Interactive Visual Analysis of Climate Data

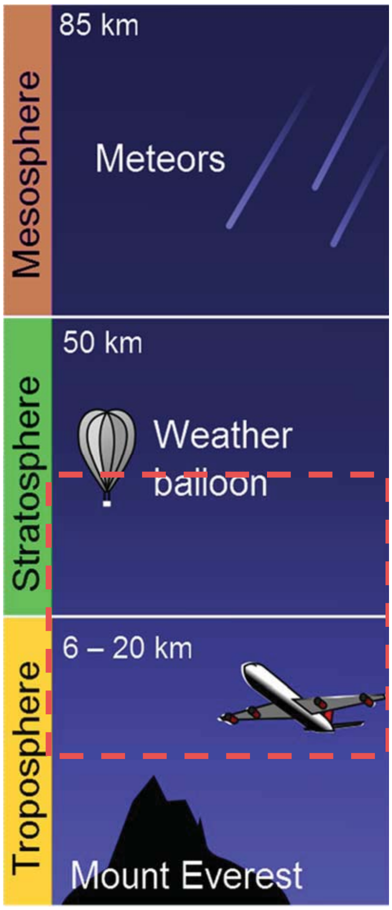
Johannes Kehrer, Helwig Hauser, *et al.*

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

www.ii.UiB.no/vis

Climatological Background

- Investigation and detection of climate change
- Upper troposphere-lower stratosphere
  - known to be sensitive
  - investigate key climate parameters
- Hypothesis generation
  - identify potential sensitive & robust **indicator regions** for climate change (e.g., certain height layers, latitudes)
  - characteristic climate signals, which deviate from natural climate variability
  - useful to monitor atmospheric change



Usual Workflow

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

large-cycle iterations

Challenging to come up with new hypotheses

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

Climate Data and Challenges

- Data sources
  → improved measurements & extensive simulations
- Challenges
  - large, multi-variate data
  - time-dependent
  - deficiencies within data
- Difficult to analyze / understand
  - usually statistical methods used
  - require prior knowledge
  - difficult to find “right” parameter settings

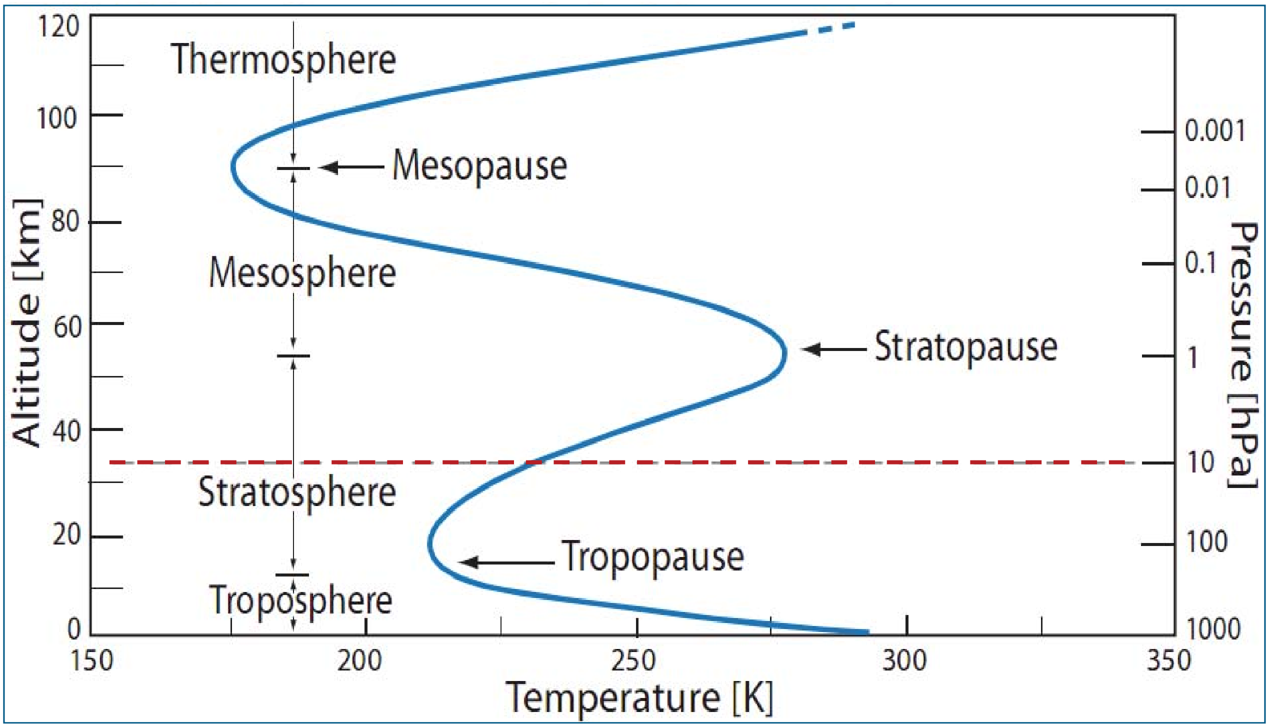
Data used in our Study

Climate Simulation Data

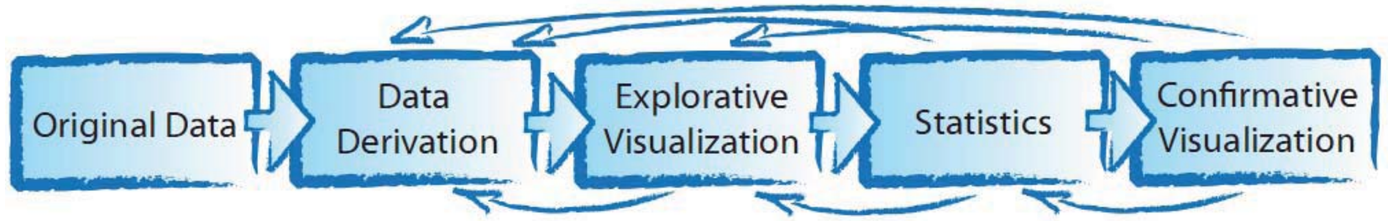
- ECHAM5 climate model, A2 scenario [MPI-M Hamburg] (IPCC 4th assessment report)
- temperature, years 1961–2061
- IPCC 20th century run before 2001
- 180.000 simulation cells
  → 2.5º x 2.5º, 18 pressure levels
- 108 time steps

