across an Interface

- **Relate** grid cells across data parts
- **Transfer** features (DOI values) in both directions

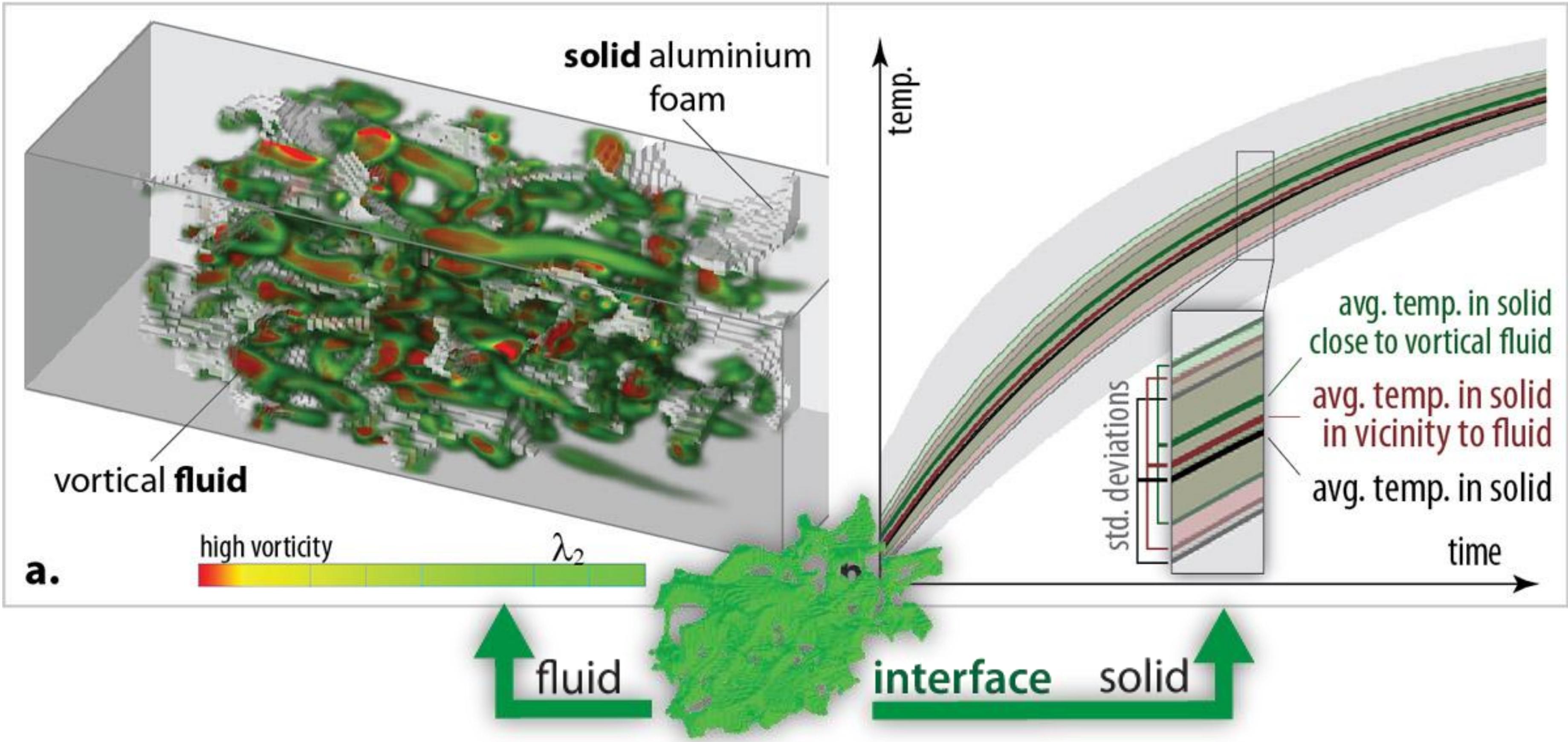


- Keep feature specification **up to date**



Heat Exchange in an FSI Scenario

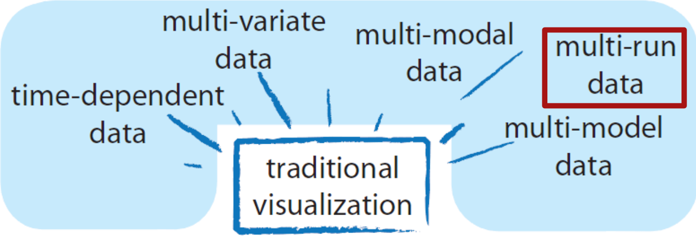
Transfer vortex feature to solid



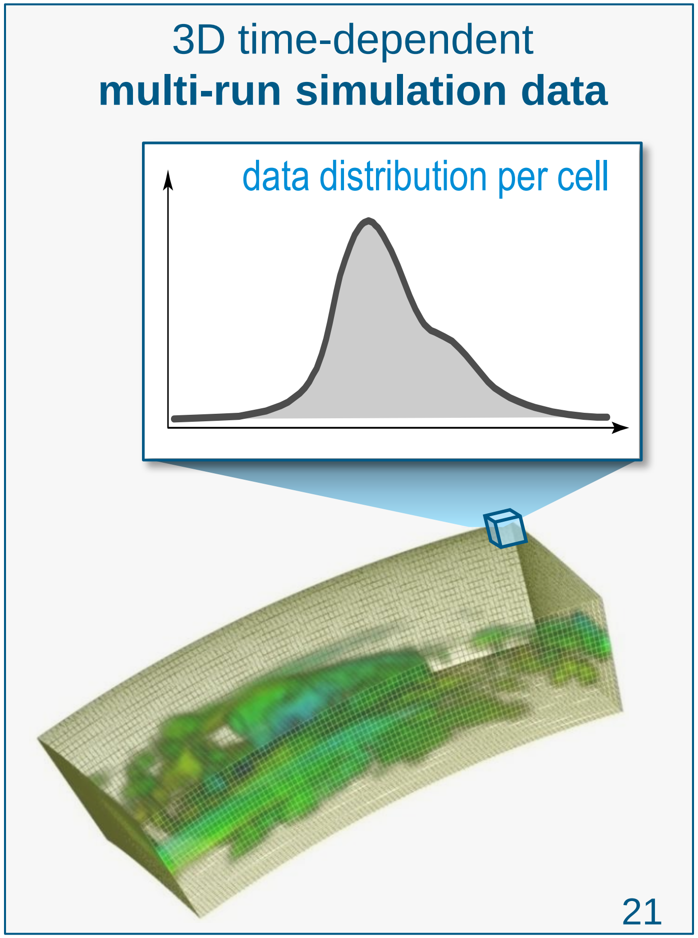
Relation: vortical flow ⇔ heating in solid

Higher-dimensional Scientific Data

- “Scientific” data:
  - some **data values *f(p)***
    (e.g., temperature, pressure values)
  - measured/simulated wrt. a **domain *p***
    (e.g., 2D/3D space, time, simulation input parameters)



- If dimensionality of ***p*** > 3, then traditional visual analysis is hard
- Reducing the data dimensionality can help (e.g., computing stat. aggregates)

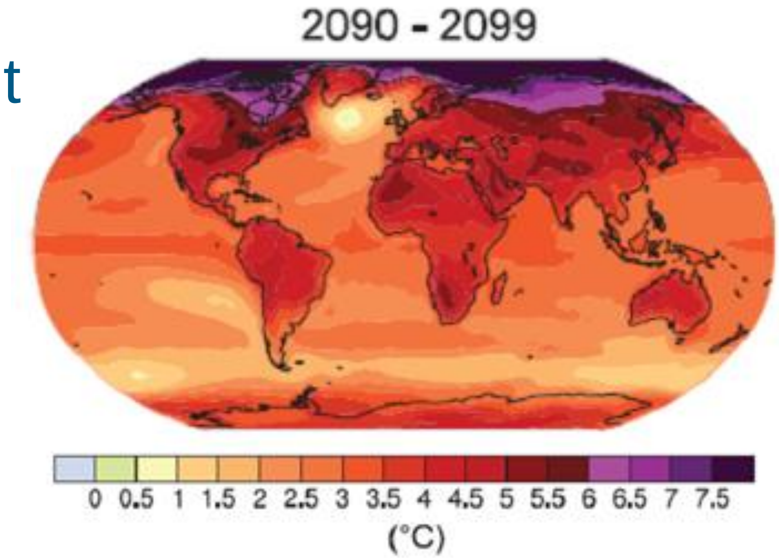


Reducing the Data Dimensionality

- **Statistics:** assess distributional characteristics along an independent dimension (e.g., time, spatial axes)

