Estimating normals for series without normals (Part II)

Timothy Osborn, Michael Taylor*, Phil Jones, David Lister
Climatic Research Unit, School of Environmental Science, UEA

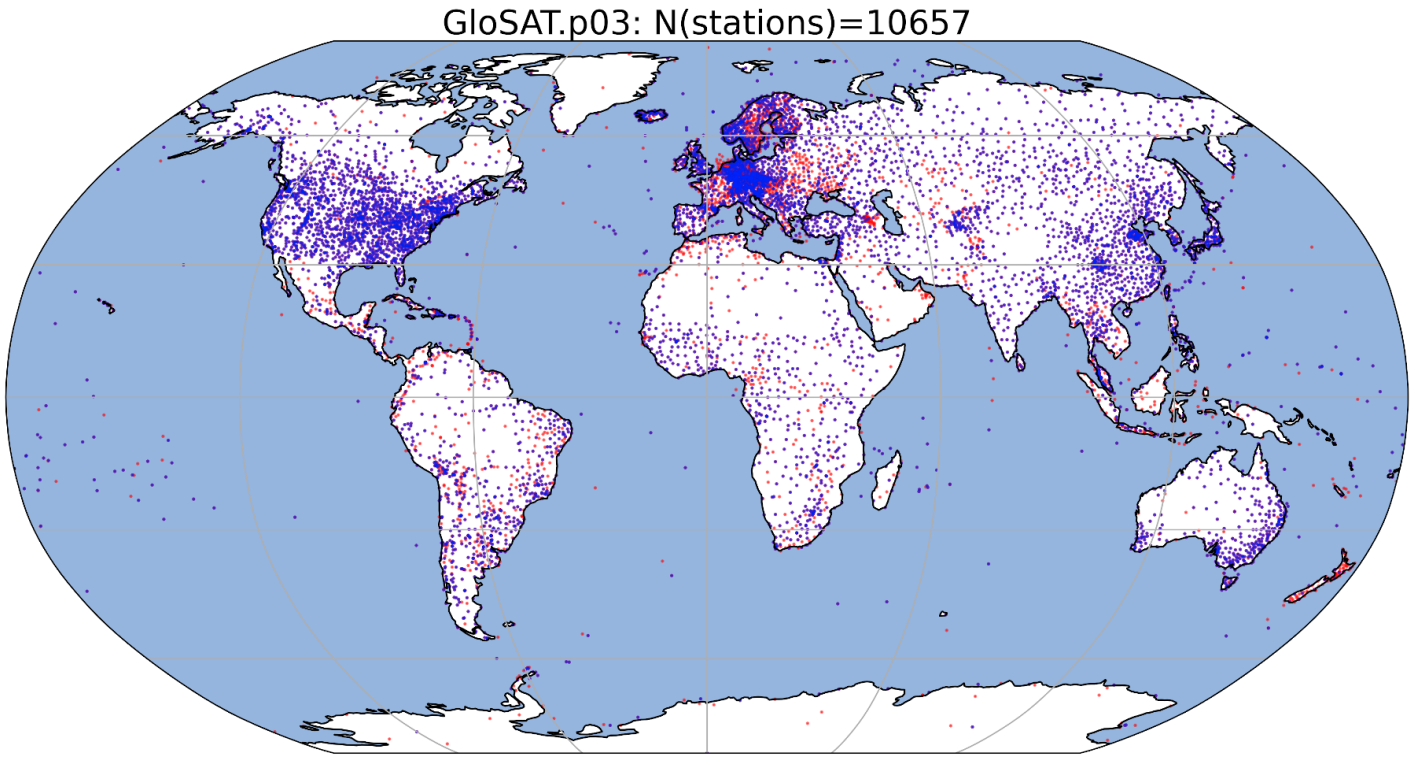
* michael.a.taylor@uea.ac.uk

University of East Anglia

GlobSAT
Global Surface Air Temperature

Why are we doing this ?

We need the normals to incorporate anomaly timeseries into the global SAT record



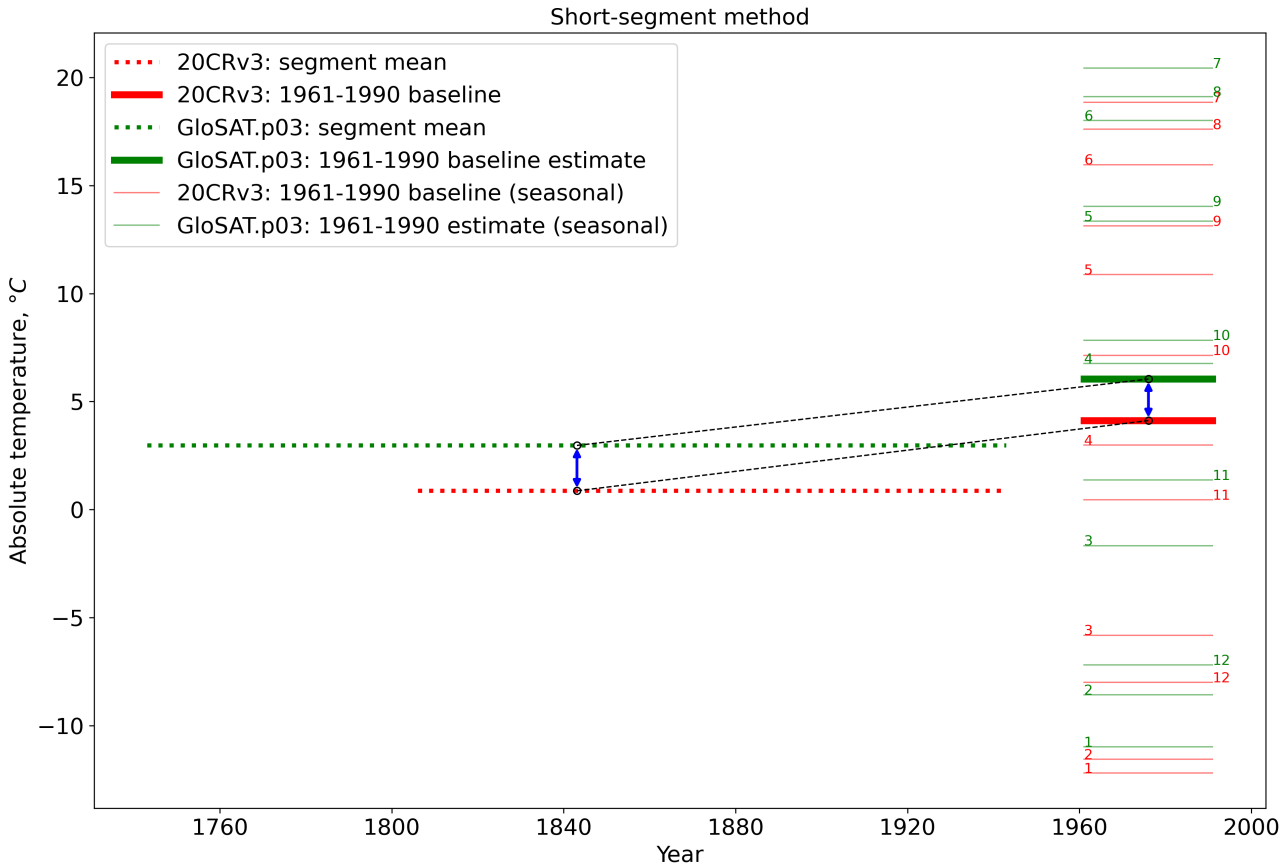
- Global: N(short-segment stations)=2618
- Global: N(stations with 1961-1990 normals)=8039

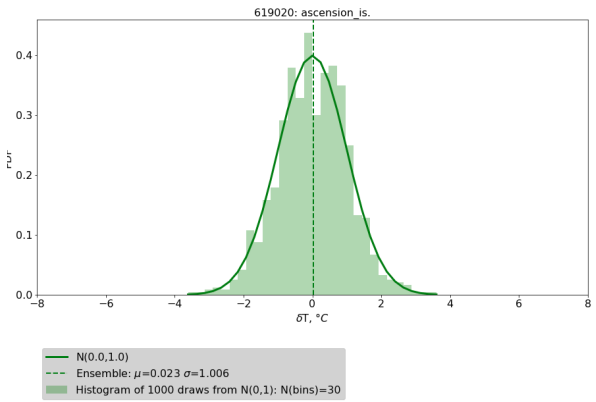
Potential for calculating new anomaly timeseries for **~24% of stations** in GloSAT.p03

First attempt: normals using reanalysis (+ Monte-Carlo)

Co-located 20CRv3 0.7°x0.7° extracted reference series – test on African stations

Station Normal (1961-1990) = 20CRv3 Normal (1961-1990) + (Station Mean (segment) - 20CRv3 Mean (segment))