Data used in our Study

Climate Simulation Data



Our Visual Exploration Process

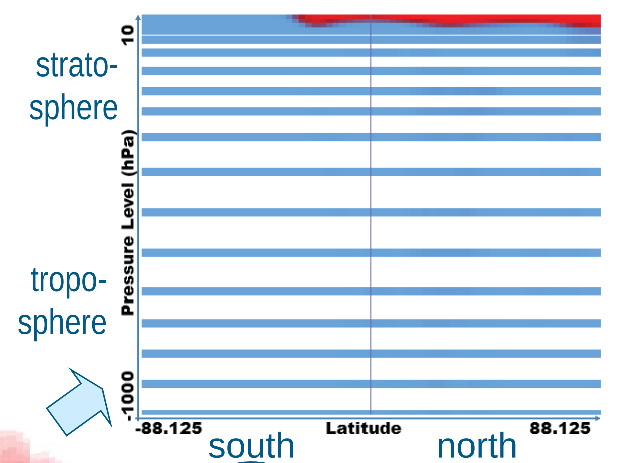


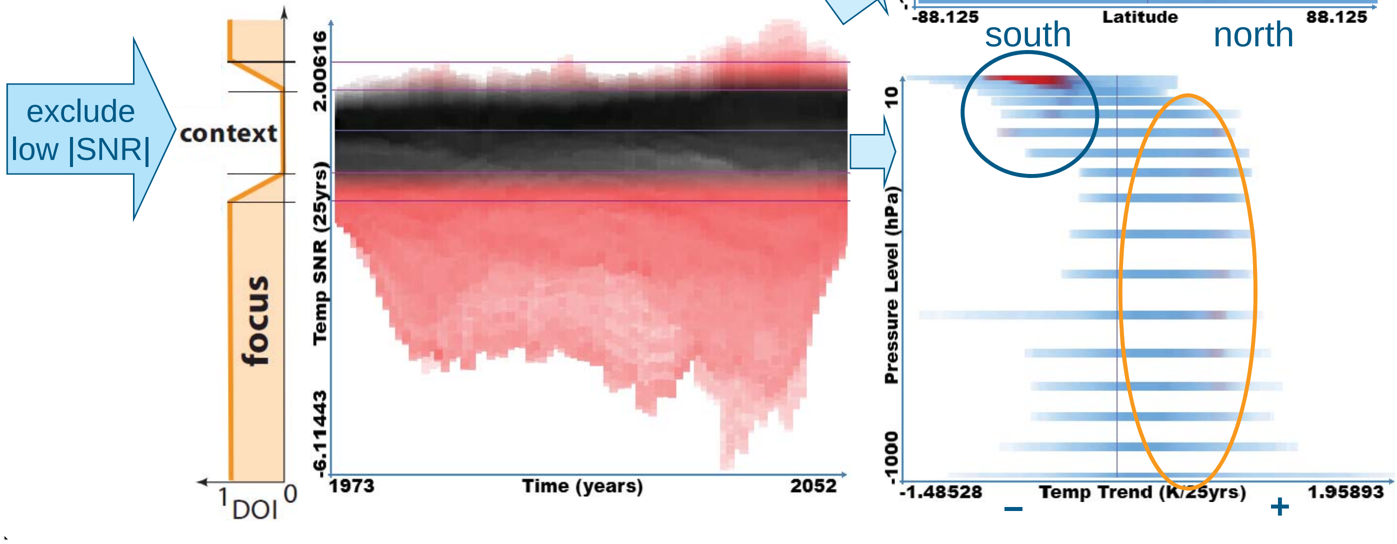
- Integrated **data derivation**
  → linear trends & signal to noise ratios (SNR)
- Interactive **visual exploration** for quick and flexible data investigation ("preview on statistics")
- Generated hypotheses evaluated using **statistics**
  → trend testing [Lackner et al. 08]
- Narrow down **parameters**

Focus on Expressive Data

Localize robust indicators

- areas with high significance (based on trends & SNR)
- smooth specification





Data Derivation: Linear Trends & SNR

[Ladstädter et al. 08]

- Smooth data $\tilde{y}$: moving average over N years
- Linear trend: $$trend_i = \frac{1}{N}(\tilde{y}_{i+N/2} - \tilde{y}_{i-N/2})$$
- Linear trend fit curve: $$fit_{i,j} = \tilde{y}_{i-N/2} + [j - (i - N/2)] \cdot trend_i$$
- Detrended standard deviation: $$s_i = \left[\frac{1}{N-1} \sum_{j=i-N/2}^{i+N/2} (y_i - fit_{ij})^2 \right]^{1/2}$$
- Signal-to-noise ratio: $$SNR_i = trend_i / s_i$$

Further Refinement

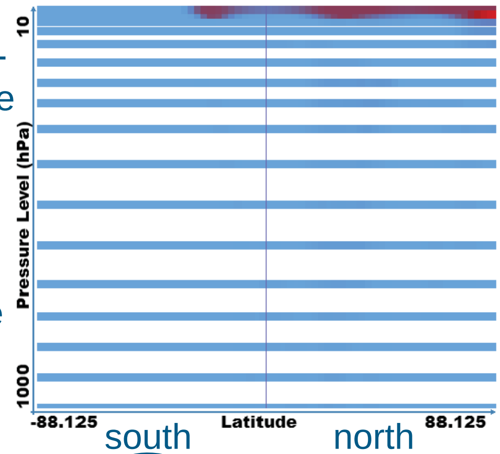
Exclude upper pressure levels

→ known deficiencies

[Cordero & Forster '06]