- Integrate into IVA through **attribute derivation**

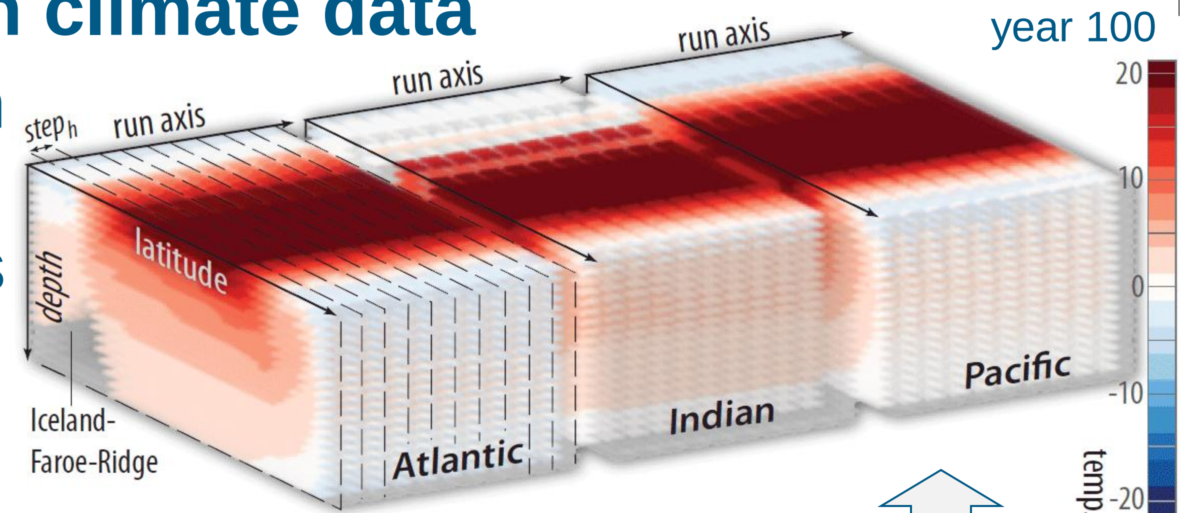
[from IPCC AR #4, 2007]



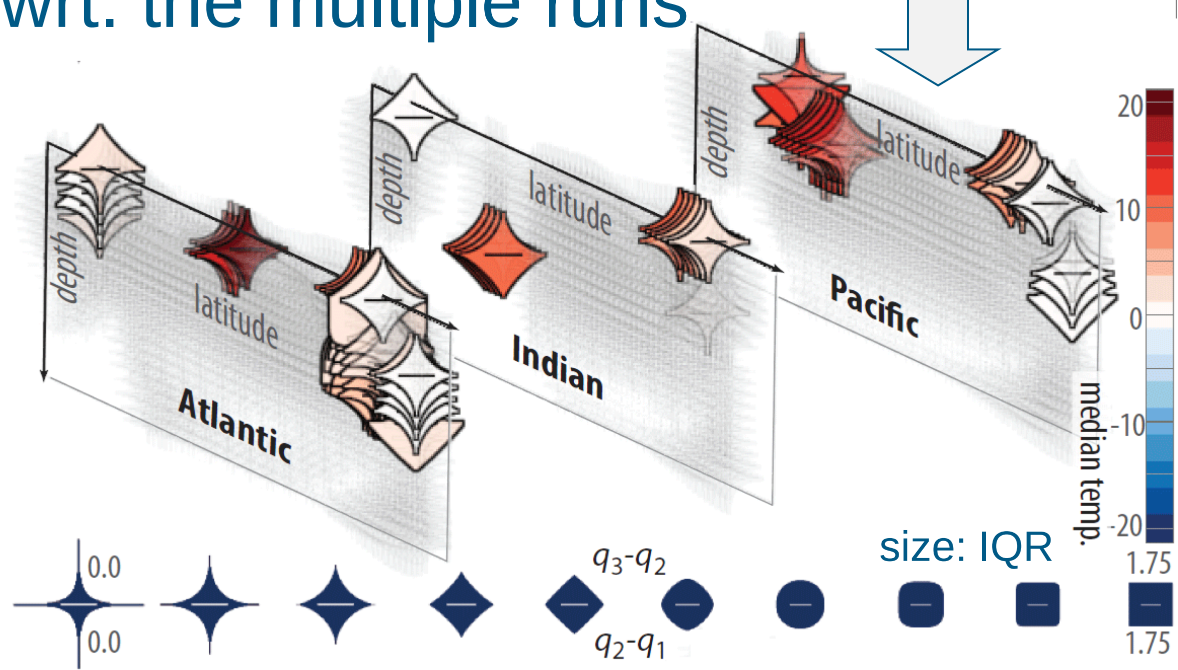
average temp. in ten years

Integrating Statistics and IVA

- Example: **Multi-run climate data**
  - ocean simulation
    (2D sections)
  - 10 x 10 = 100 runs
  - time-dependent
    (250 time steps)



