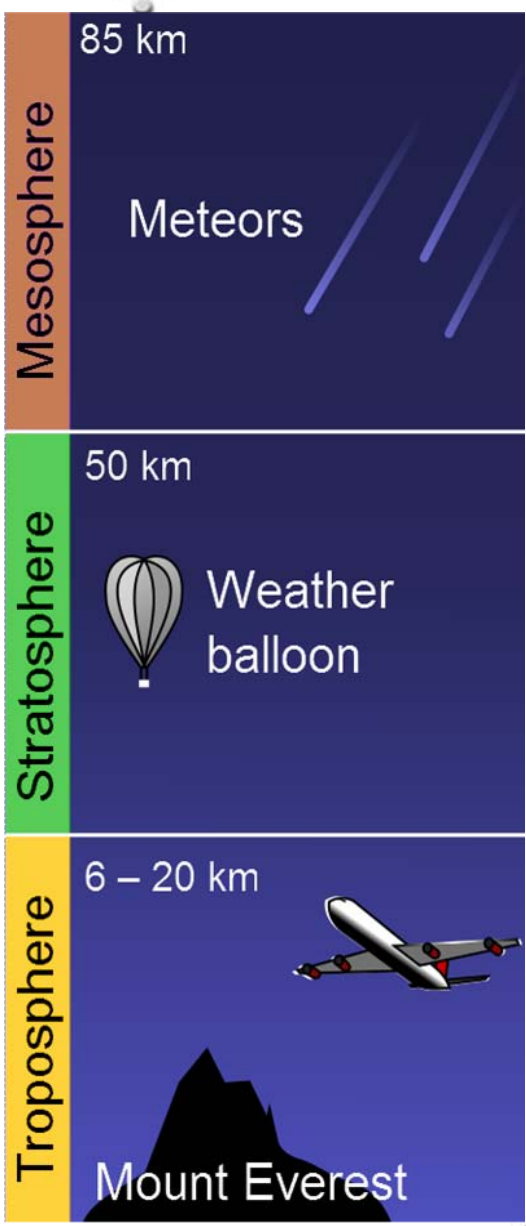
Hypothesis Generation in Climate Research with Interactive Visual Data Exploration

Johannes Kehrer[1], Florian Ladstädter[2], Philipp Muigg[3], Helmut Doleisch[3], Andrea Steiner[2], Helwig Hauser[1]

[1] Department of Informatics, University of Bergen, Bergen, Norway
[2] Wegener Center for Climate and Global Change, Graz, Austria
[3] VRVis Research Center and SimVis GmbH, Vienna, Austria

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



Kehrer et al.

Usual Workflow and Goal

- Set research focus
- Acquire data
- Explore / investigate data
- Formulate particular hypothesis
- Evaluate with statistics
- Iterate

large-cycle iterations

Challenging to come up with new hypotheses
→ intuition of expert, scientific trial & error

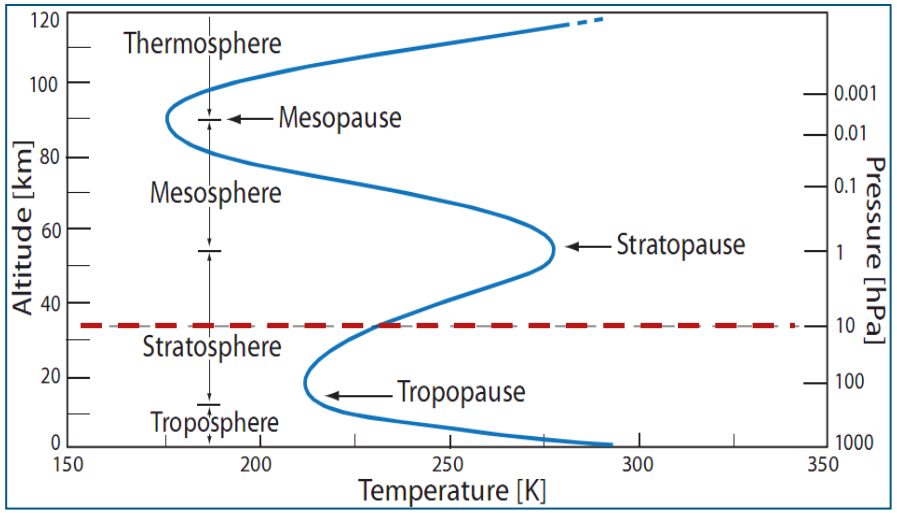
Goal: accelerate process (fast interactive visualization, more informed partner → more directed search)