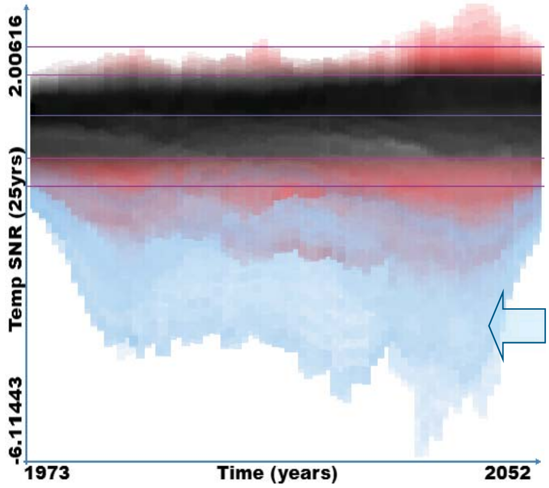
strato-sphere

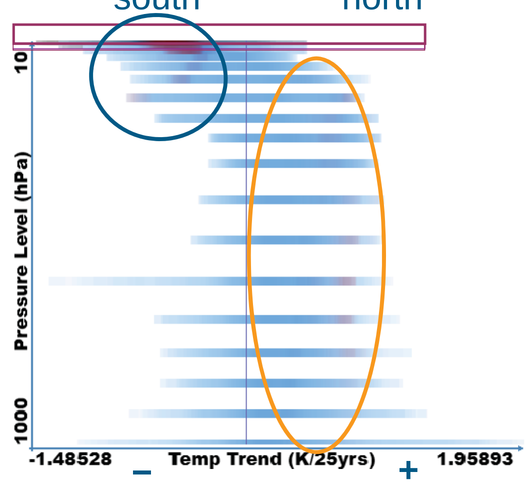
tropo-sphere



south

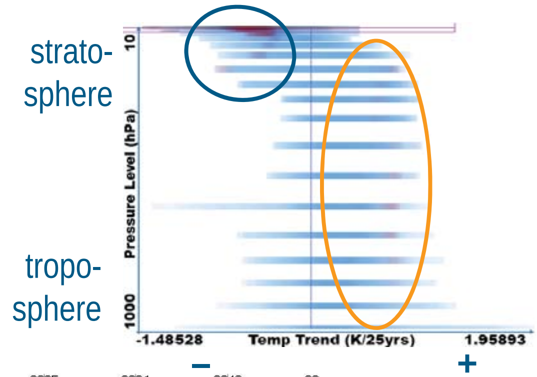
north

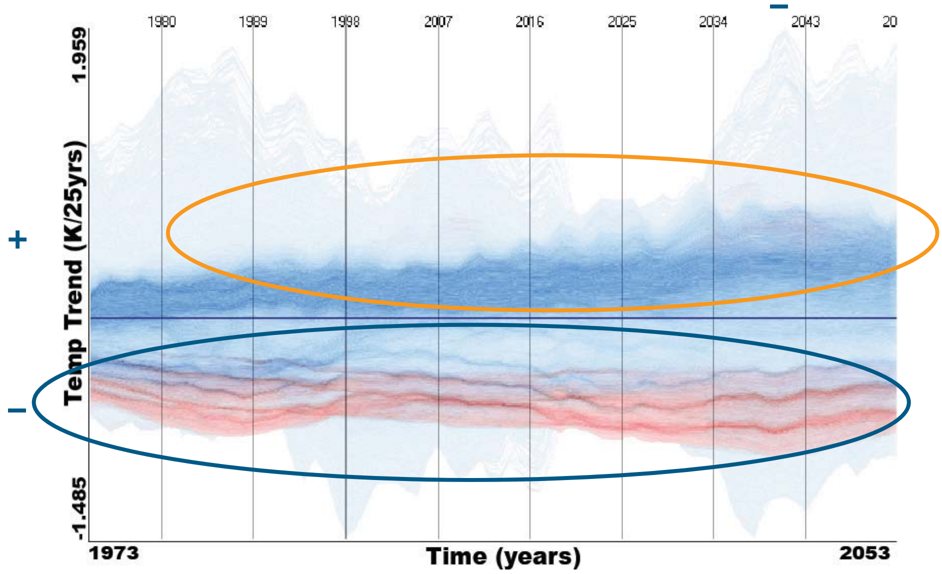




Explore Trend Variation over Time

Robust cooling trends

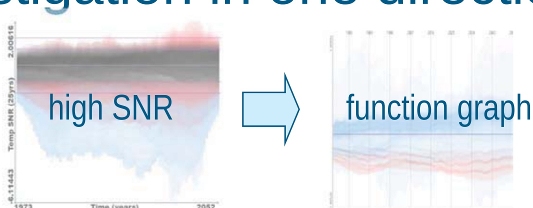




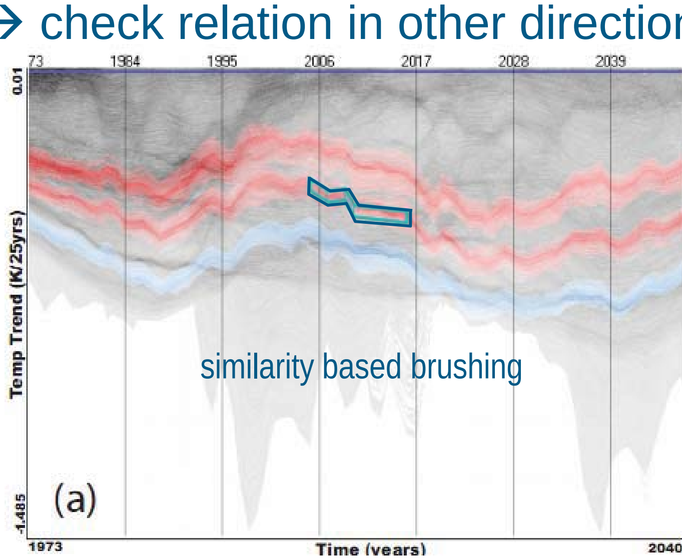