- Compute statistics wrt. the multiple runs

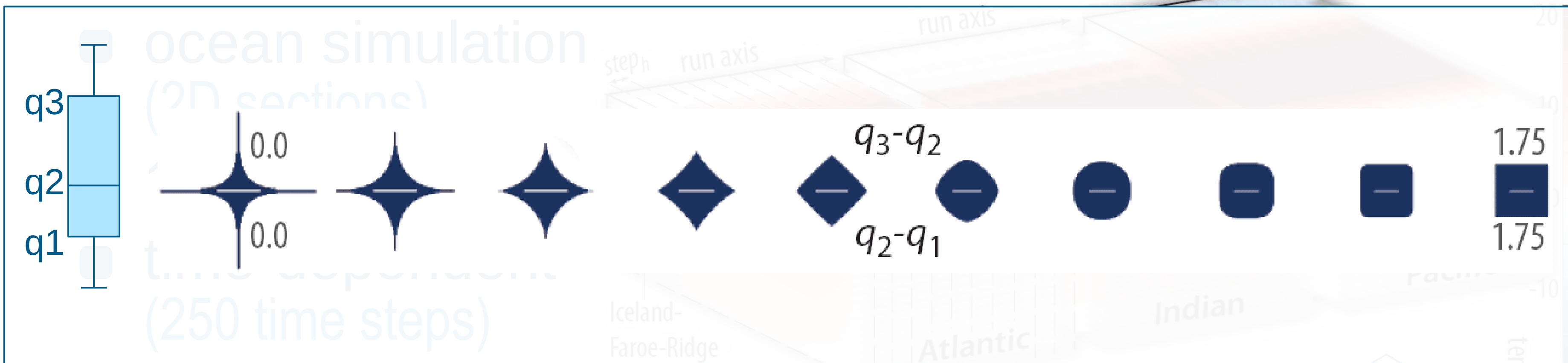


Integrating Statistics and IVA

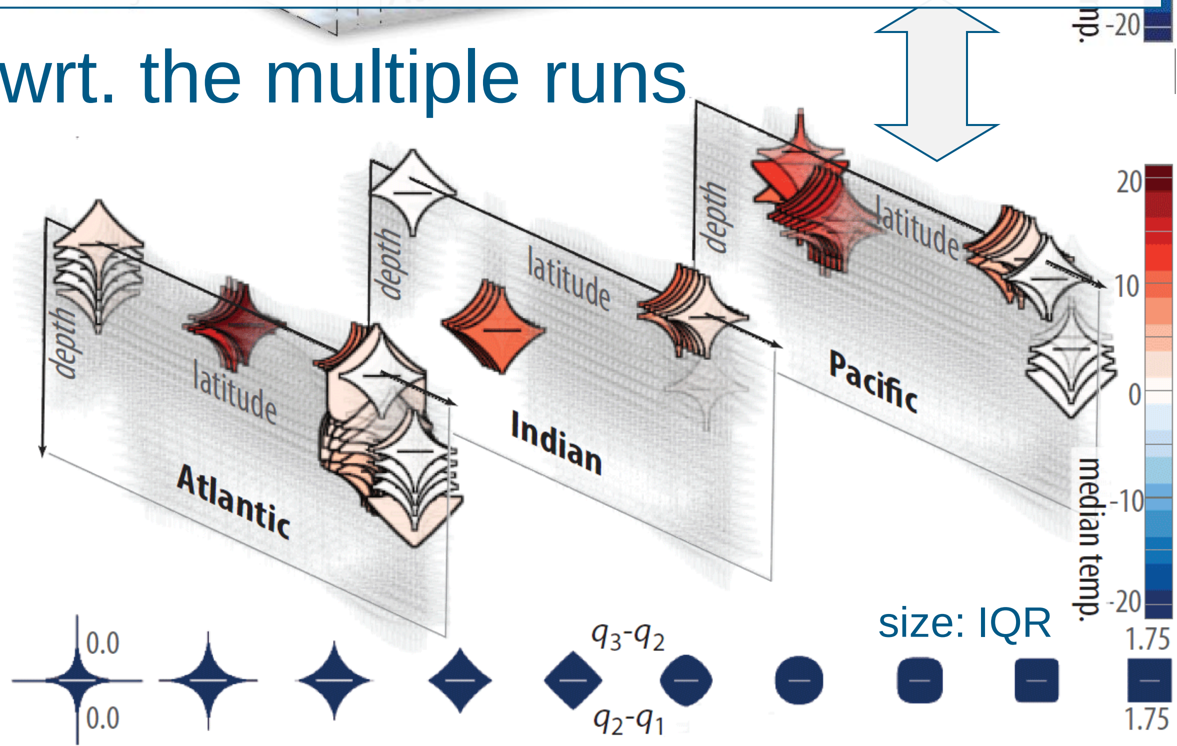
- Example: **Multi-run climate data**

run axis

year 100



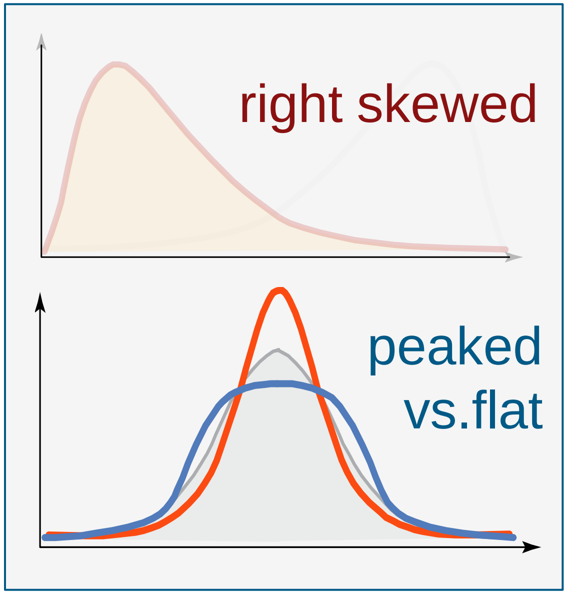
- Compute statistics wrt. the multiple runs



Moment-based Visual Analysis

- Get big picture (data trends & outliers)
- Multitude of choices, e.g,
  - 4 statistical moments (mean, std. deviation, skewness, kurtosis)
  - ×3 traditional and 2 robust estimates
  - ×2 compute relation (e.g., differences, ratio)
  - ×3 change scale (e.g., data normalization, log. scaling, measure of "outlyingness")

= 72 possible configurations per axis

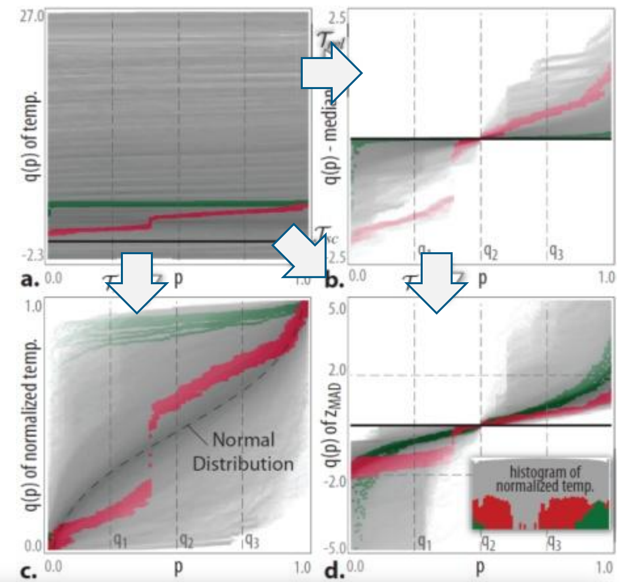


How to deal with this "management challenge"?

Structured approach to manage complexity

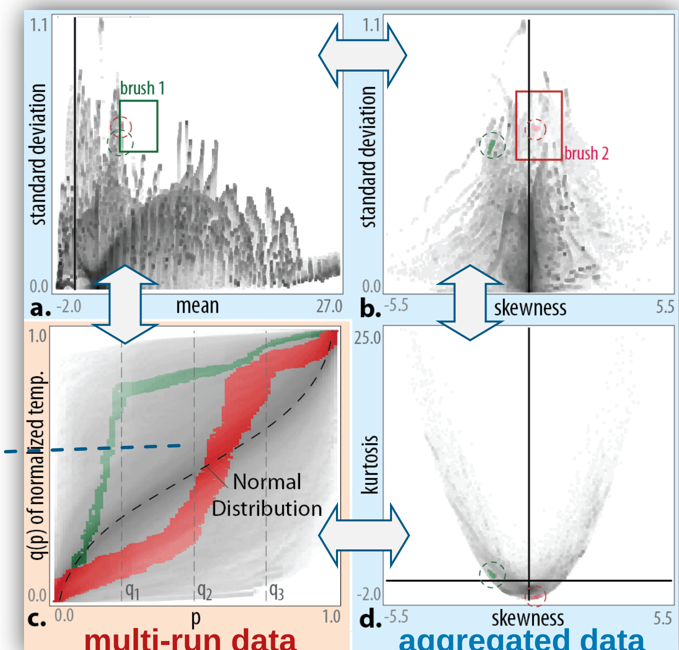
Moment-based Visual Analysis

- Iterative view transformations
  - alter axis/attribute configuration
    (construct a multitude of informative views)
  - maintain mental model of views
  - classification of moment-based views



- Relate
  multi-run data ⇔ aggregated data

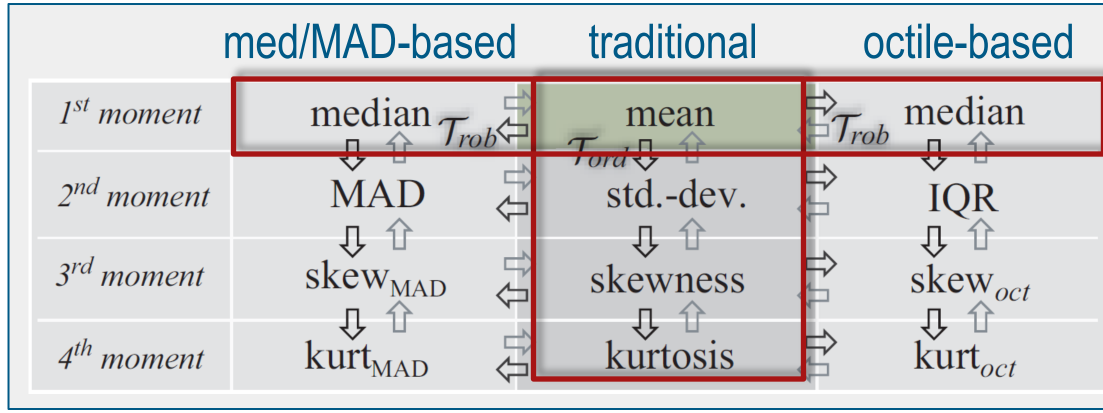
quantile plot
(focus+context)



Iterative View Transformations

Change axis/attribute configuration of view

- **change order** of moment
- **robustify moment**

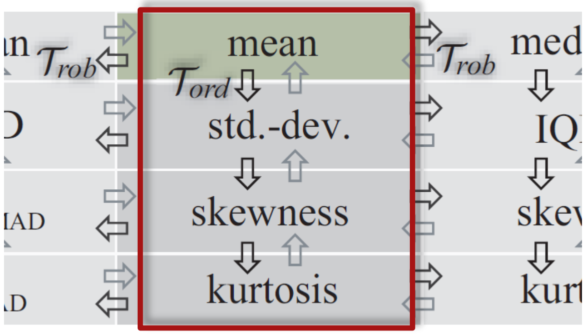


- **compute relation**
  (e.g., difference or ratio)
- **change scale**
  (e.g., normalize, z-standardization)