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

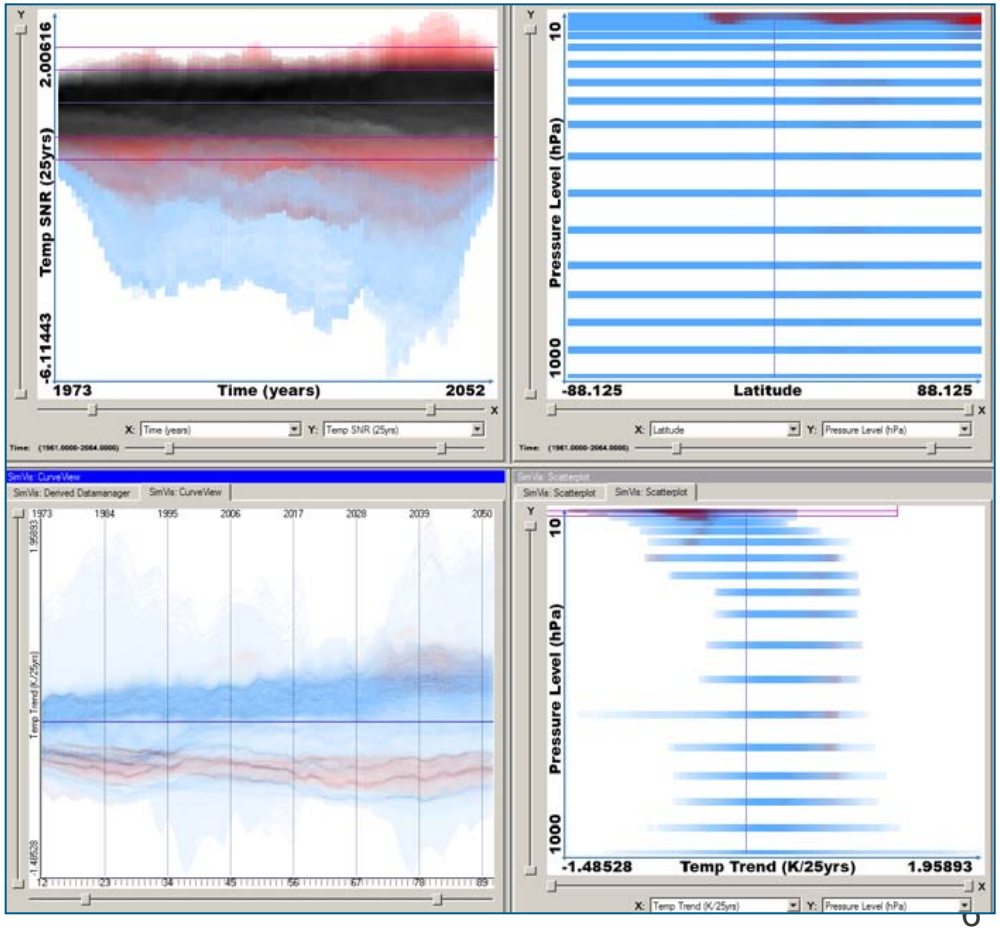


- ECHAM5 climate model, A2 scenario [MPI-M Hamburg] (IPCC 4th assessment report)
- temperature, years 1961–2061
- IPCC 20th century run before 2001
- 180.000 simulation cells
  → 2.5º x 2.5º, 18 pressure levels
- 108 time steps

Interactive Visual Data Exploration

Enables visual dialogue between user and data

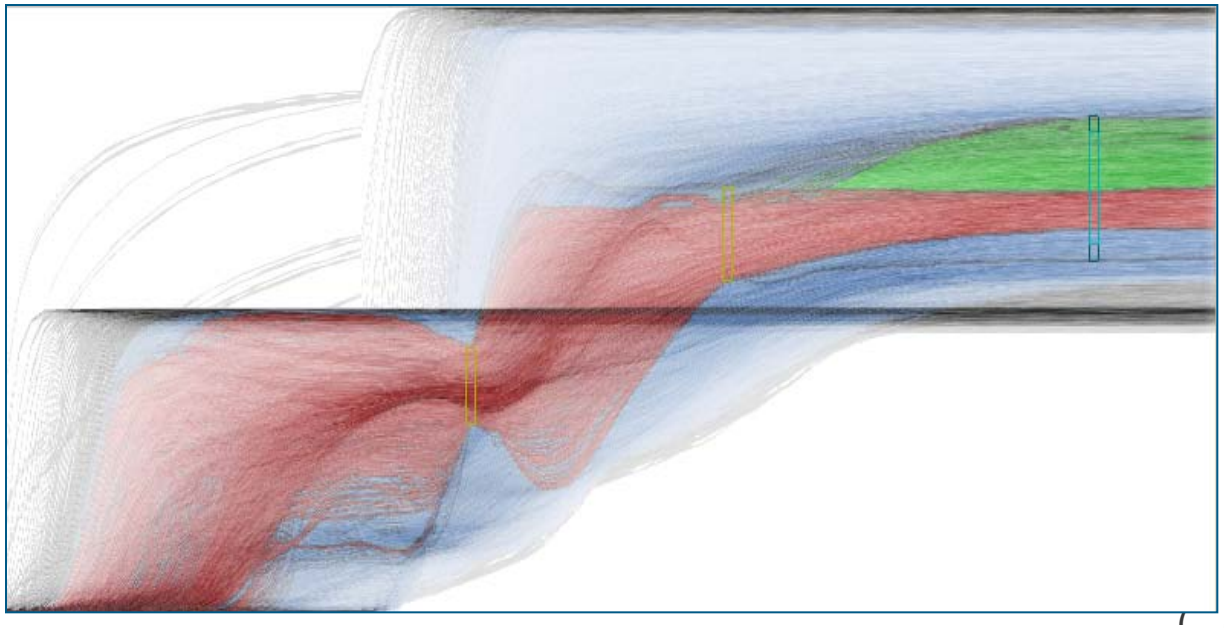
- SimVis: coordinated, multiple views framework (histograms, scatterplots, 3D/4D views, etc.)
- focus+context vis.
- degree-of-interest (DOI) data attribution $\in$ [0, 1]
- hierarchical feature definition language
- on-the-fly data derivation
- interactivity, etc.