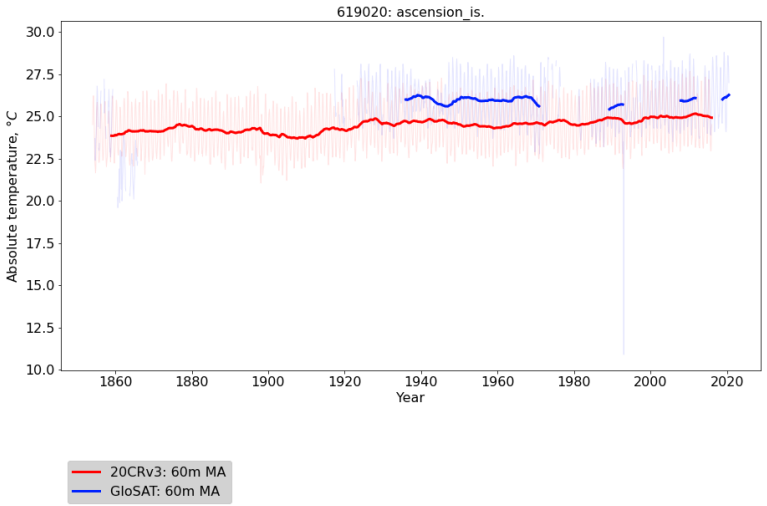


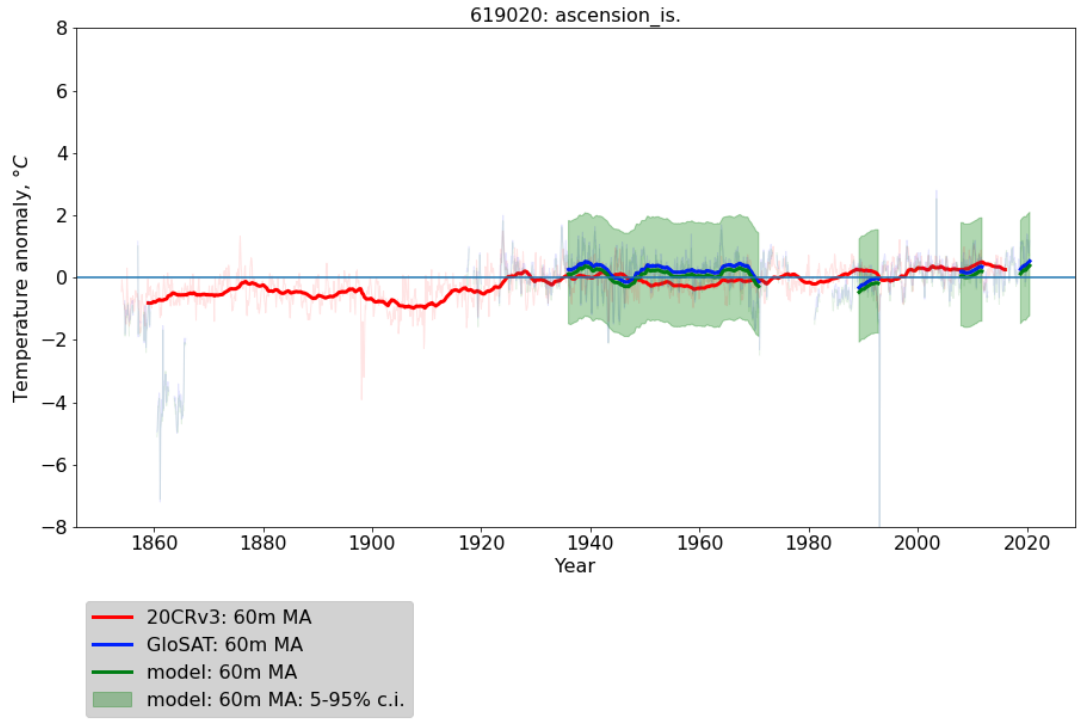
Monte Carlo:
1000 draws from N(0,1) → ensemble of station and reanalysis segment members to estimate a 5-95% c.i. on the station baseline normal estimate

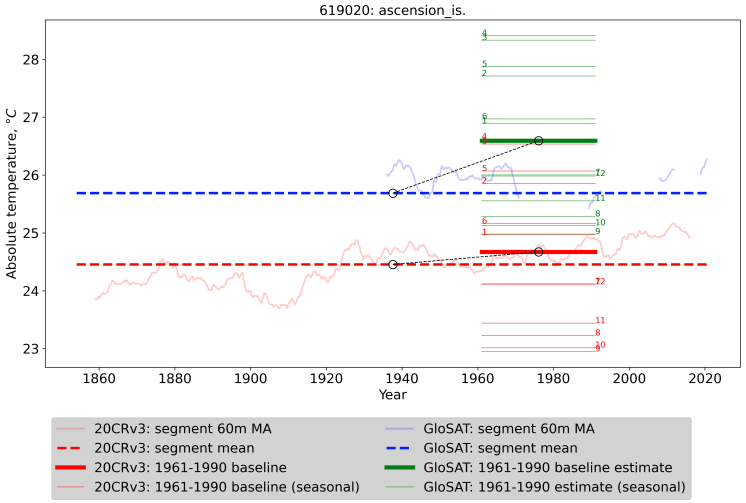
Q. But what is a 'suitable' choice of uncertainty level for this ?
Q. Use 20CRv3 percentiles ?

Example: Ascension Island

Test suitability of reanalysis in a data sparse region





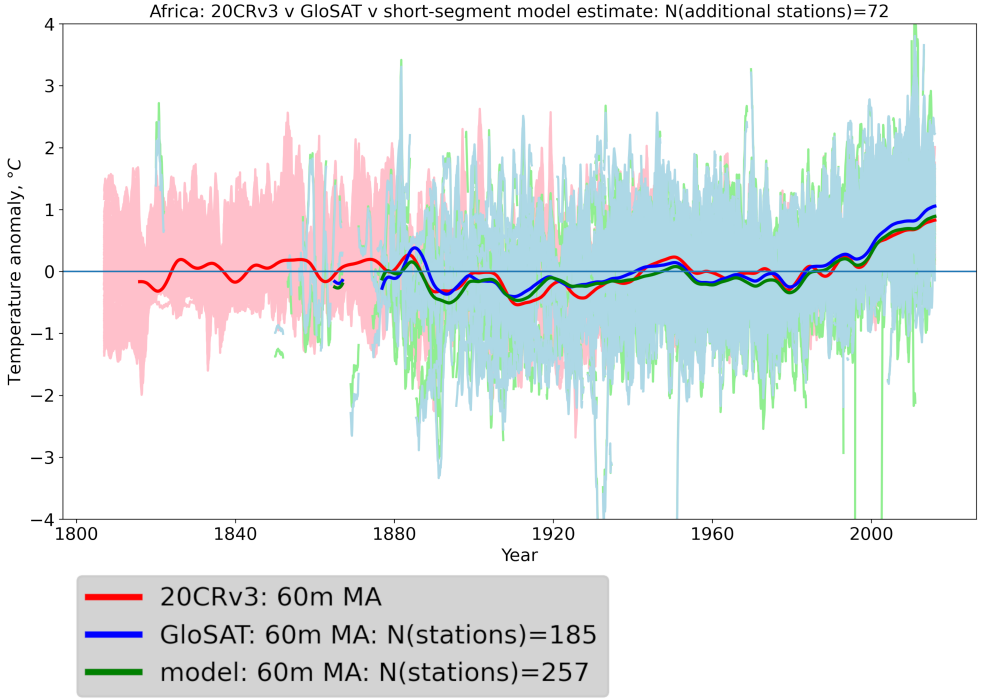


Reanalysis + this approach allows us to estimate normals for remote places like Ascension Island (the nearest island is St Helena 1259 km away)

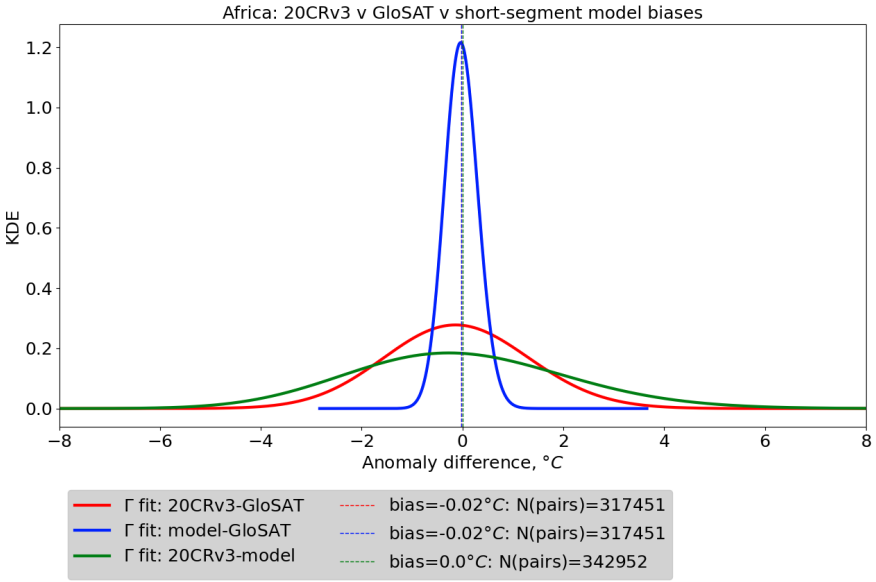
It is a dot in a sparse region like many stations in Africa are