Closer related to data tranformations

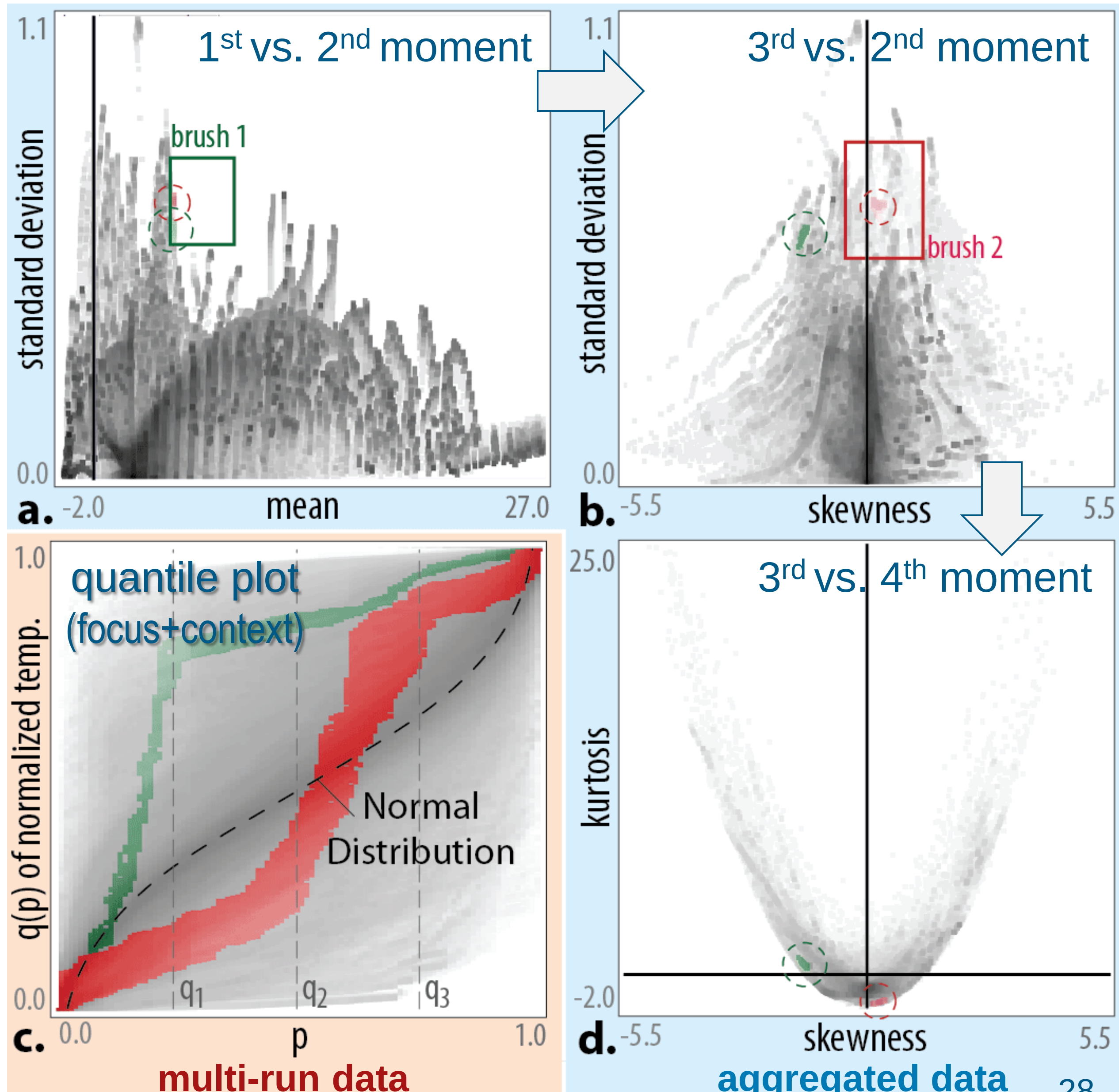
Basic View Setup: Opposing Different Moments