Analyze Relations between Dimensions

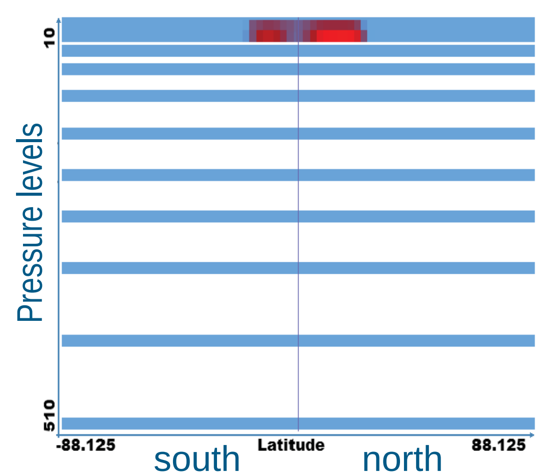
Up to now:

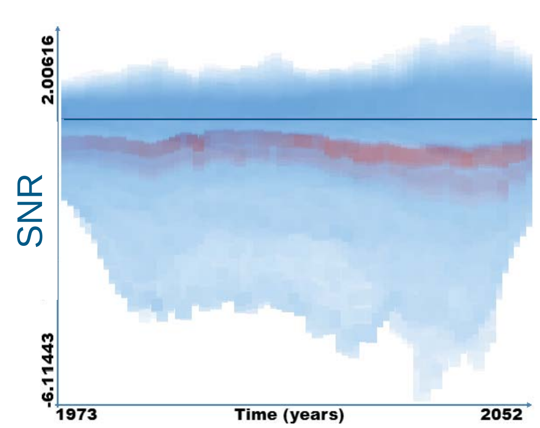
→ investigation in one direction



→ check relation in other direction



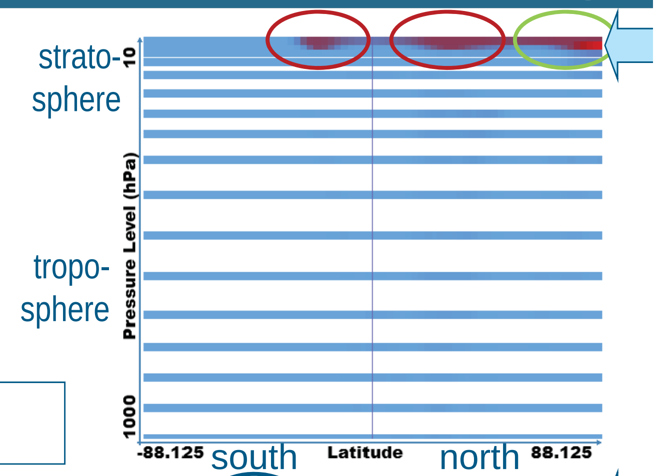


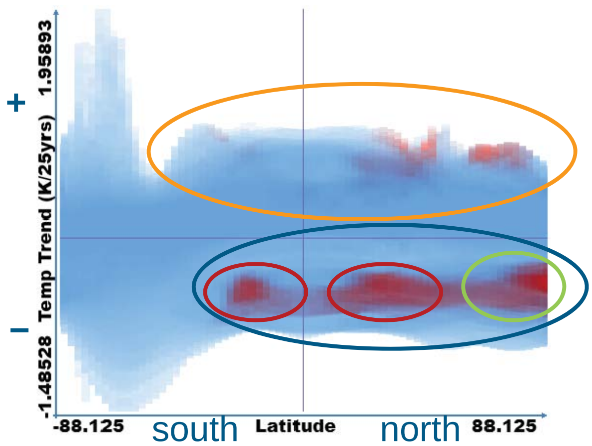


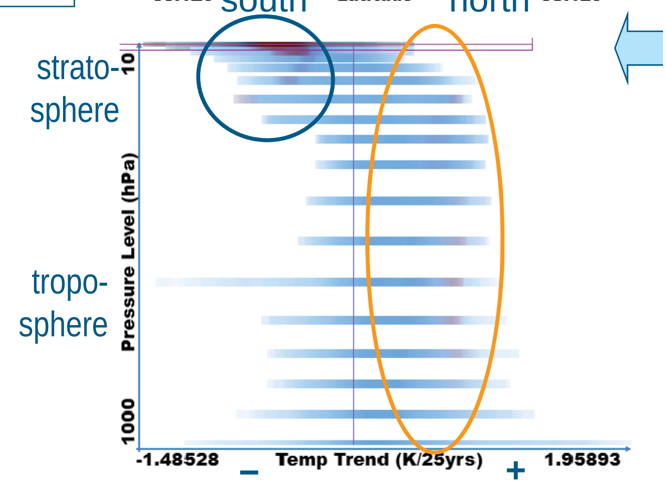
Generated Hypothesis / ECHAM5 temp.