Africa anomaly timeseries

We are able to grow the anomaly record but a lot of spread (+ outliers) wrt reanalysis



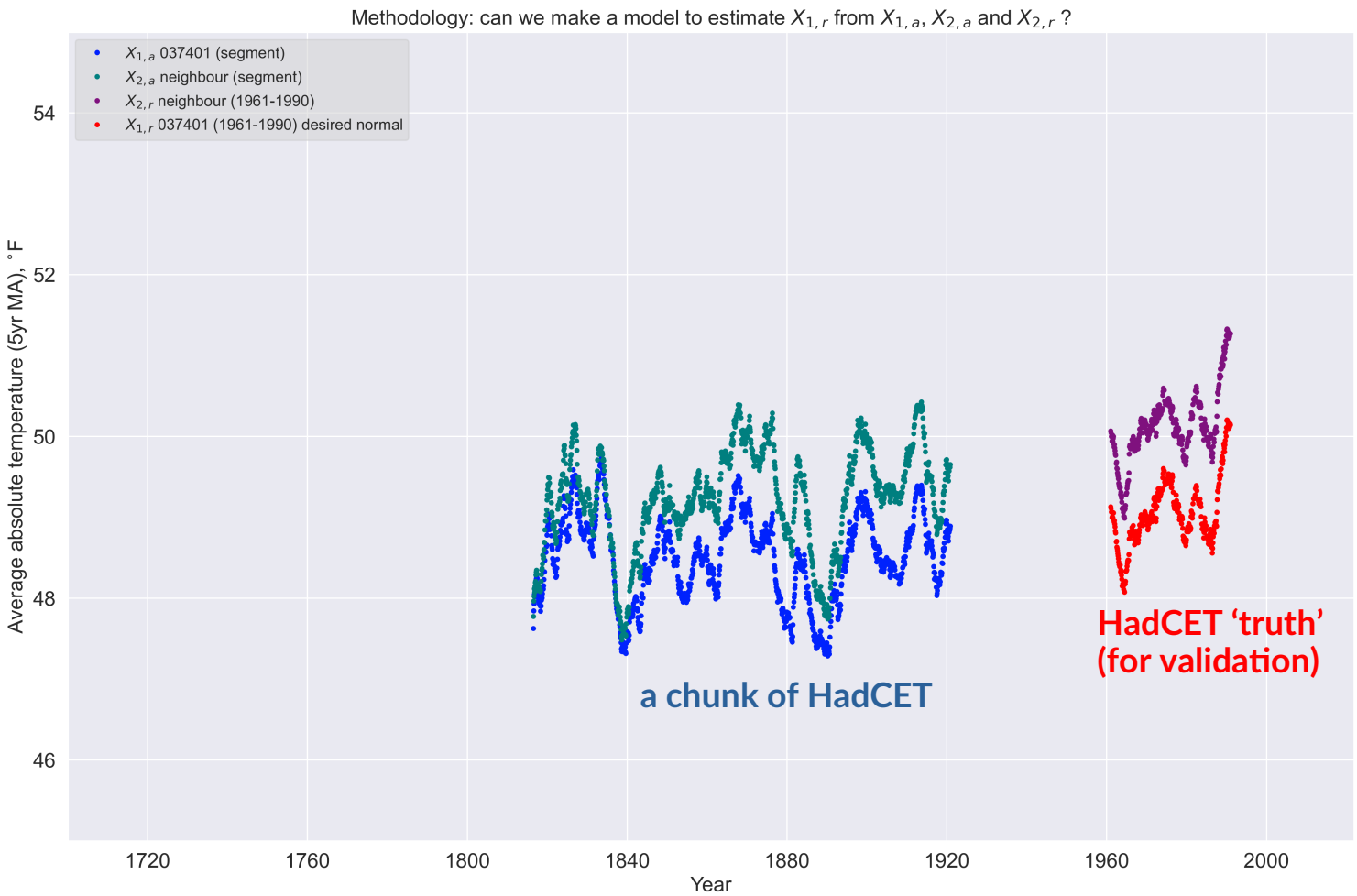
There is still a **lot of spread** though (**red** and **green** KDEs below) compared to both 20CRv3 and the existing record of observations with normals



The procedure allows for growth of the anomaly timeseries record and calculation of an **African regional anomaly timeseries** which is fairly consistent with trends in reanalysis

Second attempt: back to theory

We want an estimate of the baseline mean level <u>+ its standard error</u>



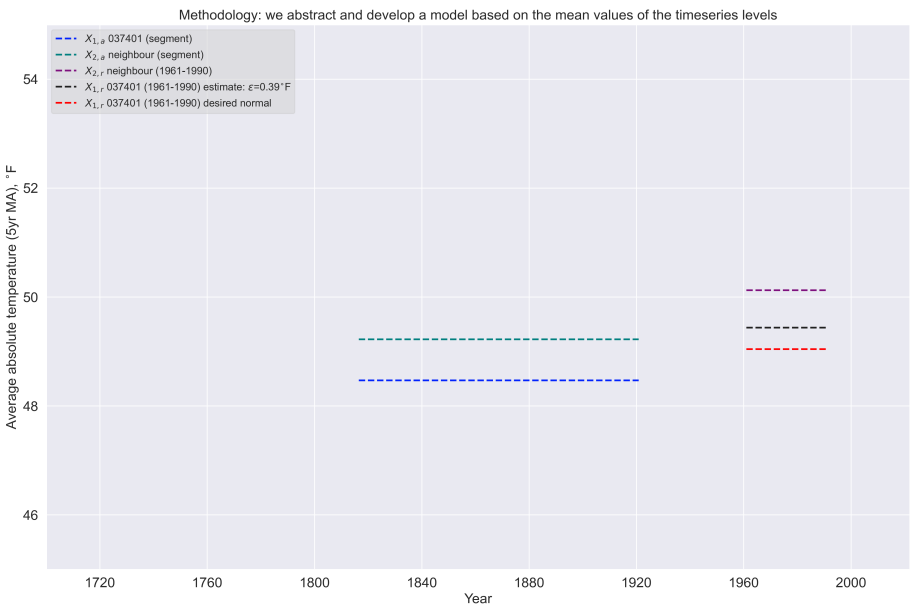
(and reduce the error if we can!)

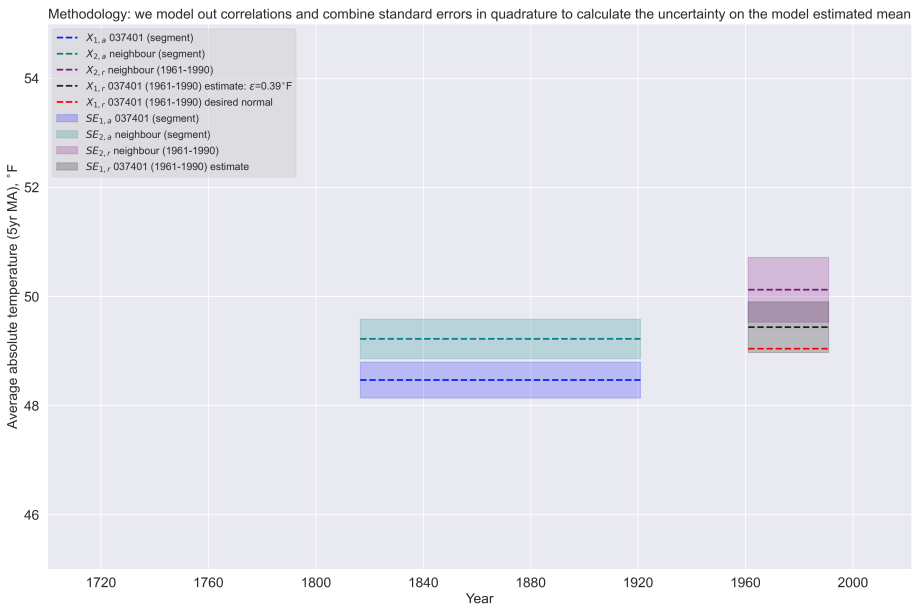
University of East Anglia

GlobSAT Global Surface Air Temperature

Various modeling approaches

We abstract the problem and work with mean levels and optimize on the standard error