change order of moment



→ study relations betw. moments

→ investigate basic characteristics of distributions

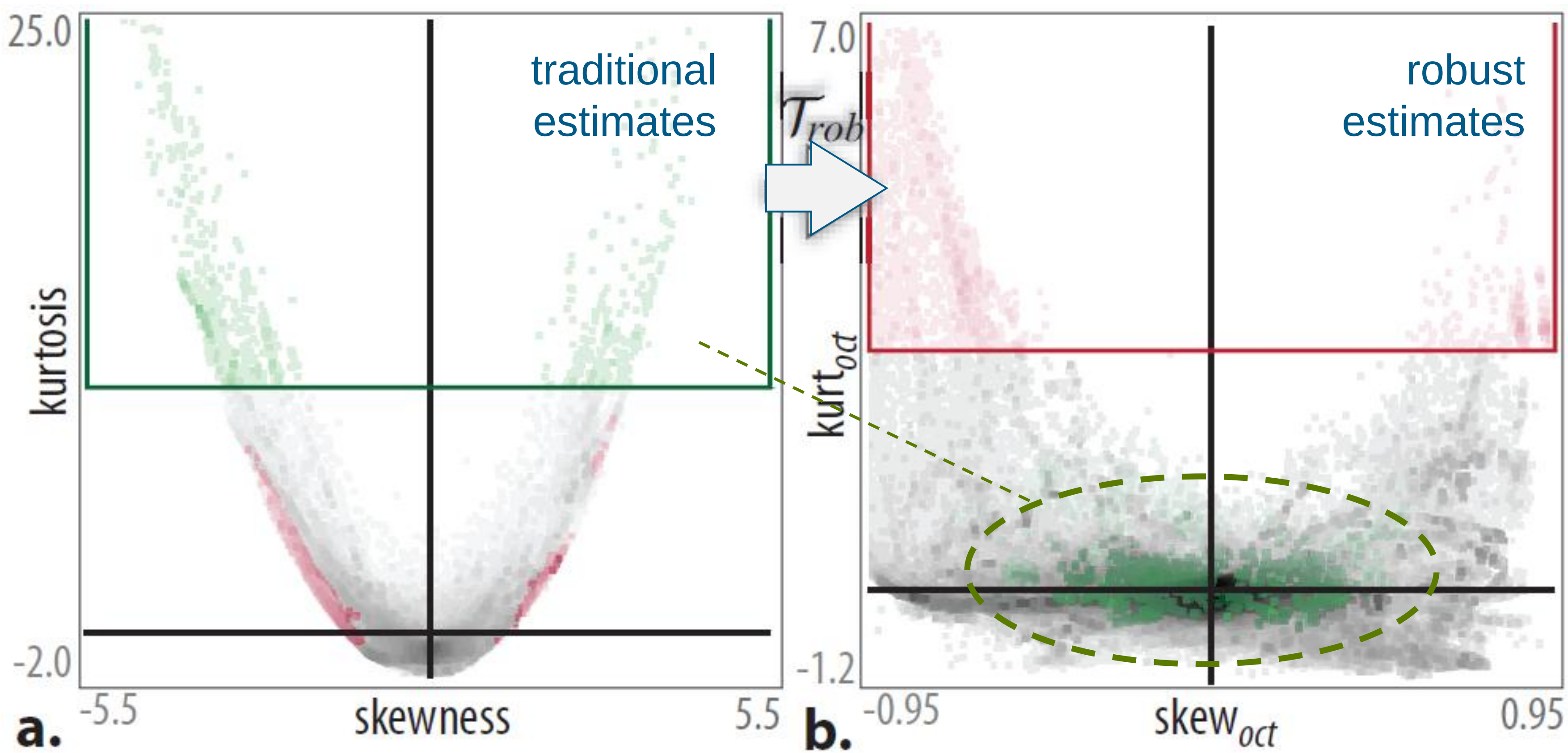


Views: Opposing Different Moments

robustify moment

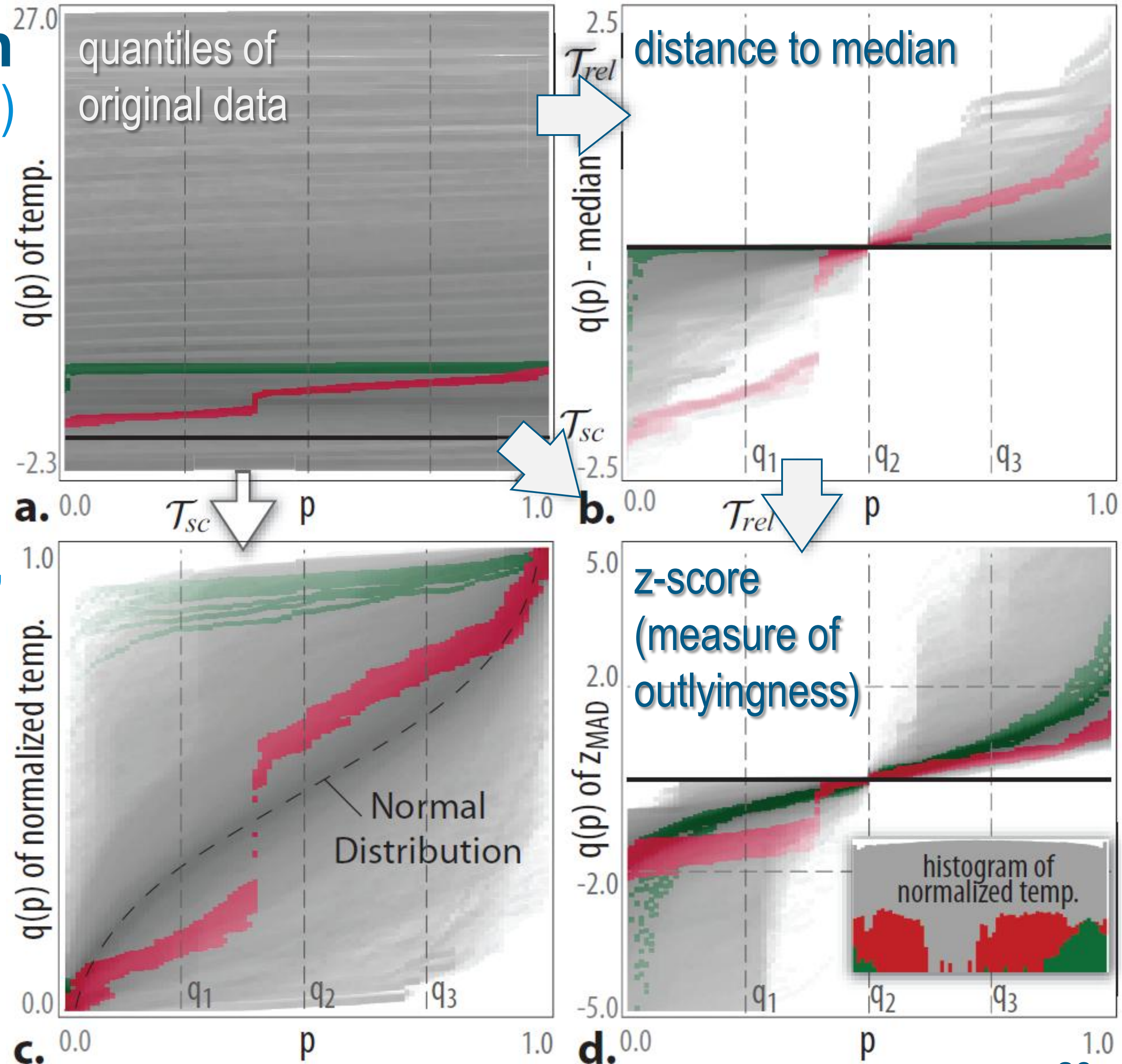
→ assess influence of outliers

1^{st} *moment*	median	mean	median
2^{nd} *moment*	MAD	std.-dev.	IQR
3^{rd} *moment*	skew_{MAD}	skewness	skew_{oct}
4^{th} *moment*	kurt_{MAD}	kurtosis	kurt_{oct}



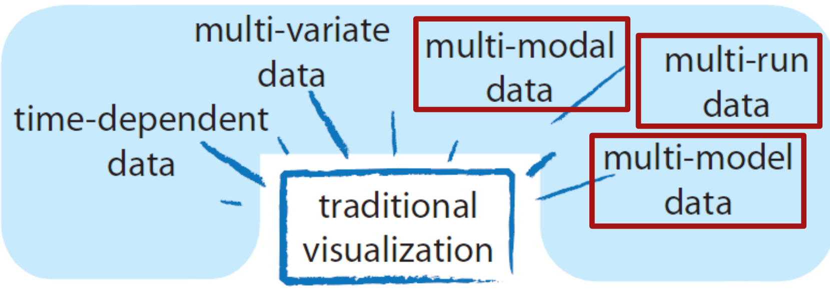
Other View Transformations

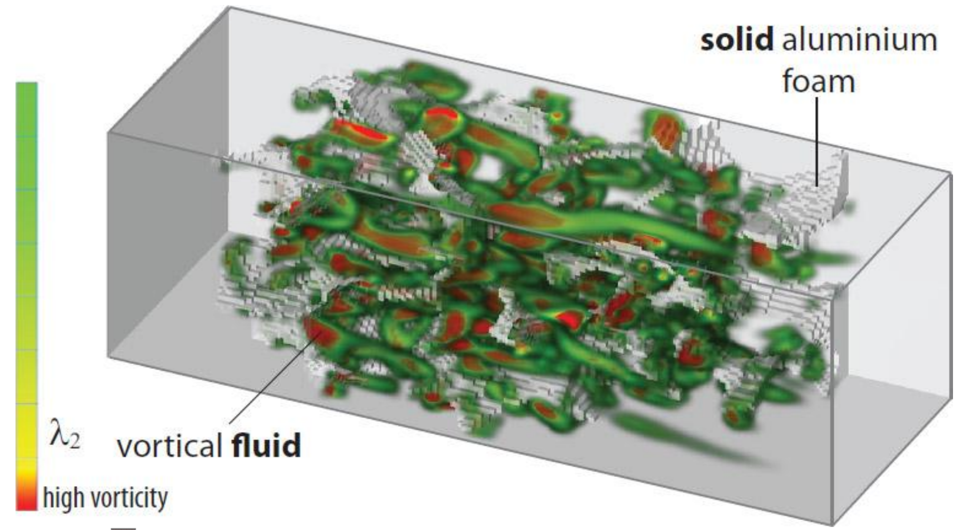
- **compute relation**
  (e.g., difference or ratio)
- **change scale**
  (e.g., z-standardization, normalize to [0,1])

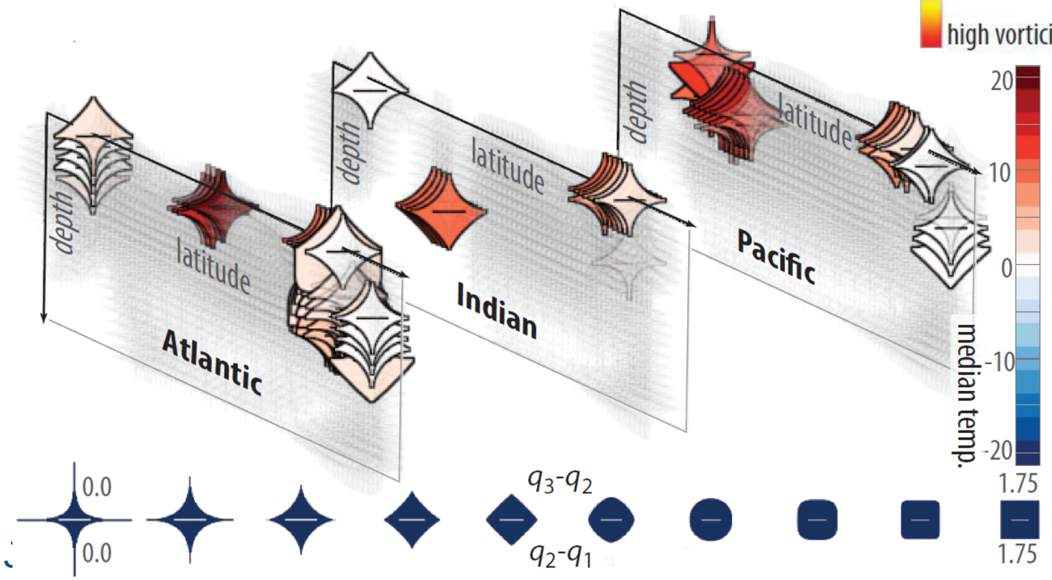


IVA across two Parts of Scientific Data

J. Kehrer, P. Muigg, H. Doleisch, and H. Hauser. **Interactive visual analysis of heterogeneous scientific data across an interface**. *IEEE TVCG, 17(7):934–946, 2011.*