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

hypothesis then to statistics

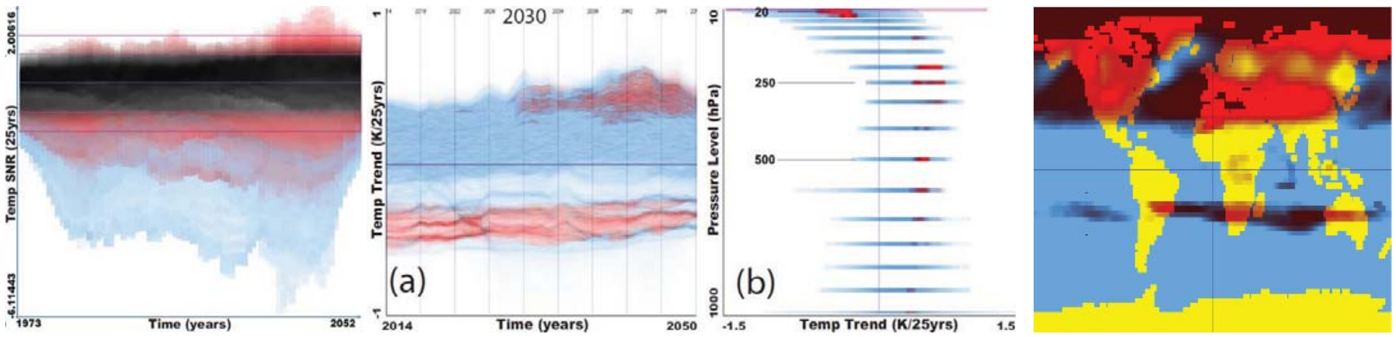






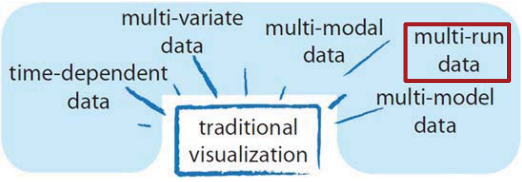
Hypothesis Generation with Visual Exploration

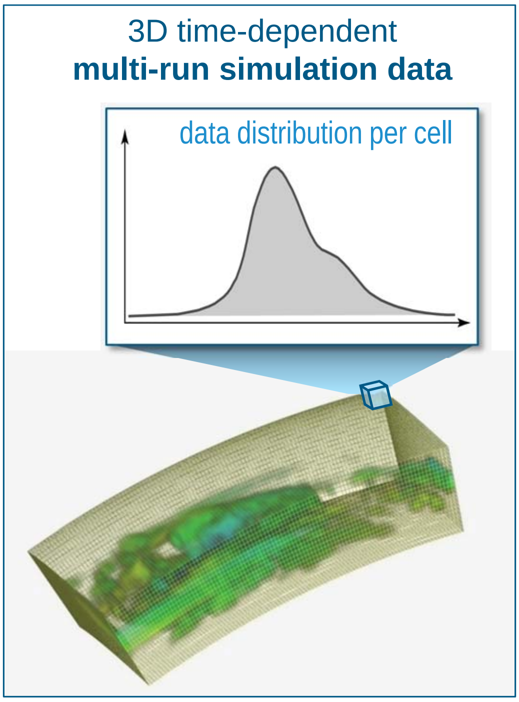
- Kehrer et al. **Hypothesis generation in climate research with interactive visual data exploration.** *IEEE TVCG, 14(6):1579–1586, 2008.*
- Ladstädter et al. **SimVis: an interactive visual field exploration tool applied to climate research**. In *New Horizons in Occultation Research,* pages 235–245.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



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)

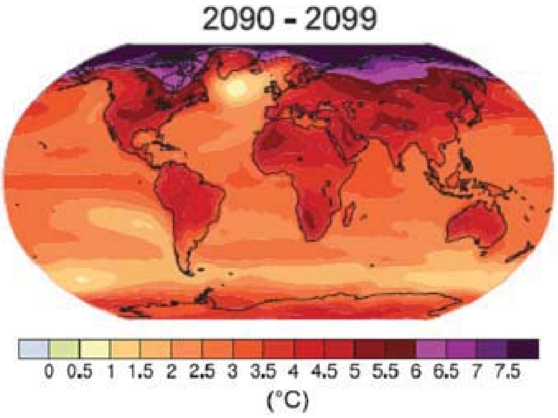




Reducing the Data Dimensionality

- **Statistics:** assess distributional characteristics along an independent dimension (e.g., time, spatial axes)
- Integrate into IVA through **attribute derivation**

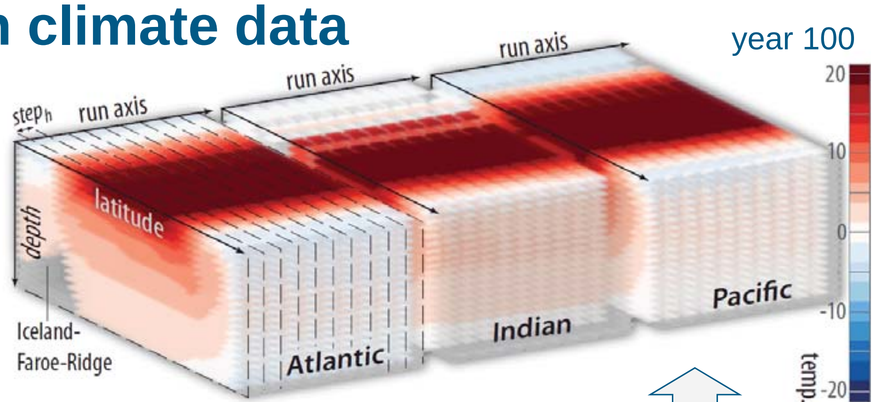
[from IPCC AR #4, 2007]