Recent Extensions of SimVis

Function Graphs View [Muigg et al. 08]

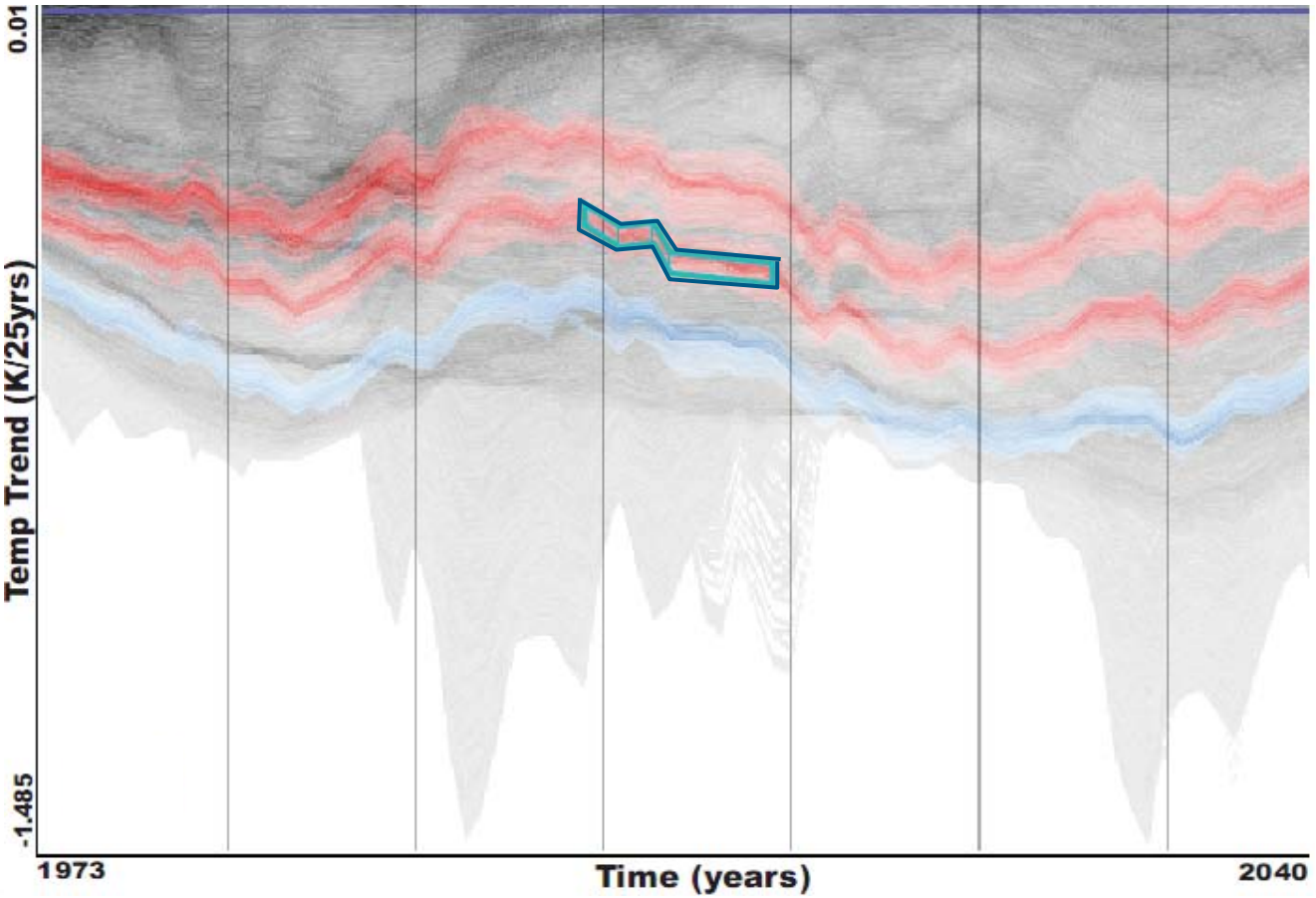
- visualize large amounts of time-dependent data
- focus+context coloring
  → color coding of features specified in different views
- transfer functions [Johansson et al. 05]
  → map line count to pixel‘s luminance
- data aggregation (frequency binmaps) [Novotný & Hauser 06]



Advanced Brushing Techniques

Select function graphs based on **similarity**

- pattern sketched by user
- similarity evaluated on gradients (1st derivative)