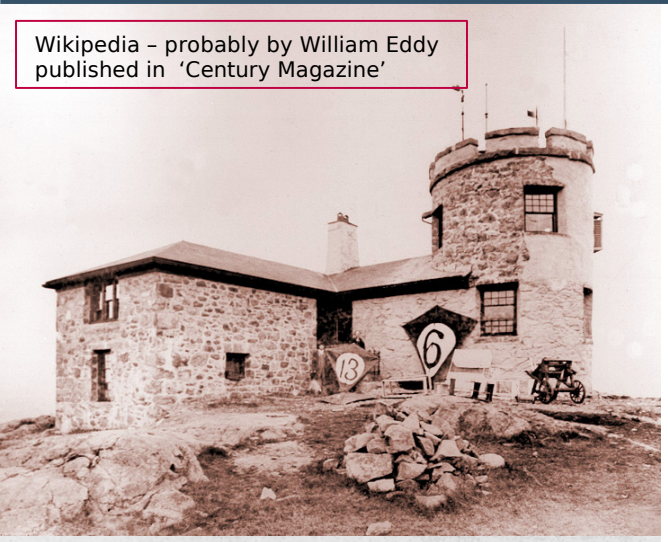




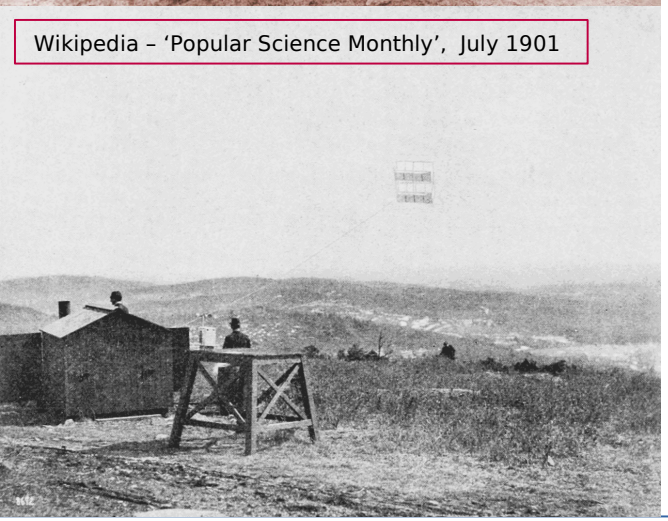
- Model 1A: single neighbour (uncorrelated segments) $\overline{X_{1,r}} = \overline{X_{1,a}} + \left(\overline{X_{2,r}} - \overline{X_{2,a}}\right) + \varepsilon_{1|2,a|r}$ $SE^2_{1,r} = SE^2_{1,a} + SE^2_{2,r} + SE^2_{2,a}$
- Model 1B: single neighbour (modeling out the correlation) $\overline{X_{1,r}} = \overline{X_{2,r}} + \overline{\left(X_{1,a} - X_{2,a}\right)} + \varepsilon_{1|2,a|r}$ $SE^2_{1,r} = SE^2_{2,r} + SE^2_{1-2,a}$
- Model 2A: mean of ensemble of neighbours
- Model 2B: mean of core neighbours from ensemble of neighbours
- Model 2C: mean of co-located max density neighbour ensemble

Results centered on Blue Hill Observatory (BHO)

Longest continuous record of meteorological observations in North America 1885-2021



Wikipedia – probably by William Eddy published in 'Century Magazine'



Wikipedia – 'Popular Science Monthly', July 1901

- Rural site at (42.2°N, 71.1°W) 193 m above sea level

- Established in 1885

- **Kite soundings**. On August 4th 1894 William Eddy launched 5 bowed kites with a total area of 9 m^2 to carry a ½ lb (~ 0.23 kg) Marvin meteograph up to 1400 ft (~ 427 m)

- BHO recorded the strongest ever wind gust in the US at 186 mph (~ 83 m/s) during the **Great New England Hurricane** that made landfall on 21 September 1938

PS: We're working on a back-extension from 1885-1786 ...

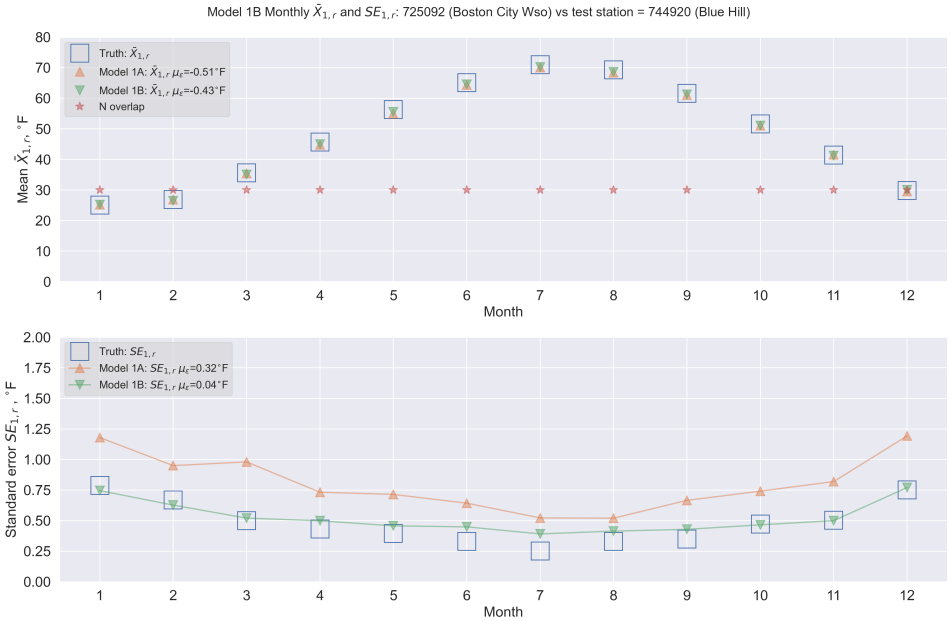
University of East Anglia

GlobSAT Global Surface Air Temperature

Models 1A and 1B

A ‘good’ neighbour isn’t necessarily the closest

Boston is only 19 km from BHO but the estimated mean is 0.52 °F lower than the ‘true’ observed mean

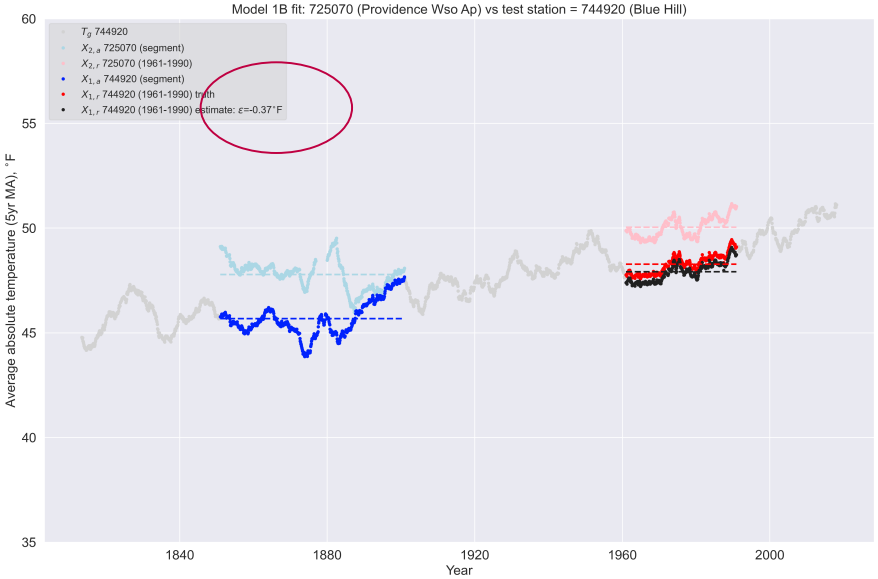


Model 1B is much better than Model 1A because the Boston reference is correlated with BHO

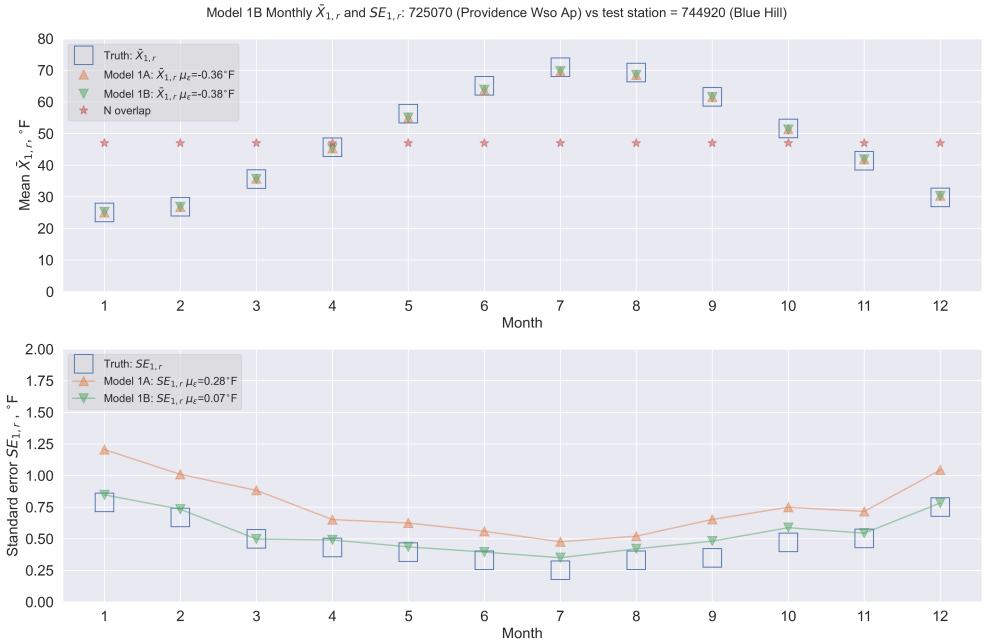
But there are better neighbours further away …