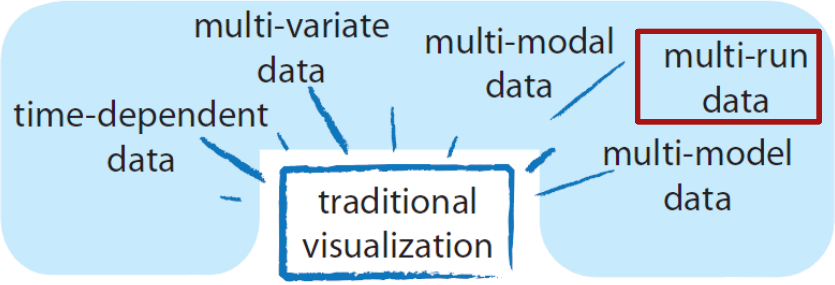


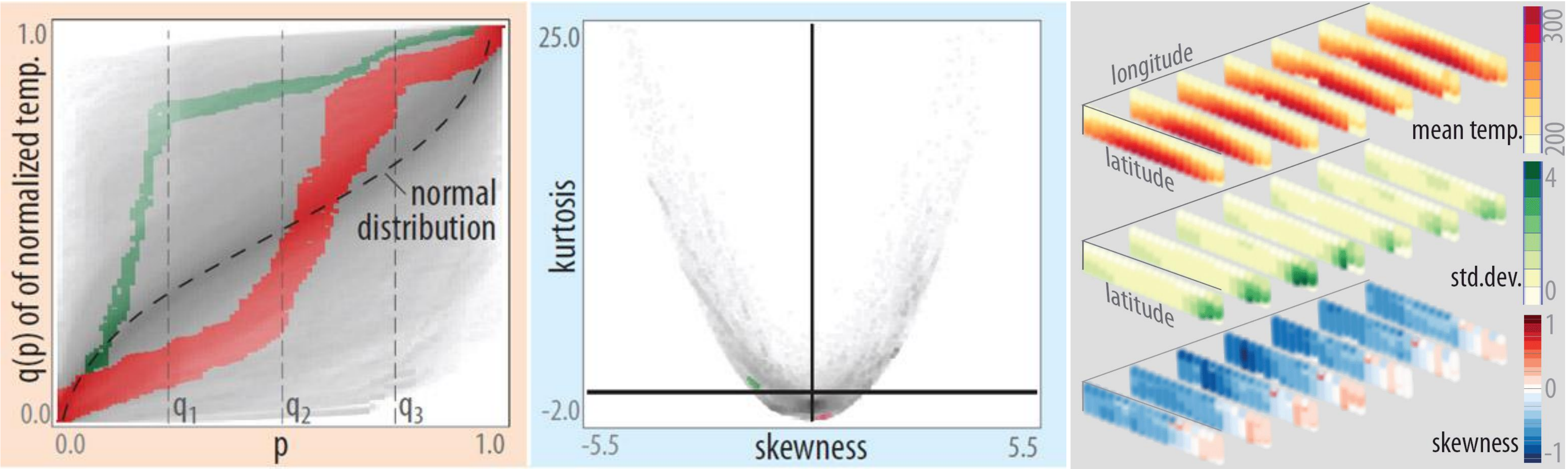




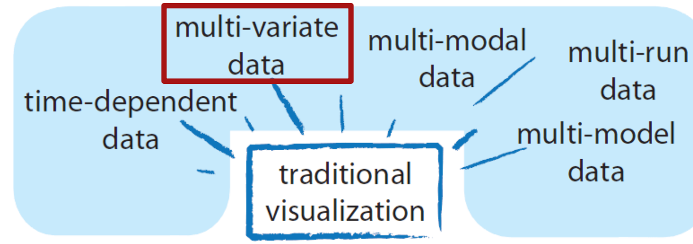
Moment-based Visual Analysis

J. Kehrer, P. Filzmoser, and H. Hauser. **Brushing moments in interactive visual analysis.** *CGF, 29(3):813–822, 2010.*



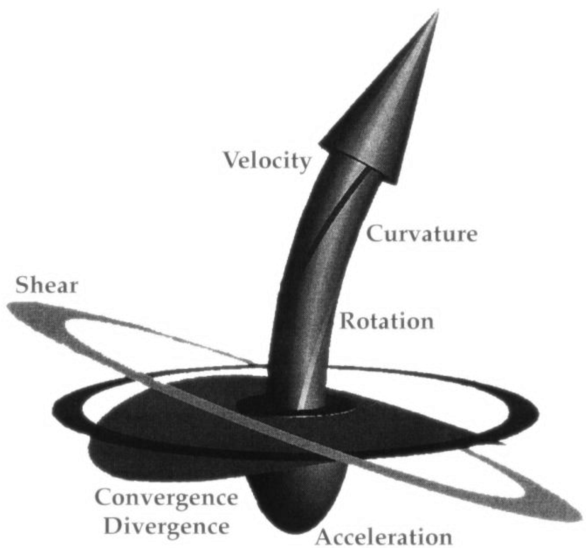


Design aspects of glyph-based 3D vis.

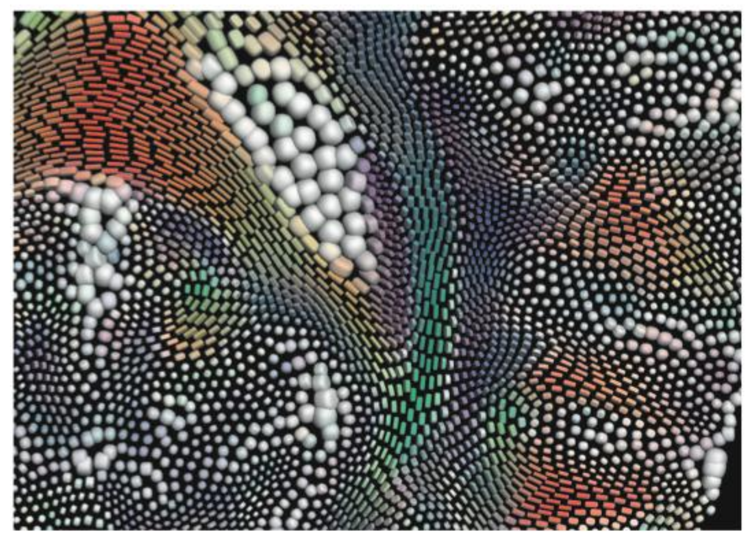


Glyphs

- Map data variate → visual property
  (e.g., color, size, shape, orientation, curvature)
- "Just" combining different visual properties is not enough

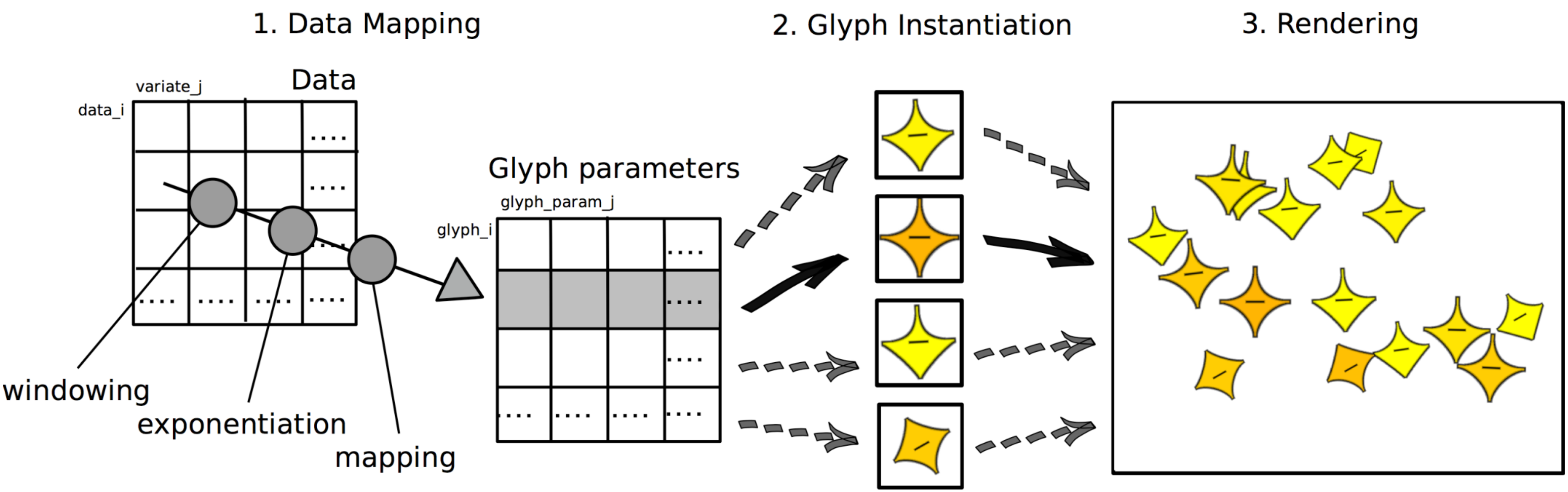


[De Leeuw and van Wijk 1993]