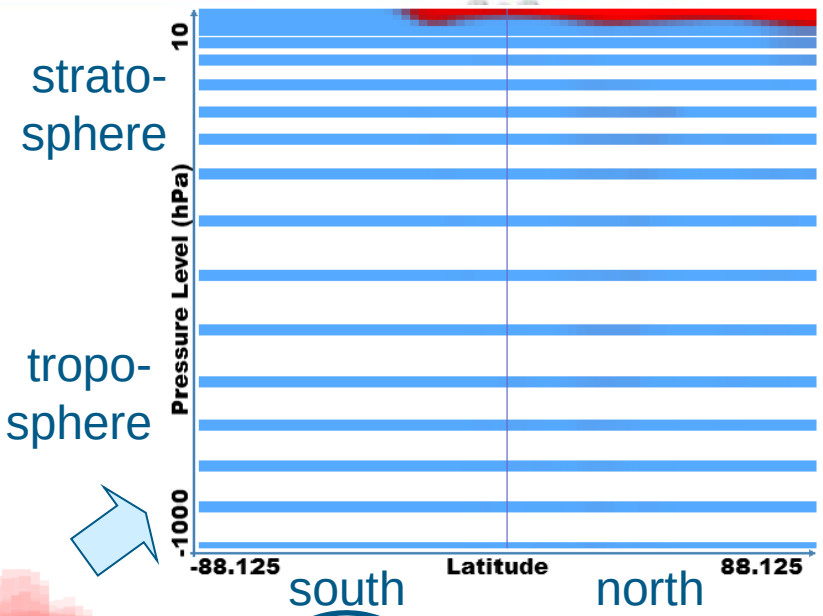
Our Exploration Process

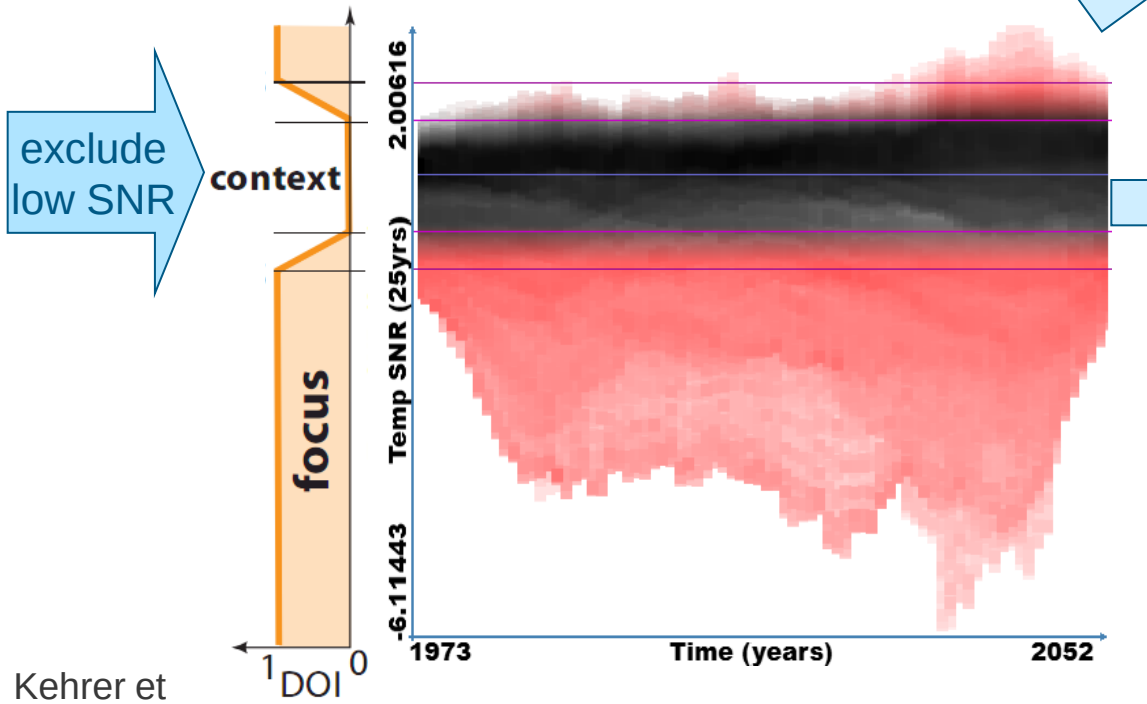
- Interactive visual exploration for quick and flexible data investigation
- Integrated data derivation [Ladstädter et al. 08]
  - linear trends
    → moving differences computed on smoothed data
  - signal to noise ratios (SNR)
    → determine significance
- Generated hypotheses evaluated using statistics
  → trend testing [Lackner et al. 08]