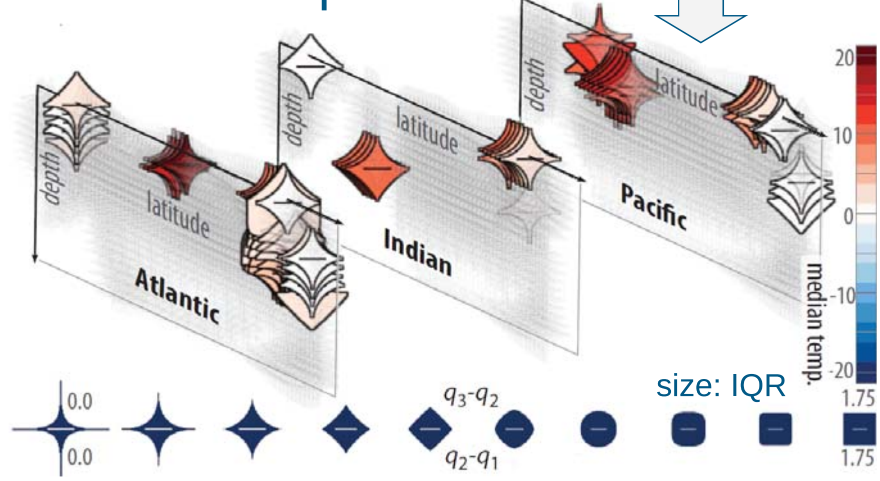
average temp. in ten years

Integrating Statistics and IVA

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



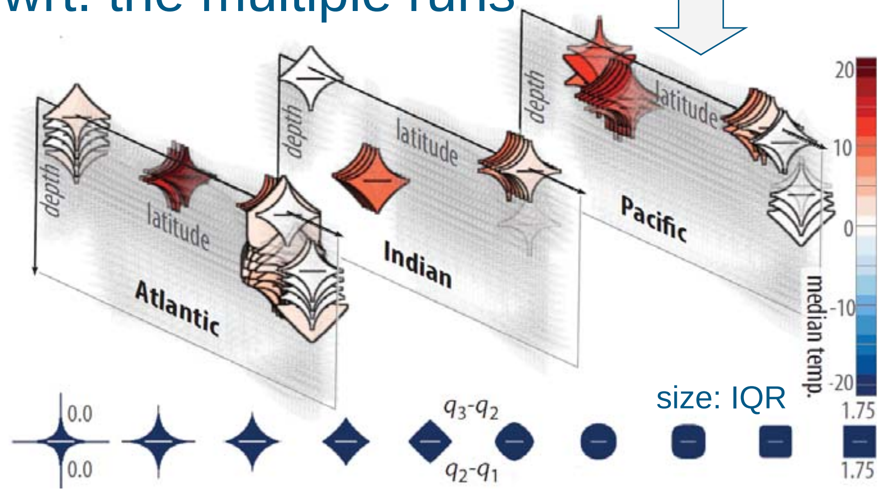
- Compute statistics wrt. the multiple runs



Integrating Statistics and IVA

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

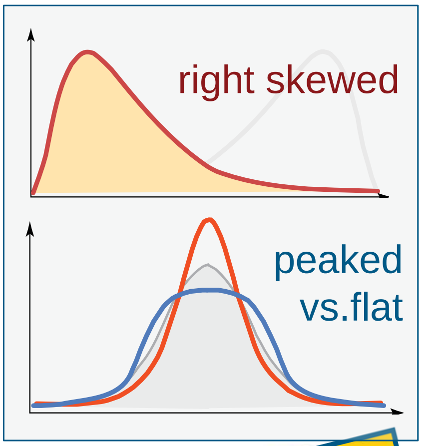
- Compute statistics wrt. the multiple runs



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

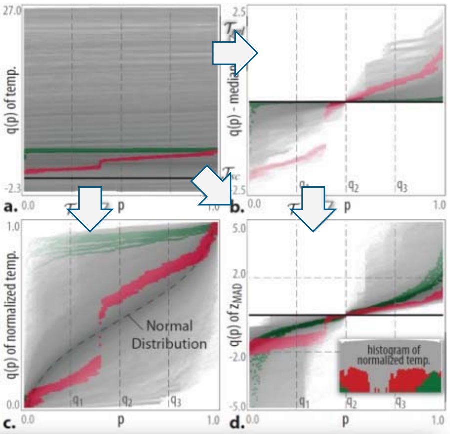


Structured approach to manage complexity

How to deal with this “management challenge”?

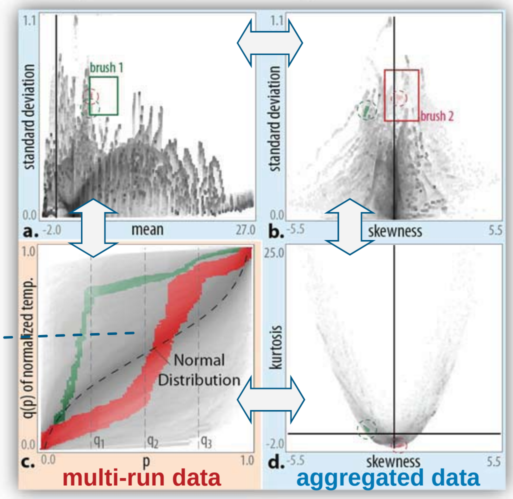
Moment-based Visual Analysis

- 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**

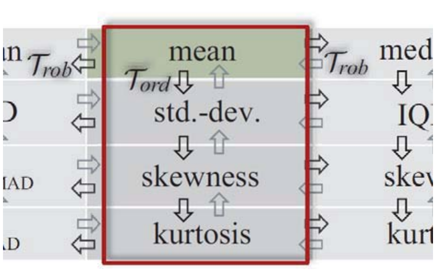
	med/MAD-based		traditional		octile-based
1st moment	median	$\mathcal{T}_{rob}$	mean	$\mathcal{T}_{rob}$	median
2nd moment	MAD		std.-dev.		IQR
3rd moment	skew_{MAD}		skewness		skew_{oct}
4th moment	kurt_{MAD}		kurtosis		kurt_{oct}