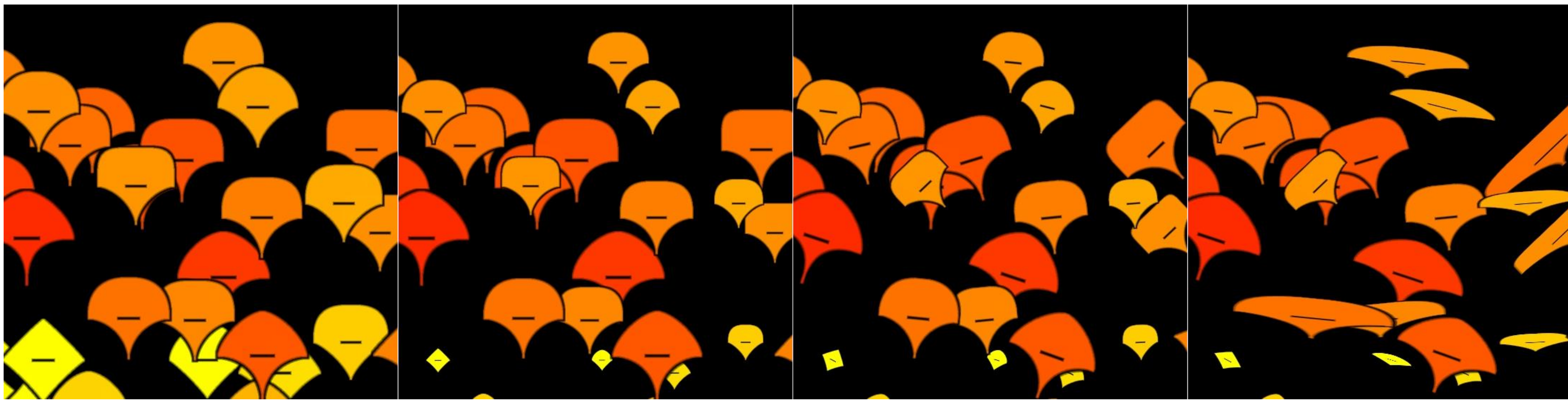
[Kindlmann and Westin 06]

Glyph Pipeline



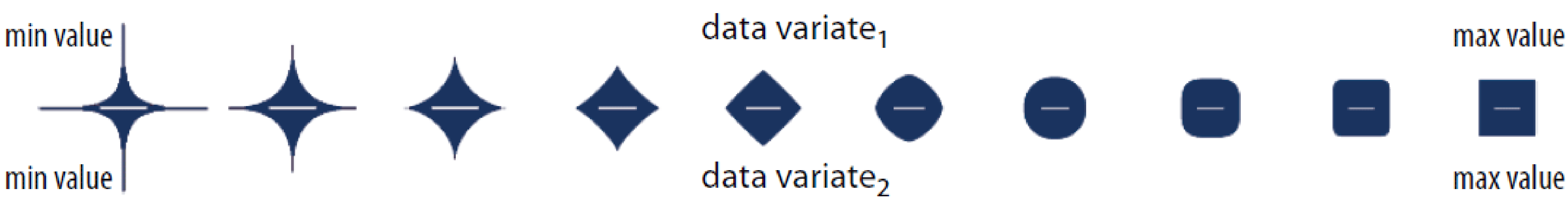
Glyph Instantiation

- Glyph orthogonality (perceive each property individually)



upper/lower shape +size +rotation +aspect ratio

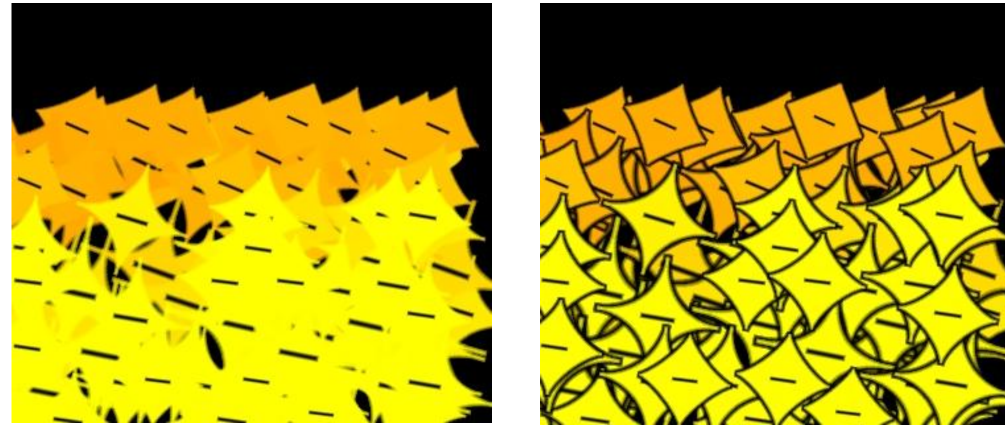
- Glyph normalization (e.g., size)