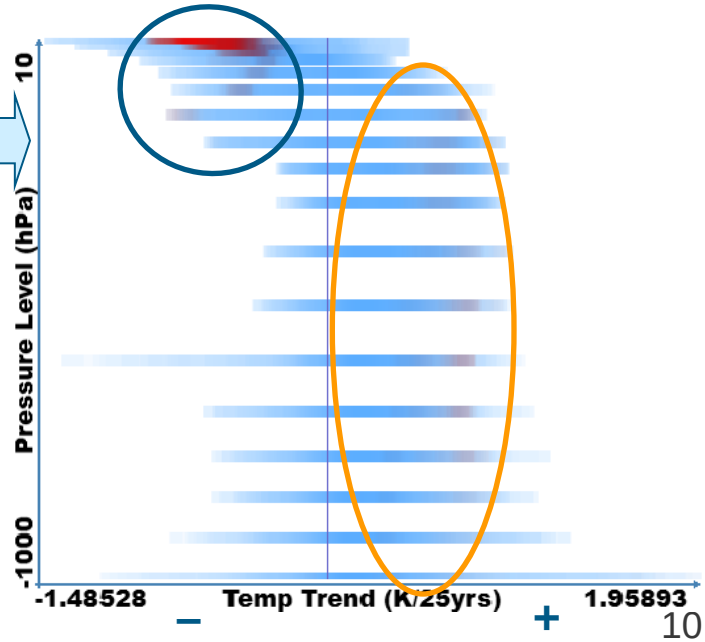
Start: Focus on Expressive Data

Localize robust indicators

- area with high significance
  → exclude low |SNR|
- smooth specification







Further Refinement

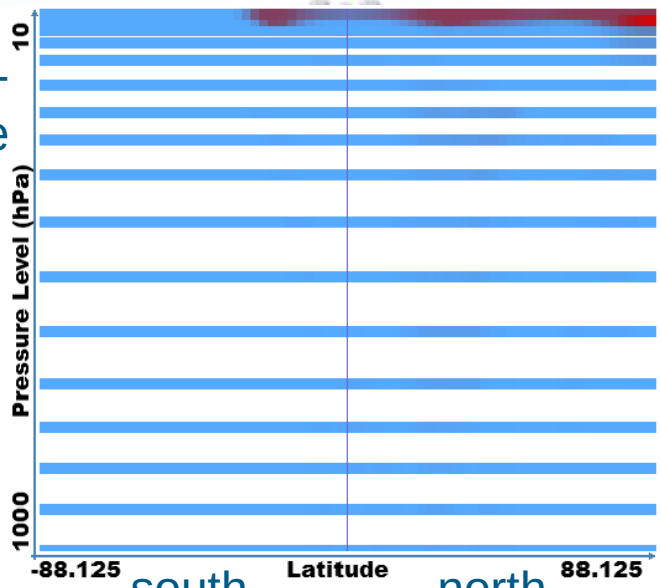
Exclude upper pressure levels

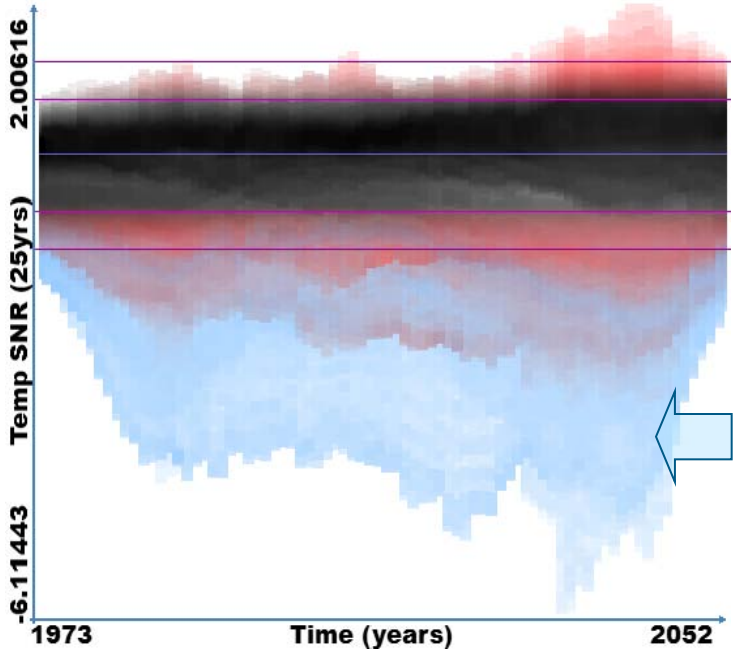
→ known deficiencies

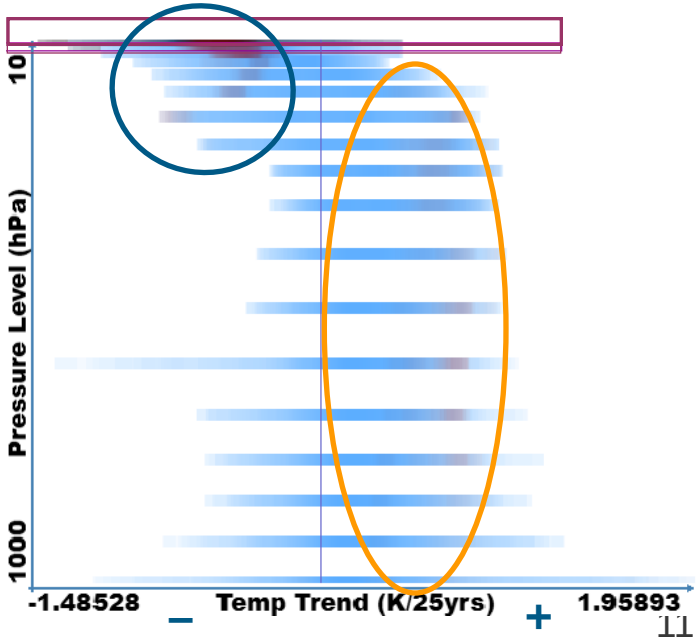
[Cordero & Forster ’06]

strato-sphere

tropo-sphere





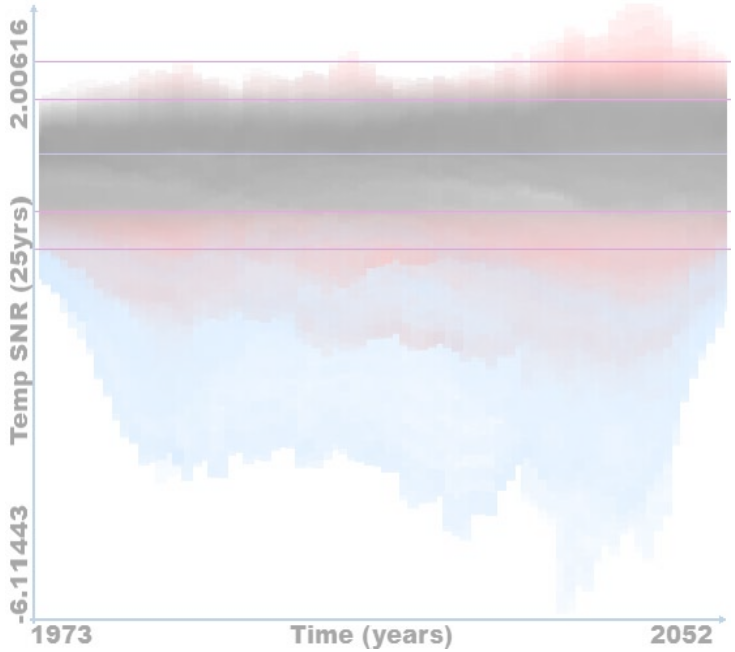


Kehrer et al.