Models 1A and 1B

… it also needs a high number of overlap years in the segment

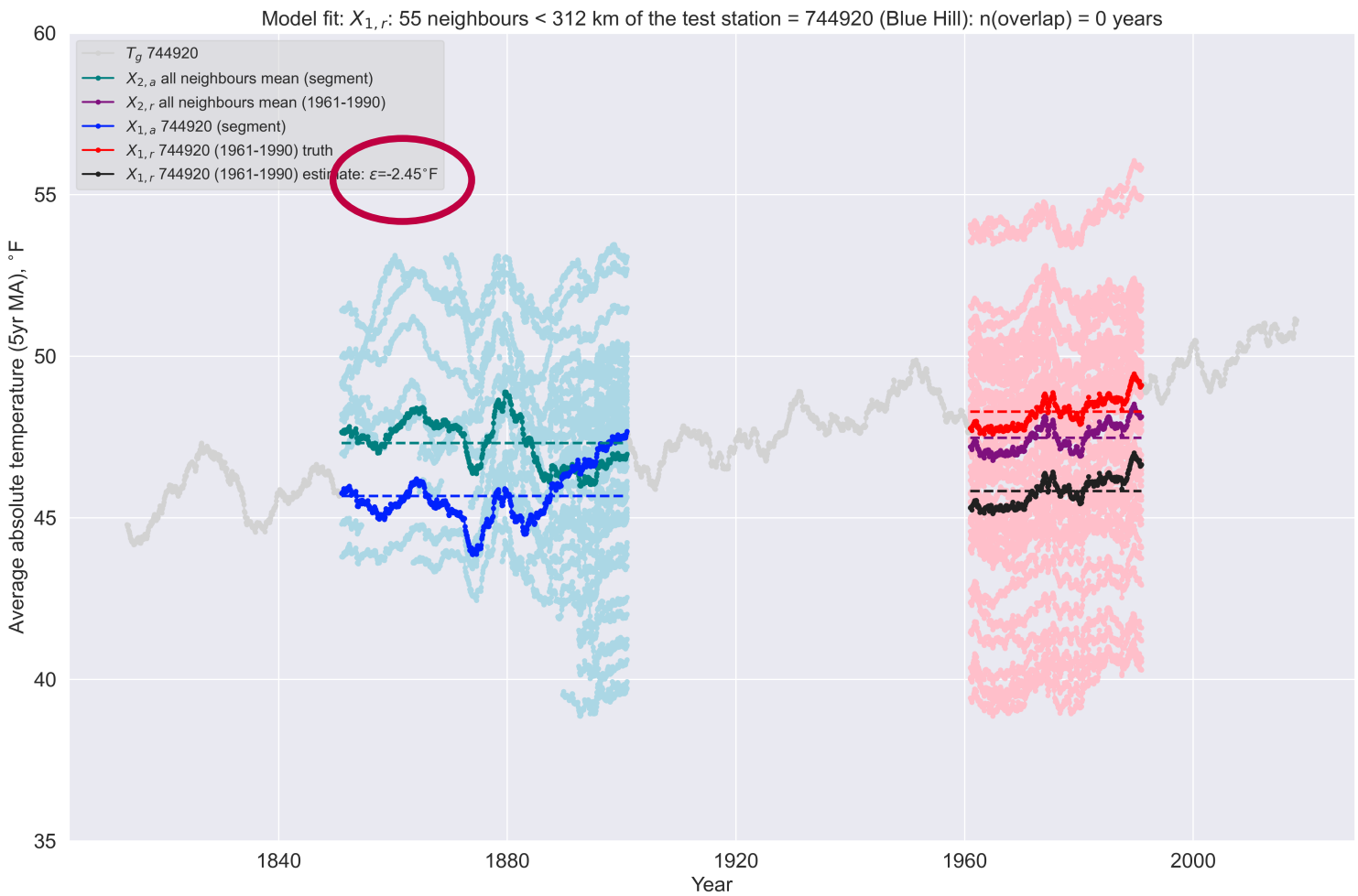


Providence, RI is 56 km away from BHO but has more overlap (~ 50 Januaries etc in the 1850-1900 segment) and Model 1B leads to a better estimate (0.37°F vs 0.52°F error)