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

Closer related to data tranformations

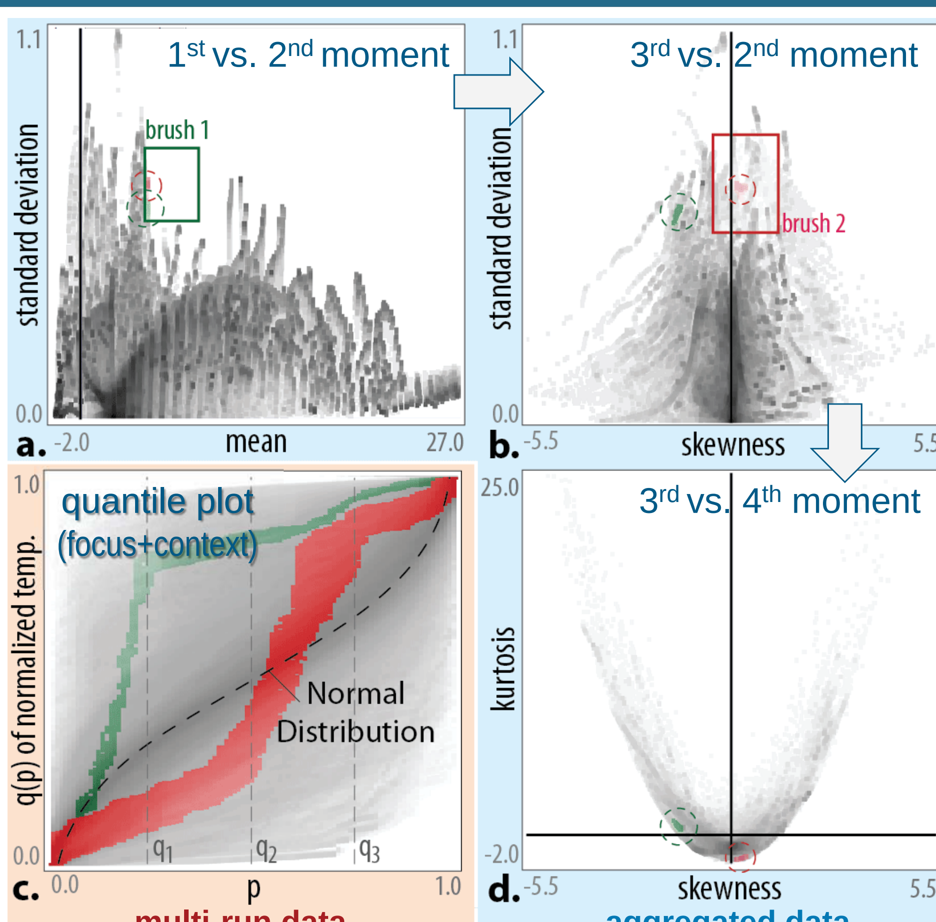
Basic View Setup: Opposing Different Moments