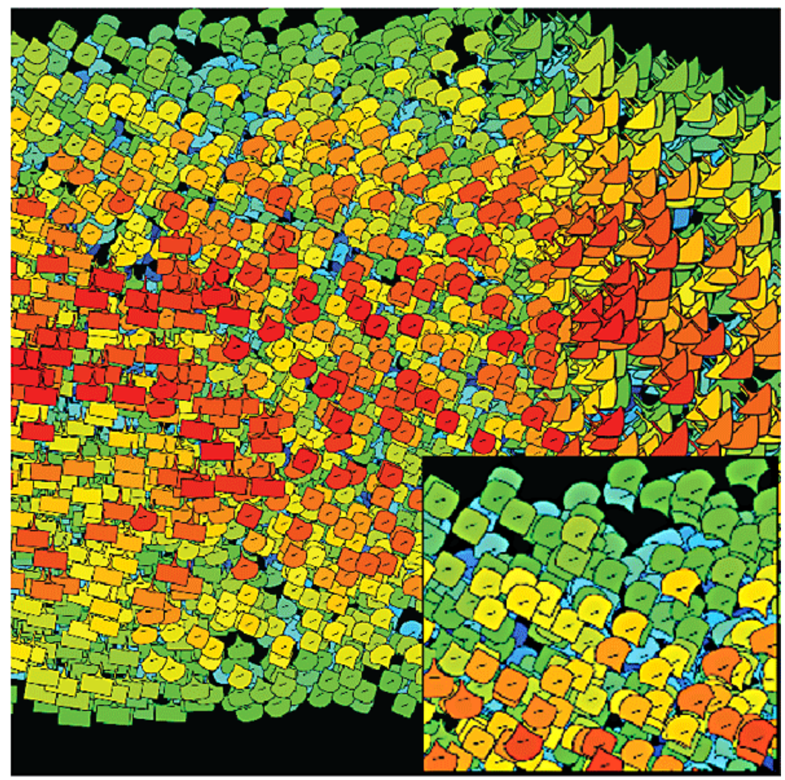
Rendering

Enhance depth perception

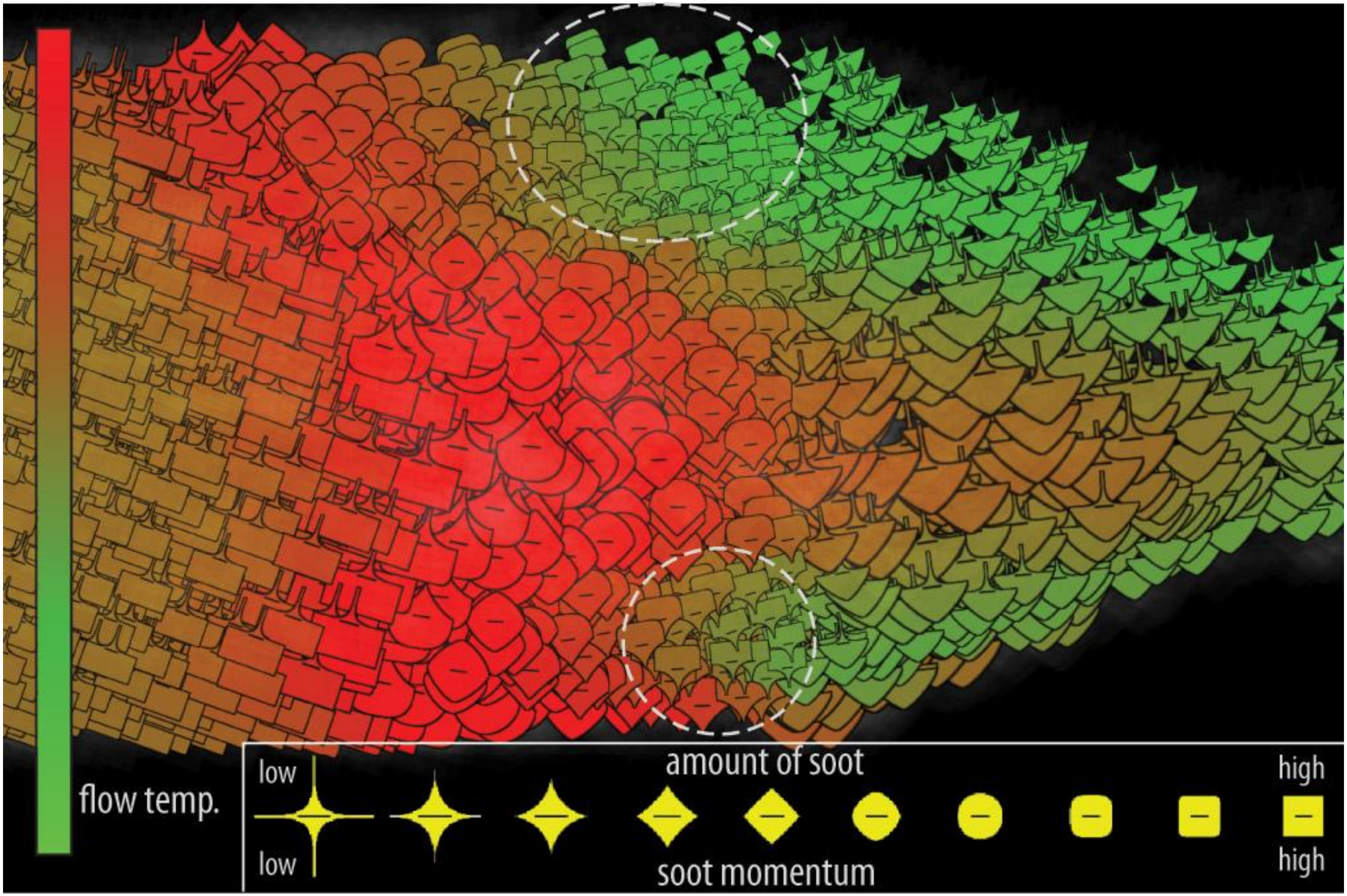
- halos/contours



- chroma depth



Diesel Particulate Filter

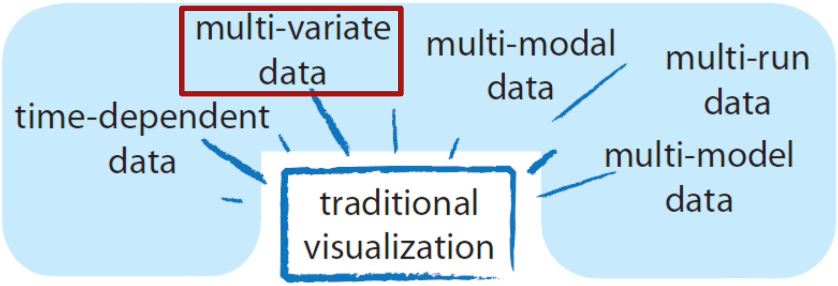


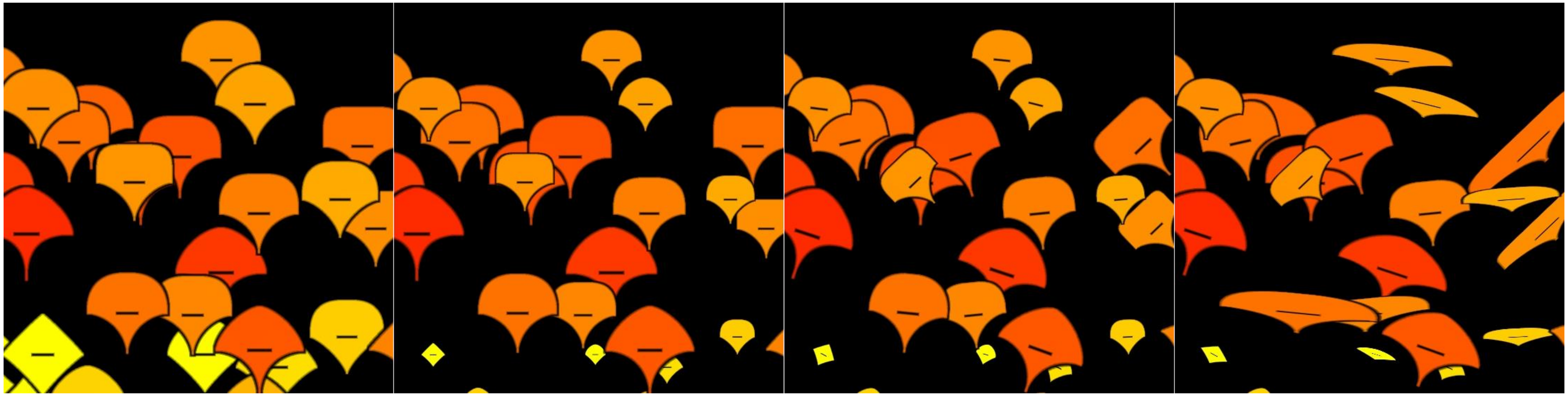
Size & color: flow temp.

Glyph rotation (-45°, 45°): O_2 fraction

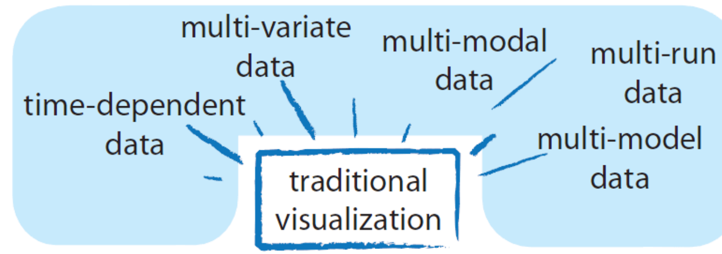
Glyph-based 3D Visualization

A. Lie, J. Kehrer, and H. Hauser. **Critical design and realization aspects of glyph-based 3D data visualization**. In *Proc. Spring Conference on Computer Graphics (SCCG 2009), pages 27–34, 2009.*





Conclusions



- Study of **multi-faceted data**
- IVA across **2 data parts**
  - relating multi-run data ⇔ aggregated statistics
  - analyst can work with both parts (e.g., check validity)
- Integration of **statistical moments**
  - traditional vs. robust statistics, outliers
  - iterative view transformations
  - interactive statistical plots (linking & brushing)
- Workflow for **hypothesis generation**
- Design considerations for **glyph-based 3D vis.**

Acknowledgements

- Helwig Hauser, VisGroup @ UiB
- Helmut Doleisch, Philipp Muigg, Wolfgang Freiler
- Florian Ladstädter, Andrea Steiner, Bettina Lackner, Barbara Pirscher, Gottfried Kirchengast
- Peter Filzmoser, Andreas Lie, Ove Daae Lampe
- Thomas Nocke, Michael Flechsig
- Armin Pobitzer, C. Turkay, Stian Eikeland
- many others

SimVis

Wegener Center
www.wegcenter.at

TU WIEN
TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

PIK
Potsdam Institute for Climate Impact Research