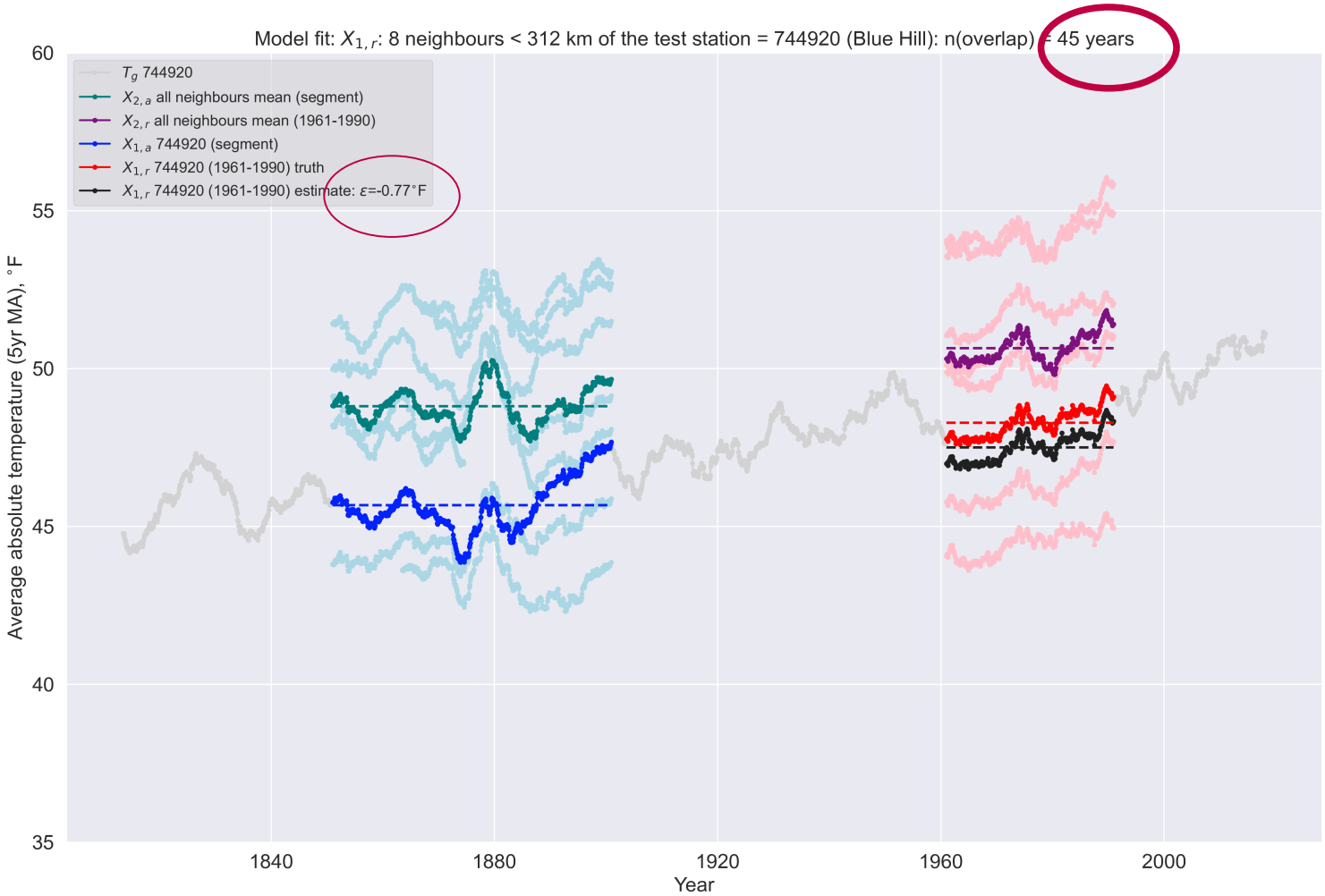
Model 2A

Just averaging neighbour overlaps (of differing length) doesn't work well



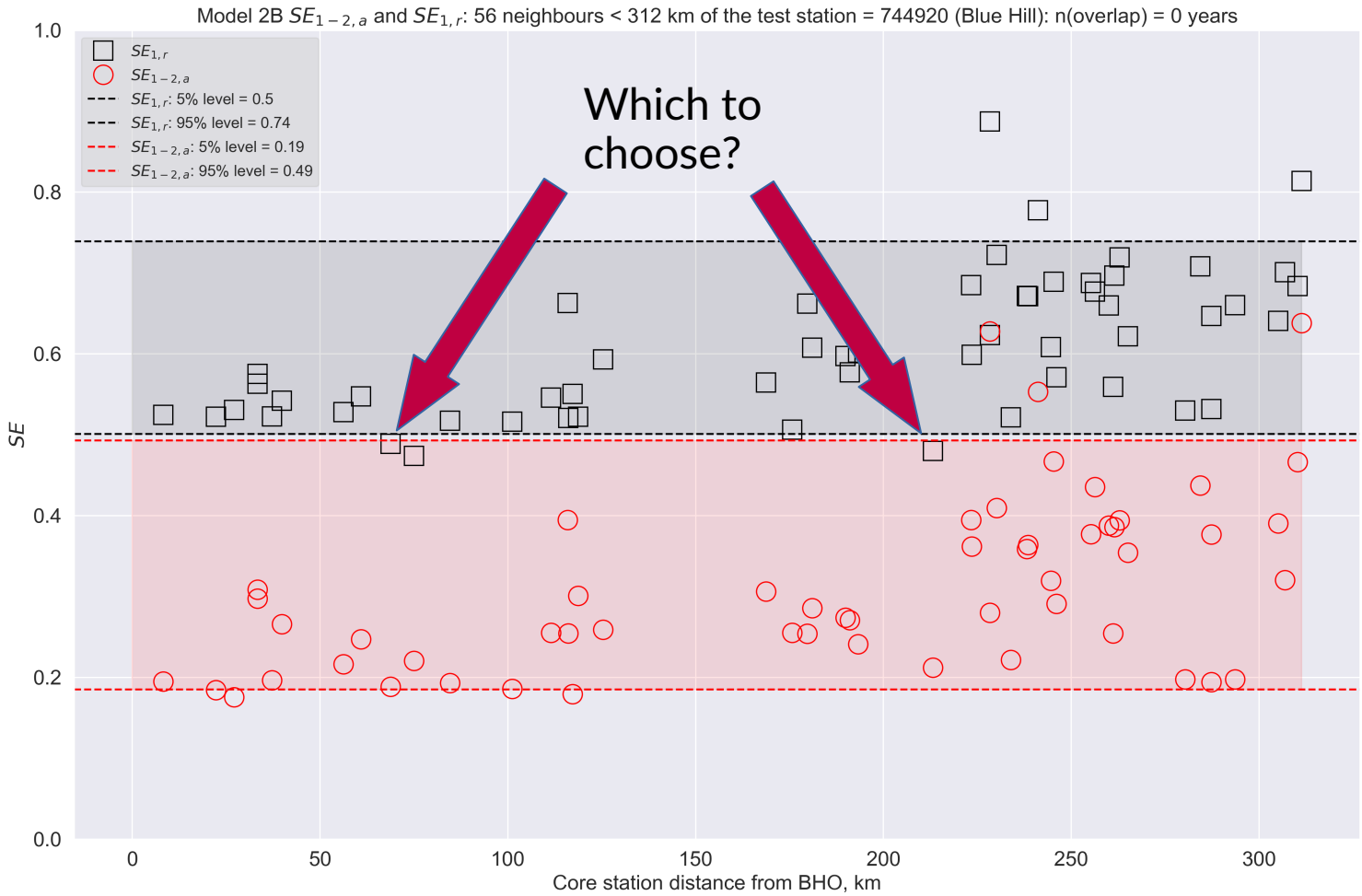
Model 2A

… but requiring high number of overlaps in the segment reduces the error



Model 2B

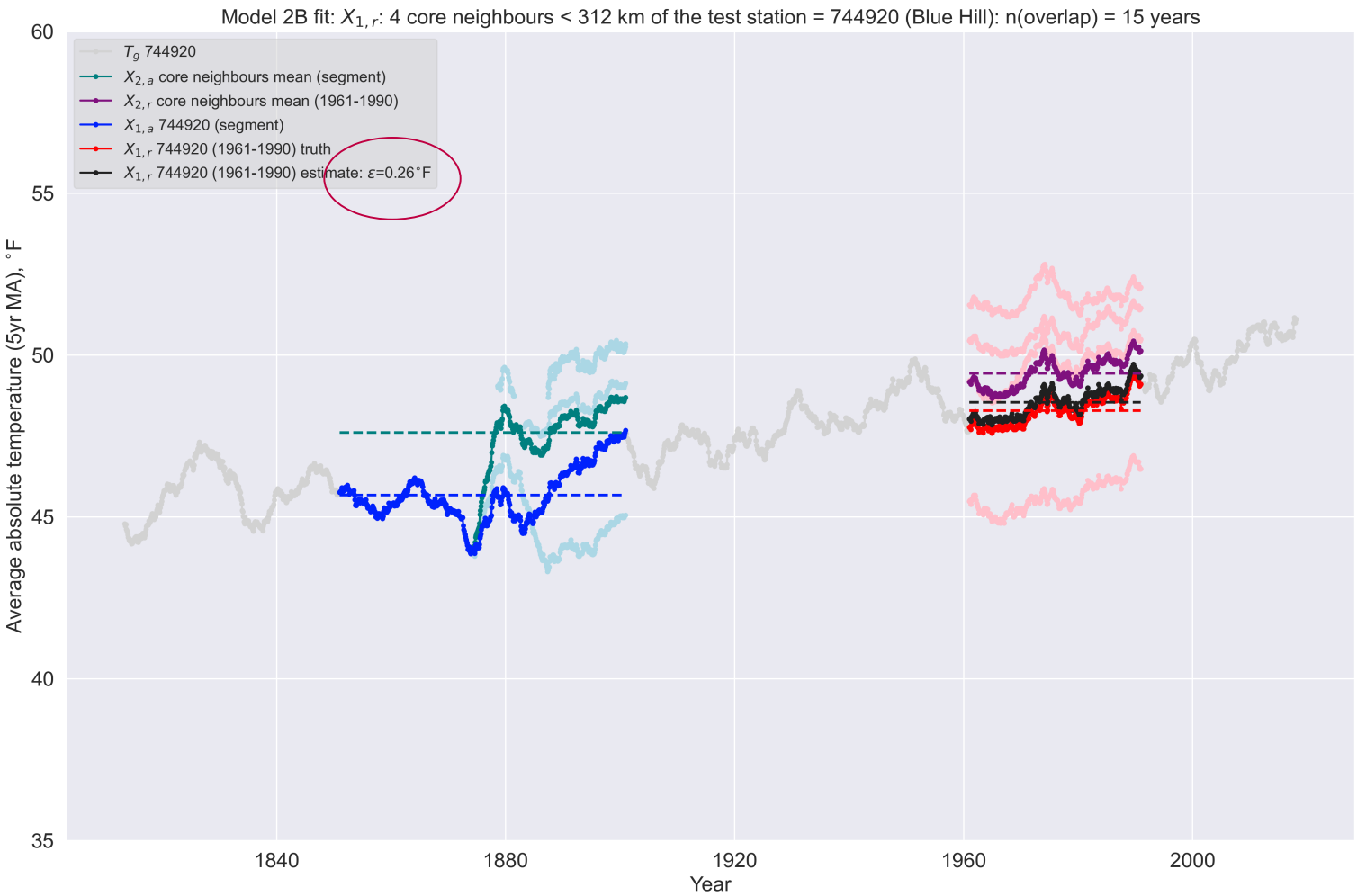
How about being choosey and selecting ‘core’ neighbours ?



… note that the dependence on distance is weak (at this scale of 312 km lasso radius neighbour selection)

