

Methods note · Land and freshwater

Comparing ground-temperature records across depth and siting

Elín Bjarnadóttir · 1 October 2026

Fictional author and institute · Editorial demonstration, not peer reviewed

ABSTRACT

Ground-temperature records can acquire apparent changes when sensors move, depths change, or surface conditions differ. This methods note synthesises a documented Arctic borehole dataset, permafrost-network metadata, and measurement-uncertainty guidance to examine when records are comparable. It distinguishes instrumental transfer from relocation between environments and explains why depth changes both seasonal sensitivity and response time. Fictional worked cases show how joining two individually valid records can introduce a step, and how depth uncertainty enters a temperature budget. The note recommends preserving comparable record segments, overlap evidence, and adjustment assumptions rather than forcing visual continuity. It estimates no regional trend and provides no universal depth correction or minimum overlap. Its contribution is a practical framework for keeping continuity of meaning visible when an observing system changes.

Keywords: ground temperature · thermistor depth · station relocation · record comparability · permafrost

A source-based editorial text, not a report of original Still North research. Examples and imagery are illustrative.

The comparison is part of the measurement

A long ground-temperature record can acquire a false step without a thermometer becoming inaccurate. A sensor is replaced at a different depth, a borehole is moved away from an eroding bank, or the station becomes sheltered by a growing snowdrift. Each instrument may record its immediate surroundings correctly. The difficulty lies in treating the observations as one unchanged measurement of the same ground. This methods note asks what evidence is needed before joining such records or comparing their warming rates.

The synthesis draws on Clow's documentation of the Alaska DOI/GTN-P Deep Borehole Array, the Global Terrestrial Network for Permafrost data platform, and NIST uncertainty guidance. These sources cover an actual observing system, the distinction between temperature and active-layer products, and the treatment of measurement uncertainty. The numerical examples below are fictional teaching cases, not Still North observations. No regional temperature trend is estimated here.

Depth changes both timing and sensitivity

A seasonal surface-temperature signal propagates downwards while becoming smaller and later. In an ideal homogeneous, conduction-dominated medium, a periodic signal is attenuated approximately exponentially with depth. Real soils introduce layering, latent heat, changing water content, and sometimes moving groundwater. The simple model explains why depth matters, but does not provide a universal correction between two sensors. A temperature measured at two metres is not a noisier version of the temperature measured at twenty metres.

At shallow depths, one annual visit can land on a different part of the seasonal cycle in successive years. Even visits on the same calendar date can encounter dif-

ferent phases after contrasting snow seasons. Continuous logging allows an annual mean to be estimated more directly, provided missing months are addressed. Deep records suppress much seasonal variation, but respond more slowly to recent surface change. Comparing short trends at different depths therefore mixes different temporal filters as well as potentially different climates.

Clow's published data description is particularly useful because it separates what was measured from corrections that were not applied. The temperature profiles retain drilling disturbance rather than silently presenting a reconstructed undisturbed state. The study also documents instrument response, calibration, and changes in logging practice. Such distinctions demonstrate why a downloaded temperature column needs its processing history, not simply a place name and a unit.

Siting controls the boundary condition

Ground receives a surface energy history shaped by snow, vegetation, organic layers, drainage, and disturbance. A wind-scoured ridge and a sheltered hollow may experience similar air temperature but different winter insulation. Moving between them changes the environmental system under observation. A new access track can alter drainage or trap snow. A nearby building can introduce heat. These are not all instrument errors, and subtracting an instrumental calibration offset will not remove their effects.

Two spatial questions should be distinguished. A fixed borehole can show change at that site, including real local ecological change. A regional assessment asks how that site represents a wider population of terrain. Excluding a station because shrubs have grown around it may discard an important landscape response. Allowing that station alone to represent every terrain type is a different

mistake. Stable site identity and representative network design solve different problems.

The GTN-P platform retains separate permafrost-temperature and active-layer-thickness products and provides site attributes such as elevation and vegetation. A practical comparison should retain the actual measured depths, not infer them from a map symbol or from modelled temperature layers displayed alongside observations. A modelled ten-metre field and a thermometer near ten metres can be informative together, but are not interchangeable records with identical support.