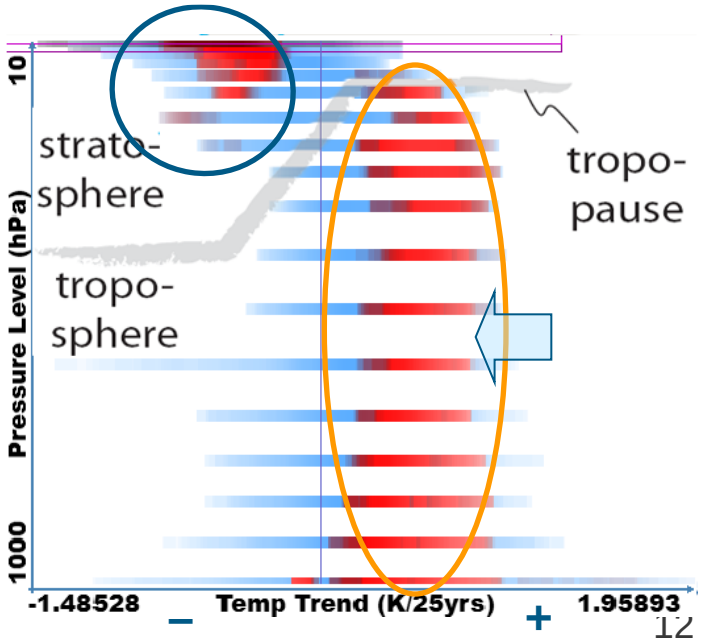
Exploring Indicators

Investigate less robust indicators

→ emphasize feature coloring

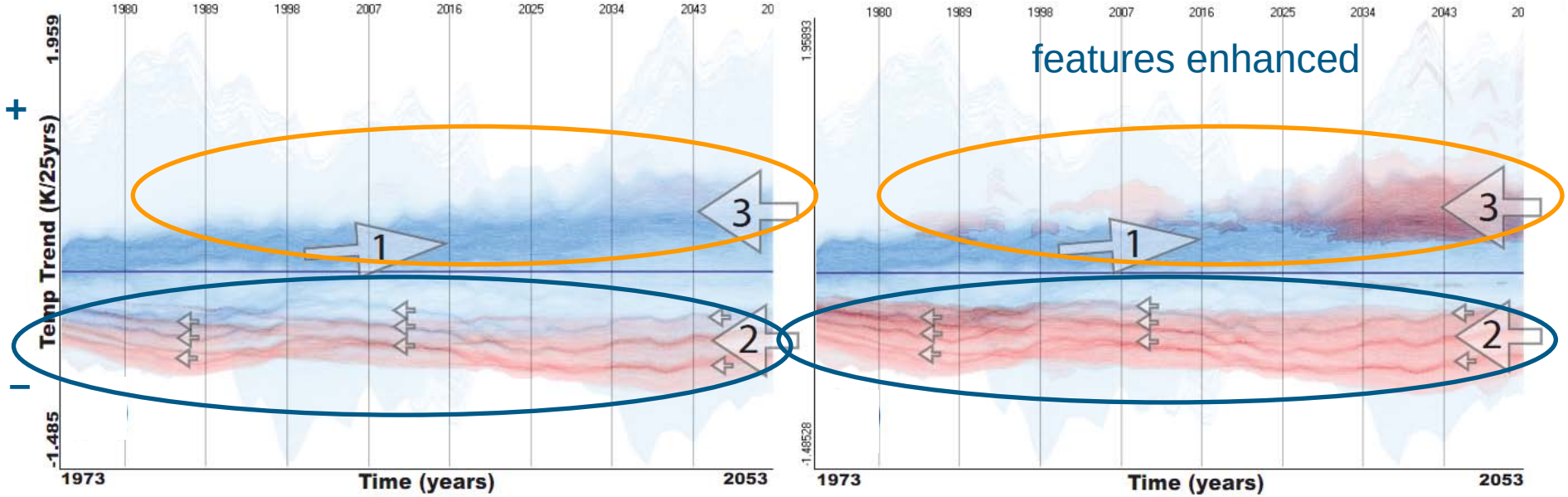


tropo-
sphere



Explore Trend Variation over Time

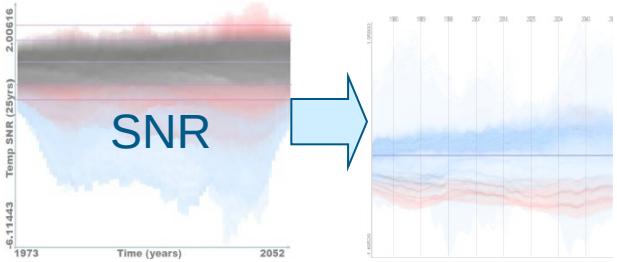
- several highlighted neg. traces (2)
  → high significance over whole investigated time span (robust)
- less robust indicators (3)

