Histogram-based Aggregation: bridging meter-scale InSAR permafrost deformation data and climate model

Zhijun Liu[1], Barbara Widhalm[2], Annett Bartsch[2], Thomas Kleinen[1], Victor Brovkin[1]

1. Max Planck Institute for Meteorology, Hamburg, Germany 2. b.geos, Korneuburg, Australia

zhijun.liu@mpimet.mpg.de

This presentation participates in OSPP

EGU

Outstanding Student & PhD candidate Presentation contest

MAX-PLANCK-INSTITUT FÜR METEOROLOGIE

Q-ARCTIC

RESEARCH MOTIVATIONS

- **Long-term Permafrost Deformation (PD)** reflects permafrost degradation under climate change. PD is affected by climate change as it occurs during melting and freezing. PD also has returned impacts on climate with change in soil water content, surface condition and carbon release.
- The driving factors of PD and how PD evolves under current climate change are not fully understood. This requires **better representation of permafrost in Eerth System Models**, which have km-scale resolutions. Instead of physical-process-based modelling, we aim for statistical-based modelling.
- **Metre-scale Resolution InSAR Data** is becoming available. To capture meter-scale PD in ESMs at km-scale resolution, a scale-bridging aggregation method is required with high computation-efficiency.

Fig 1. The Batagaika crater in eastern Russia formed dueg to thawing permafrost.Credit: Yuri Kozyrev/NOOR/eyevine

HISTOGRAM-BASED AGGREGATION

- **Advantage 1: preserves the statistical feature of original meter-scale data**

10m (original InSAR) 10km (averaged)

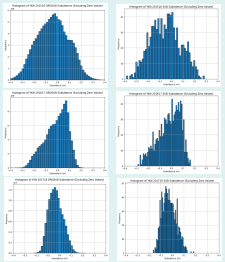


Fig 2. Distortion of Information: statistical features of subsidence data are lost or distorted during traditional aggregation.

- **Advantage 2: computational efficiency**
  - **Data Histogram**: 50-150 parameters, depending on bin number true to original data
  - **Probability Distribution Function**: 2-4 parameters, depending on PDF type capture longtail behaviour

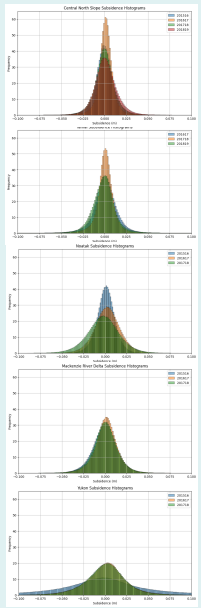
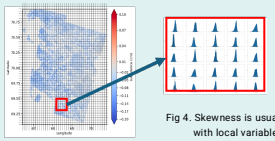


Fig 3. Data histograms displays regional and temporal variability. Yukon region has the significantly larger STD and is the only region under Atlantic not Arctic impacts.

- **Advantage 3: flexibility on aggregation resolution**



- The goal-resolution can vary between 1km and 100km depending on the need of heterogeneity scale of result. A minimum of 3000 data sample is required.

Fig 4. Skewness is usually stronger at finer resolution with local variables being more dominant.

- **Standard deviation** (STD) is not uncertainty but reflects regional characters in the local permafrost system. It indicates soil mobility and the **vulnerability of local permafrost to climate change**. Interpreting regional permafrost deformation with InSAR data STD is applied and elaborated in our submitted manuscript (inquiries welcome).
- Traditional upscaling is point oriented. **Histogram-based aggregation includes 2 or more dimensions (mean, STD, skewness, etc) of data information** without much more computational cost.

CLUSTERING HISTOGRAMS: DISENTANGLING IMPACTS

- **EP-Means clustering:** clustering on distributions instead of data points, it helps to highlight the dominating spatial patterns when multiple entangled factors are present.

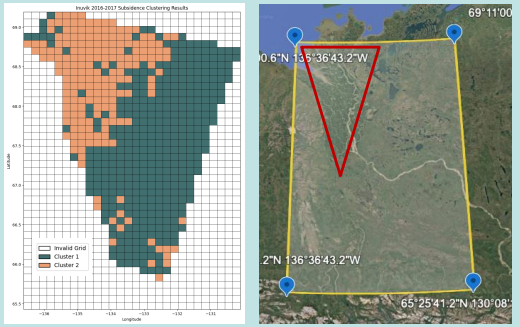


Fig 5. Example of permafrost deformation histogram clustering result in Inuvik/Mackenzie River Delta (MRD). Pattern in the left plot is consistent in all 3 years of data, even though in some years this region is smaller. This pattern aligns with the area of flood plain in MRD, demonstrating that flood plain impact is dominant in this region.

- In each studied location, the **dominating mechanism is different** depending the environmental conditions (coastal distance, geomorphology, soil types, river distribution, vegetation types, etc…)
- In Central North Slope, clustering result identifies organic-matter-rich region; in Inuvik/MRD, flood plain region is identified.
- Clustering histograms allows the data statistics show the dominant factor for **regional classifications**, without researchers making any presumptions.

ENHANCED INTERPRETATION ON CLIMATE CORRELATIONS

- Direct correlation analysis applied on traditionally aggregated permafrost deformation and climatic factors shows extremely low (lower than 0.1) correlation.
- Correlation analysis on separate cluster groups within a region shows impacts on deformation from the same factor can be significantly different depending on the local dominating condition, which is hard to pre-determine without histogram-based clustering.

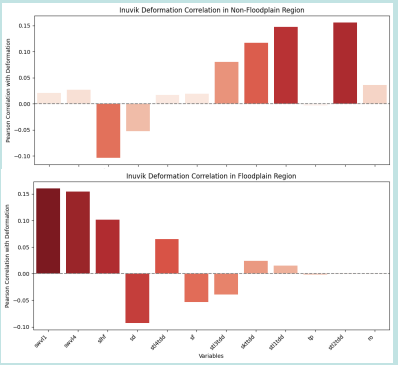


Fig 6. Correlation analysis on floodplain and non-floodplain region in MRD/Inuvik based on the histogram-based clustering result. Variables include: soil water content, surface latent heat flux, snow depth, snow fall, soil temperature, skin temperature, precipitation, river runoff. Data from ERA5-Land at ~10km resolution.

TAKE-HOME MESSAGES

- Meter-scale InSAR data provides a new perspective to examine highly heterogeneous permafrost deformation. Statistical features (mean, STD, skewness) of permafrost deformation can reveal spatial and temporal variability in local system.
- Histogram-based aggregation on InSAR permafrost deformation data preserves more information at coarser resolution efficiently. It contains more dimensions of information than traditional point-oriented aggregation.
- EP-Means clustering results and further correlation analysis result demonstrate that histogram-based aggregation can help disentangle driving factors in a complex system and enhance the interpretability on climatic forcing correlations.
- Histogram-based aggregation helps to disentangle climatic forcing impacts in building a permafrost model; using histogram as model output format to represent permafrost deformation can also capture its high heterogeneity more accurately and efficiently.