Model 2B

Choosing by optimizing on the standard error



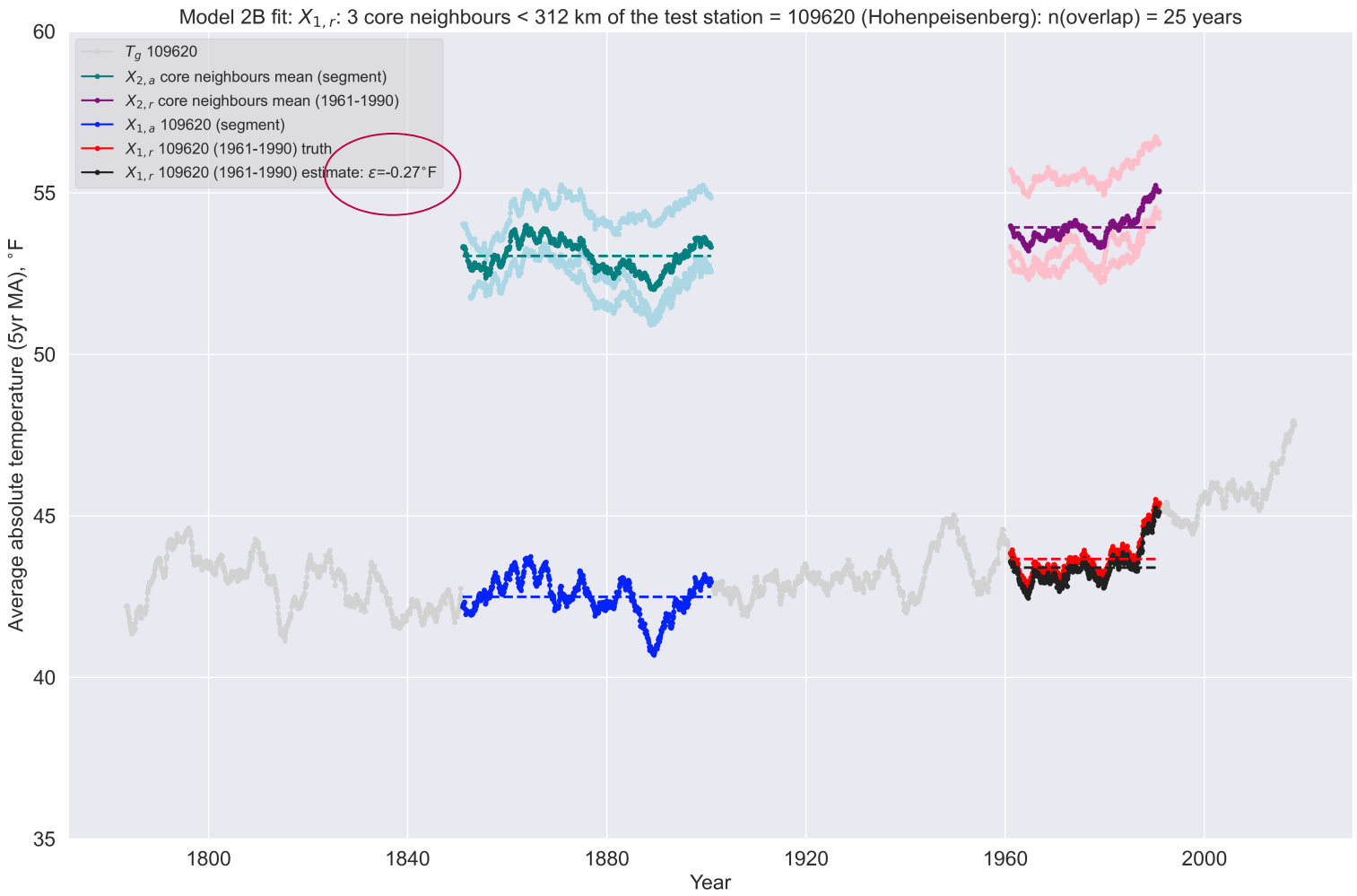
... picking the right neighbours seems to work

University of East Anglia

GlobSAT Global Surface Air Temperature

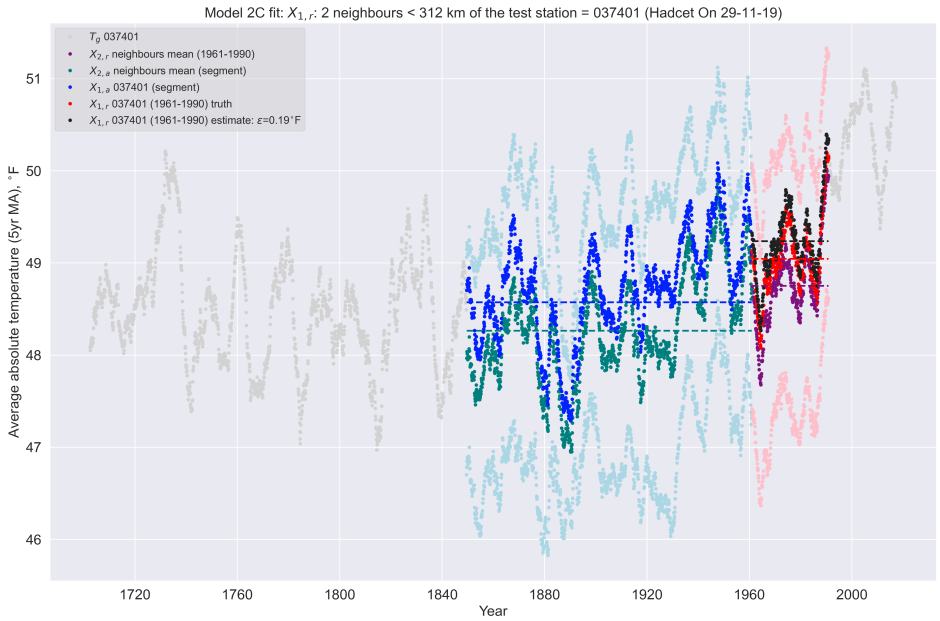
Model 2B

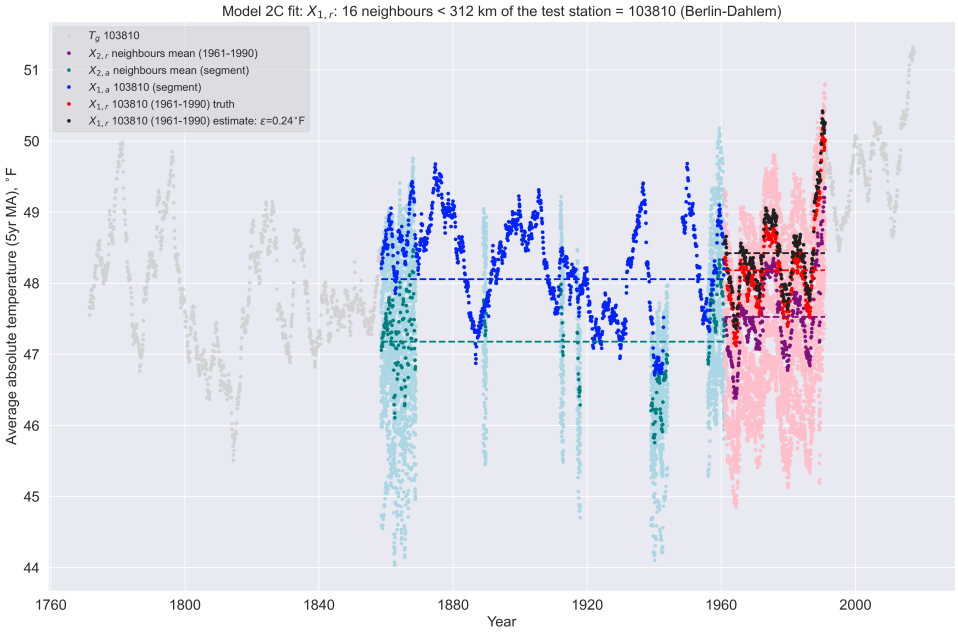
Works quite well elsewhere too which is encouraging



Model 2C

Using only co-located overlaps is working well and bodes well for Model 3





standard error on the mean estimate ~ 0.2°F

Summary

1. Using **reference series** overlapping both a segment and the climatology baseline we can estimate the **mean level** and **standard error** for a 'no baseline' series (~0.2°F error)
2. **Reanalysis** can be used for remote stations and can help grow the record in data sparse regions – which the caveat that the reanalysis also has uncertainty and is time-limited (e.g. 20CRv3 currently goes back to 1806)
3. Using an **ensemble mean of neighbouring observations** as reference works great provided that there are enough co-located overlaps with the no-baseline series
4. The optimal choice of method looks like it doesn't depend on distance but does depend on data density

Q. How do these approaches compare with PHA / Rbeast approaches ?
Q. Can we incorporate reanalysis pressure level temperatures for high elevations ?
Q. How valid / scalable is this approach to subregions and the global record ?

Suggestions more than welcome

Many thanks for listening

NOAA PSL
20CRv3 gridded monthly 2m air temperatures:
https://portal.nersc.gov/project/20C_Reanalysis/

CRU / UEA & UKMO HadObs
CRUTEM5.0.1 land surface air temperature instrumental data 1781-2020:
https://crudata.uea.ac.uk/cru/data/temperature/