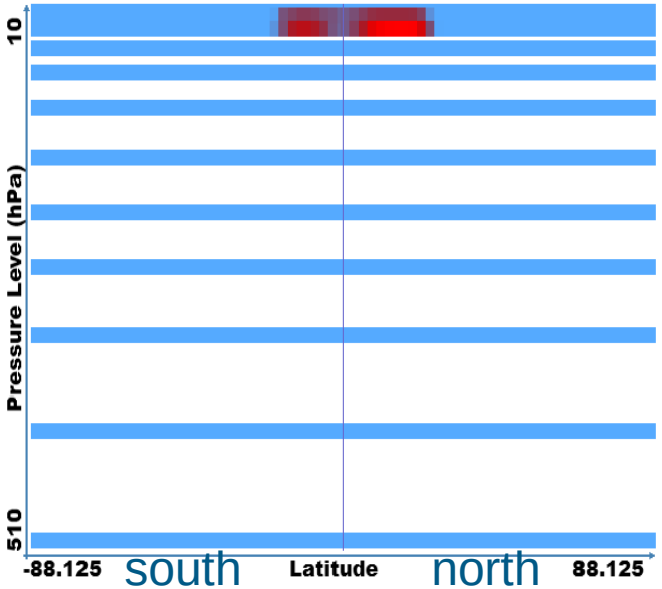


Analyze Relations between Dimensions

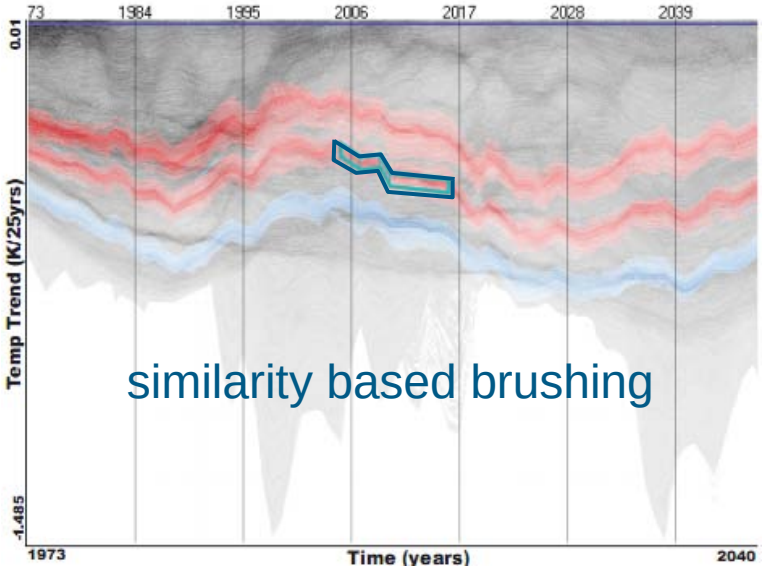
Up to now:

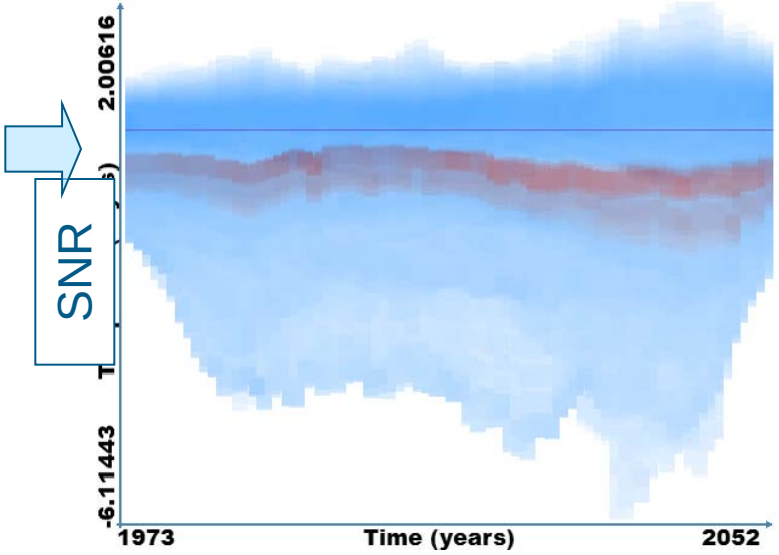
→ investigation in one direction





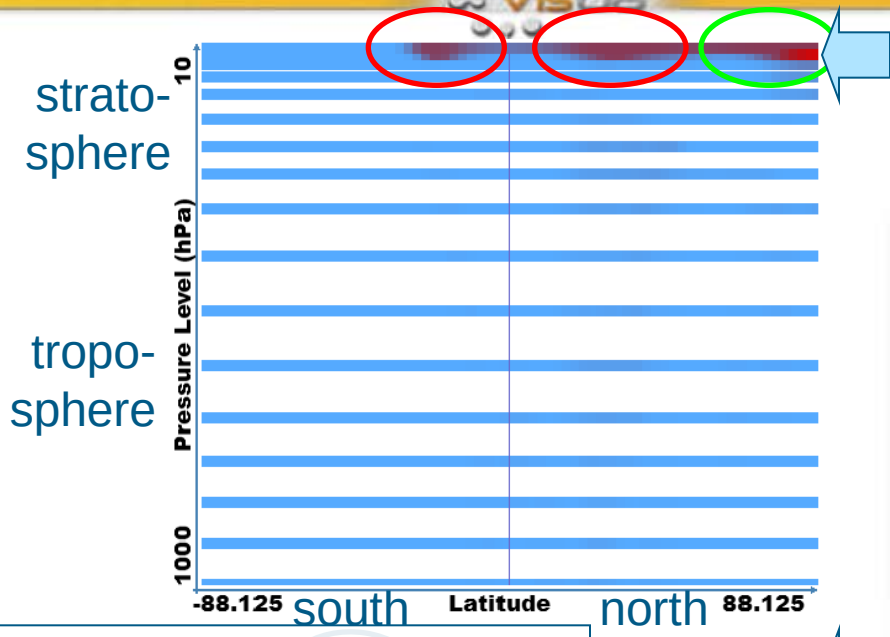
→ check relation in other direction



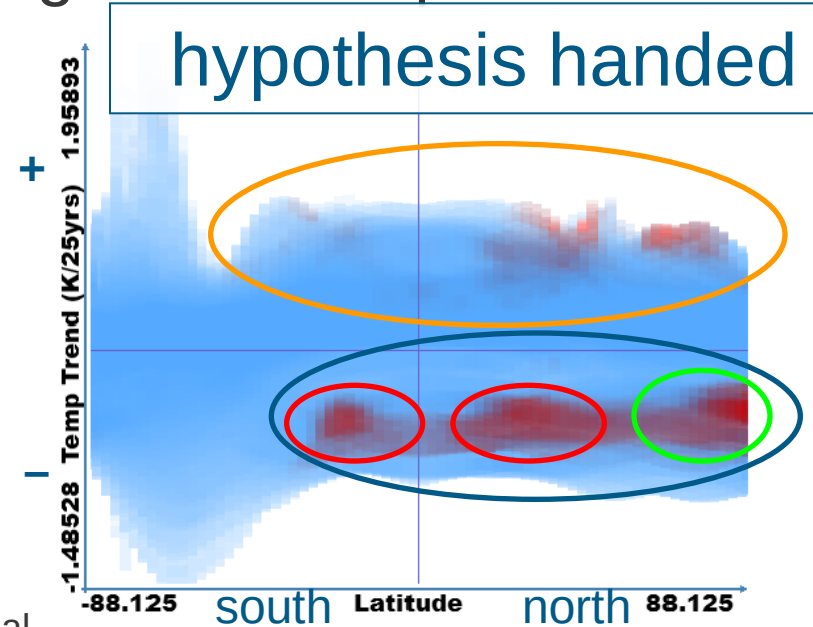


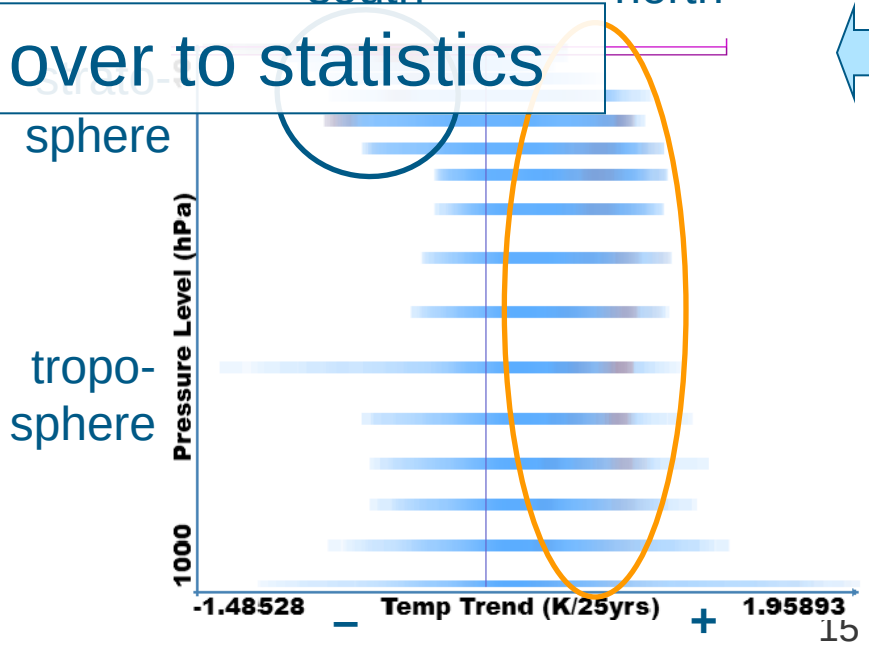
Generated Hypothesis / ECHAM5 temp.

Promising indicator region is seemingly located in **lower stratosphere**, geographically located at **northern latitudes** & **tropics**.
Corresponding **cooling trend** considered robust over whole investigated time span.



hypothesis handed over to statistics





Conclusions

Visual Exploration of derived parameters (linear trends, SNR)

- rapidly generate promising hypothesis
  → afterwards checked with classical statistics
- useful to narrow down parameter settings (statistics)
- in comparison to the original approach:
  faster, **more flexible**, **and informed** exploration

Future work

- further integration of statistical methods in our visual exploration framework
- detailed quantitative evaluation of results w. statistics

Acknowledgements

- Wegener Center for Climate and Global Change, Austria
  - Gottfried Kirchengast
  - Bettina C. Lackner
- Austrian Science Fund
  - Project INDICATE P128733-N10
- Datasets are courtesy of
  - Max-Planck-Institute for Meteorology, Hamburg
  - European Centre for Medium-Weather Forecasts

→ www.ii.UiB.no/vis