change order of moment



→ study relations betw. moments

→ investigate basic characteristics of distributions

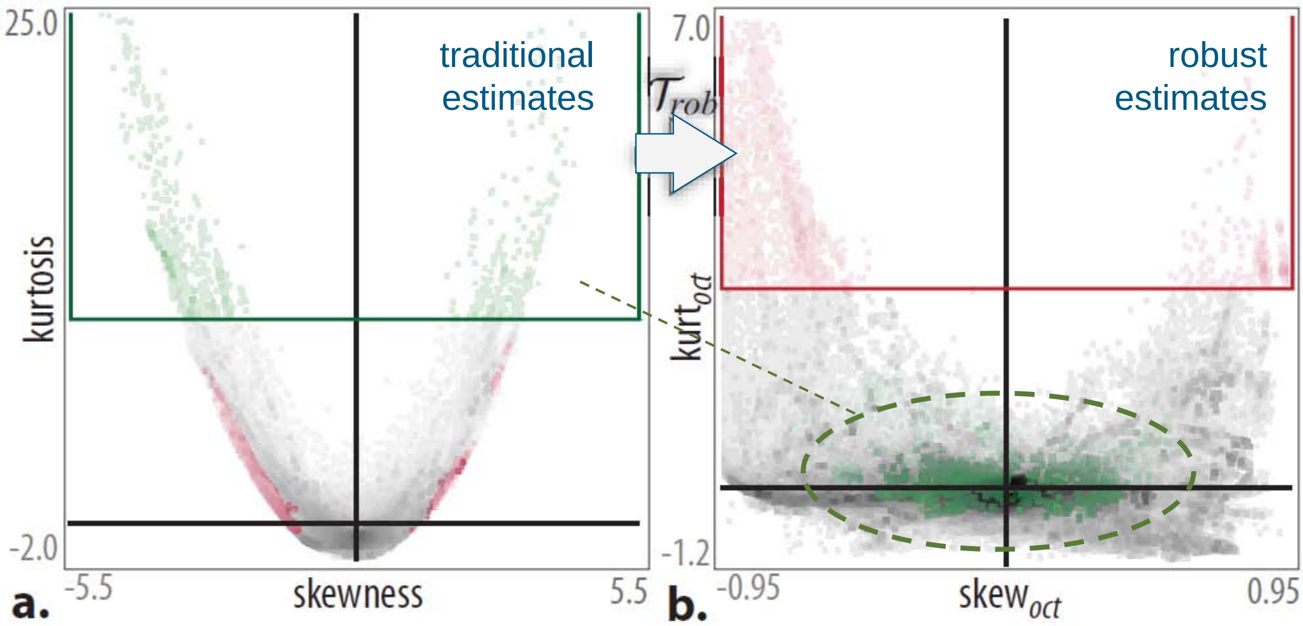


Views: Opposing Different Moments

robustify moment

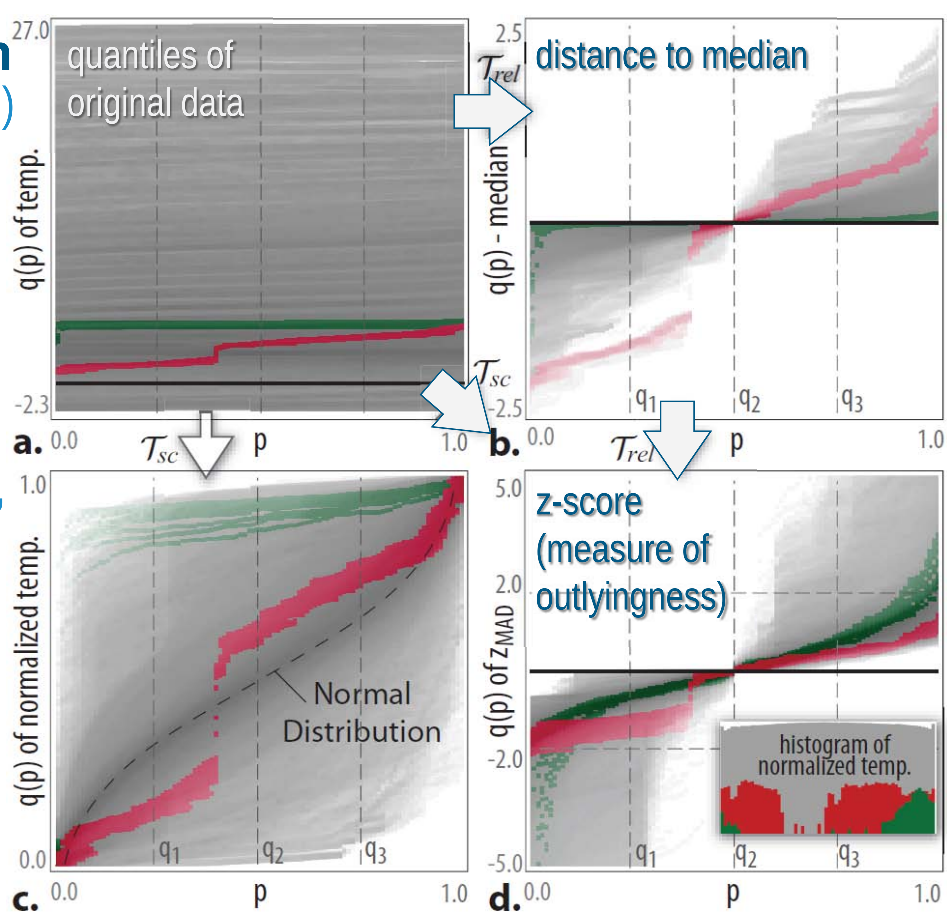
→ assess influence of outliers

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



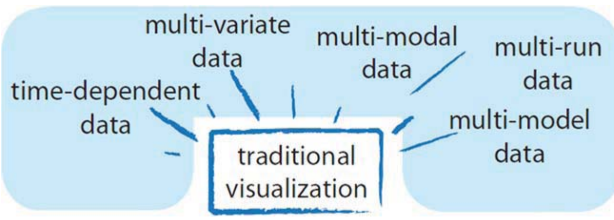
Other View Transformations

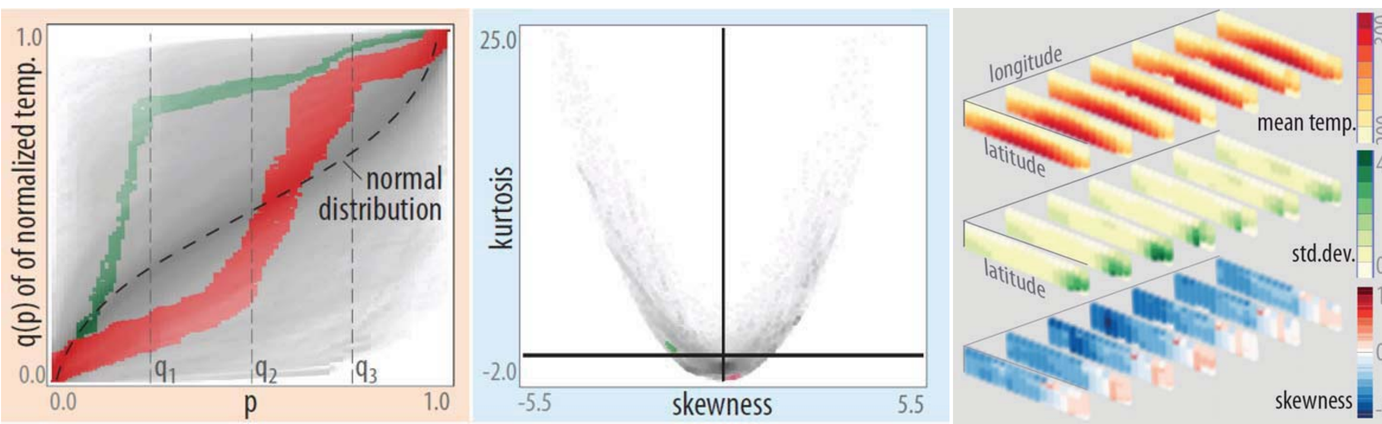
- **compute relation**
  (e.g., difference or ratio)
- **change scale**
  (e.g., z-standardization, normalize to [0,1])



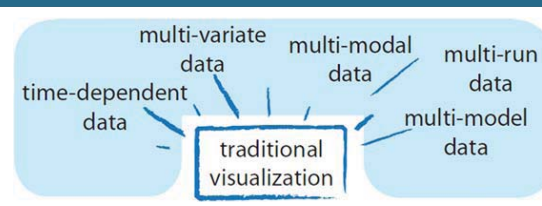
Moment-based Visual Analysis

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





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**

Acknowledgements

- Johannes Kehrer
- 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