Worked example: a replacement that resembles warming

Fictional worked example. A programme has ten complete annual means from a sensor nominally at ten metres. Erosion forces relocation to a new borehole on a more sheltered surface. During two overlapping years, the old site averages -4.2°C and -4.0°C , while the new site averages -3.6°C and -3.3°C . Joining the raw series would introduce a difference of roughly 0.6 to 0.7°C at the handover, in addition to any underlying change.

One tempting response is to subtract 0.65°C from every new reading. That would assume a constant site relationship across snow conditions, thermal adjustment, and future years. Two paired means do not establish this. They show that the records differ, not that the difference is an enduring calibration constant. Reporting separate series with an explicit handover may be more defensible than creating a seamless record. A short gap in comparability is preferable to an invented certainty.

A different replacement case is more tractable. Suppose two calibrated sensors occupy the same borehole and verified depth during a full overlapping seasonal cycle. Their paired differences show no meaningful seasonal dependence and agree with laboratory checks. An offset adjustment may then be supportable, with uncertainty from the overlap retained. This is a transfer between instruments observing substantially the same environment, rather than a transfer between landscapes. The distinction determines how much an overlap can establish.

Depth uncertainty belongs in the budget

Nominal depth is not always actual depth. Cable marks, borehole inclination, slack, reference-height changes, and ground settlement can affect the position assigned to a reading. A depth error matters most where the vertical temperature gradient is steep. For a small displacement, the associated temperature difference can be approximated as the local gradient multiplied by the depth displacement. The gradient must come from an appropriate profile, not a generic regional value.

For an explicitly hypothetical local gradient of 0.08°C per metre and a standard depth uncertainty of 0.25 metres, the corresponding standard temperature-uncertainty component is about 0.02°C . If an independent calibration component is 0.03°C , their root-sum-of-squares combination is about 0.036°C , before other contributions. This calculation illustrates sensitivity, not an instrument specification. Shared calibration errors, uncertain gradients, or a changing depth reference require a fuller treatment rather than automatic independence.

NIST guidance explicitly includes covariance when combining standard uncertainties. A common calibration bias can partly cancel when estimating change with the same stable instrument, yet remain important when comparing absolute temperatures between instruments. Conversely, independent noise may average down while a site relocation does not. Counting thousands of hourly readings cannot make a structural discontinuity disappear.

An analysis sequence that preserves the evidence

1. Define the target: site temperature at a stated depth, a depth-specific trend, or a regional terrain-weighted estimate.
2. Inventory instrument, depth-reference, logging, borehole, surface, and processing changes before fitting a trend.
3. Compare seasonal coverage and averaging windows. Treat a summer-only mean as a different statistic from a complete annual mean.
4. Use overlap and profiles to test a proposed transfer, and preserve both adjusted and original values with their rationale.
5. Present sensitivity to credible alternative handling of discontinuities, rather than choosing whichever creates the smoothest graph.

This sequence need not produce a complicated model. A small table of comparable record segments, an unjoined figure, and a clear account of exclusions may answer the question better than a heavily adjusted single line. Where physical modelling is justified, its added assumptions should be visible. A transfer model cannot manufacture the missing winter observations or site history that would have constrained it.

Limits and reporting implications

The conduction argument is least adequate where advective water movement, thawing ice, or major surface disturbance dominates. Near the melting point, additional energy can change ice content with relatively little temperature rise. A muted trend is therefore not a direct measure of small physical consequence. Deep thermal records also integrate past surface conditions, so their short-term rates should not be read as immediate responses to the latest air-temperature anomaly.

This note does not prescribe a universal minimum overlap, a standard reference depth, or an automatic break-detection threshold. Those depend on the signal, terrain, instrumentation, and decision. Statistical detection can identify a suspicious step, but cannot establish its cause without metadata and physical evidence. The practical result is a sharper reporting obligation: publish which measurements are directly comparable, which require a transfer assumption, and which remain separate. The continuity worth protecting is continuity of meaning, not merely continuity of a plotted line.

References

1. Clow, G. D. (2014). Temperature data acquired from the DOI/GTN-P Deep Borehole Array on the Arctic

- Slope of Alaska, 1973–2013. *Earth System Science Data*, 6, 201–218.
2. Global Terrestrial Network for Permafrost. GTN-P data platform: permafrost temperature, active-layer thickness, and site metadata.
 3. National Institute of Standards and Technology. Technical Note 1297, section 5: Combined standard uncertainty.