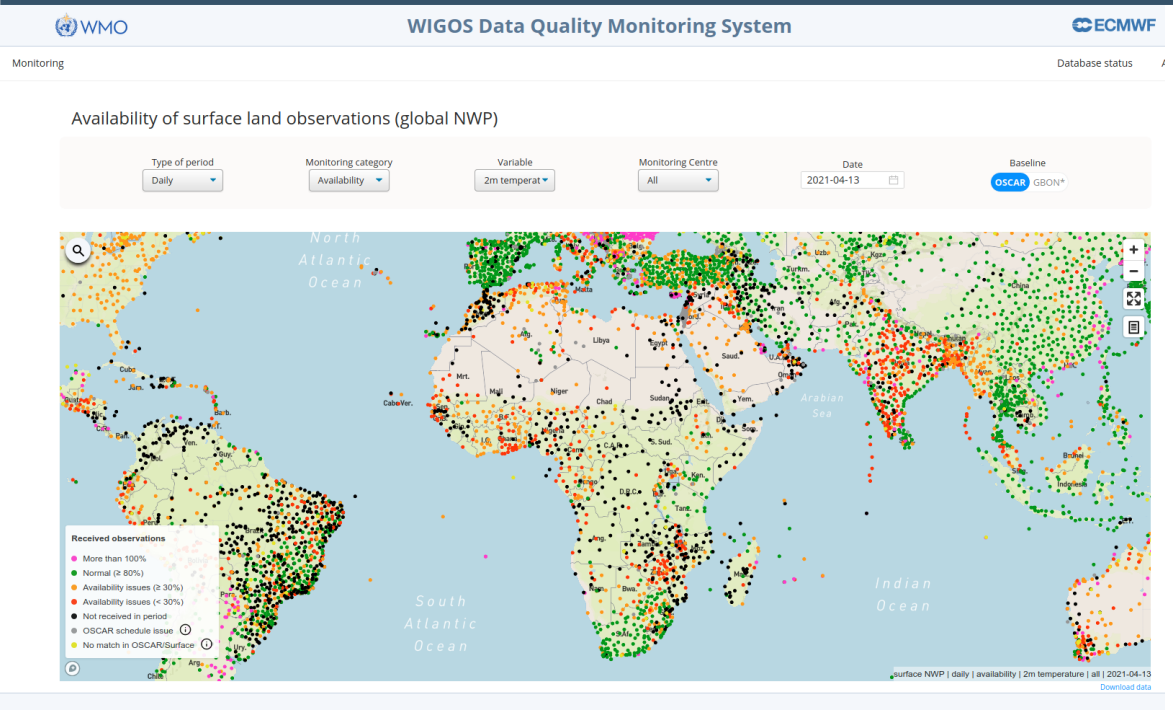
GloSAT project
https://www.glosat.org/

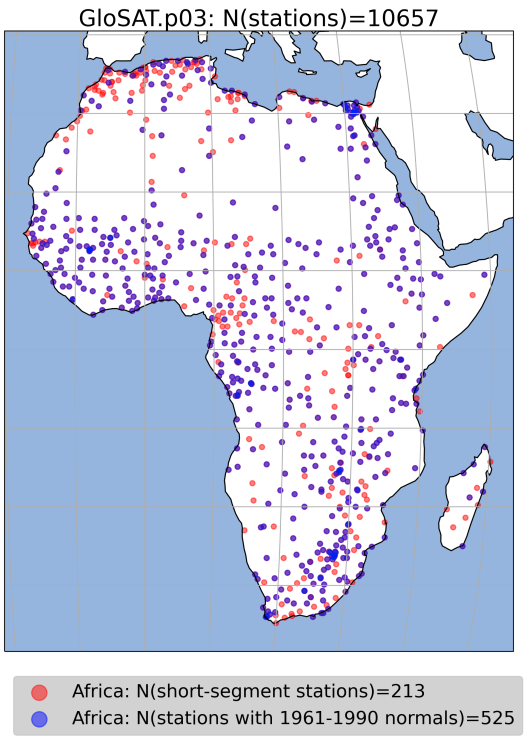
Codebase:
https://github.com/patternizer/glosat-best-fit-means
https://github.com/patternizer/glosat-short-segments
https://github.com/patternizer/glosat-new-england

EXTRA SLIDES

Africa – lots of 'no baseline' series

It's an important example of the scale of the challenge + need for accurate normals

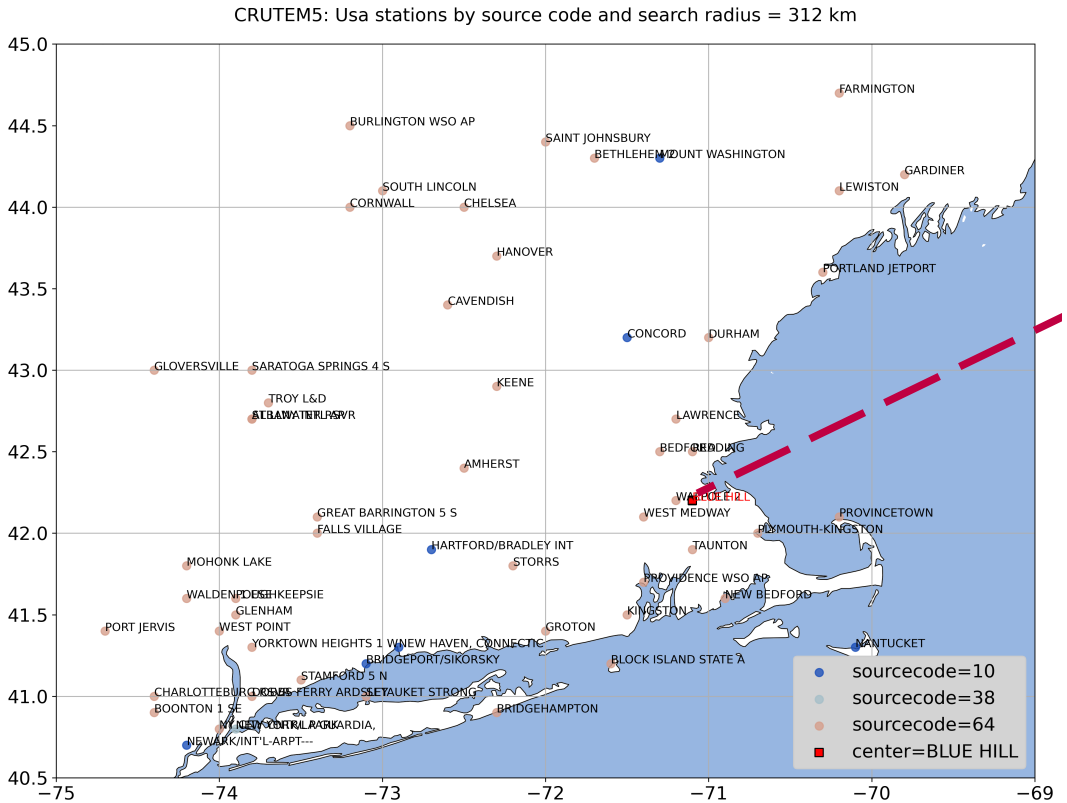




- WIGOS DQMS reports **availability issues** for stations in Algeria, Congo, Tanzania, Zimbabwe & Islands + …
- **We have lots of series to add** but they are <u>without normals</u>
- More anomaly series in data sparse regions will help support global NWP cal/val efforts

Ensemble of neighbours

312 km of Blue Hill Observatory (BHO) à la Gordon Manley



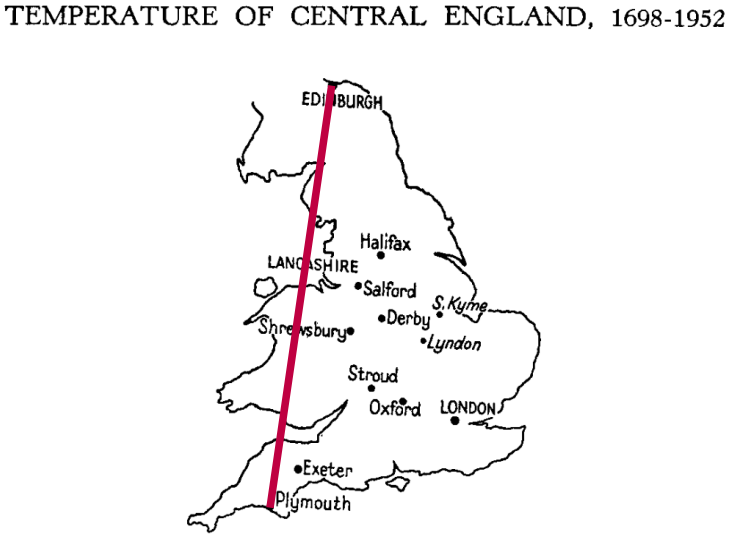


Figure 1. Location of the principal stations referred to in the text.

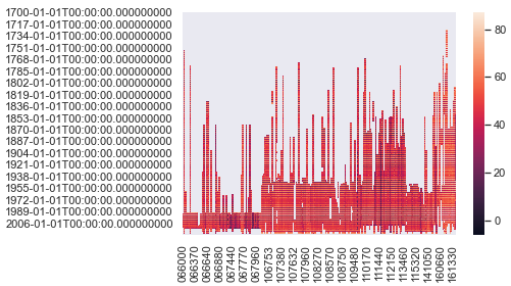
Manley, G., 1953. The mean temperature of central England, 1698–1952. Quarterly Journal of the Royal Meteorological Society, 79(340), pp.242-261.

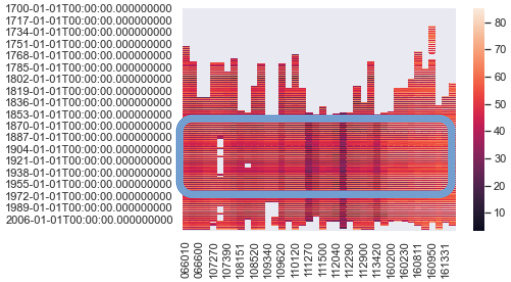
Haversine distance Edinburgh to Plymouth=624 km is used to lasso stations centered on BHO within a radius of 312 km

Model 2C

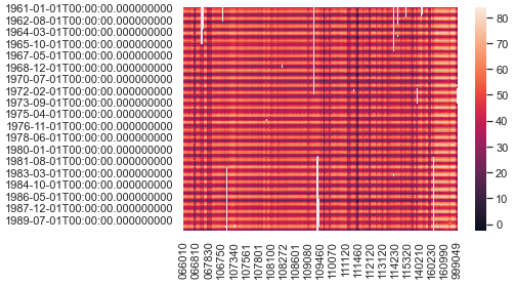
How about optimizing on the segment itself by maximizing the density of co-locations ?

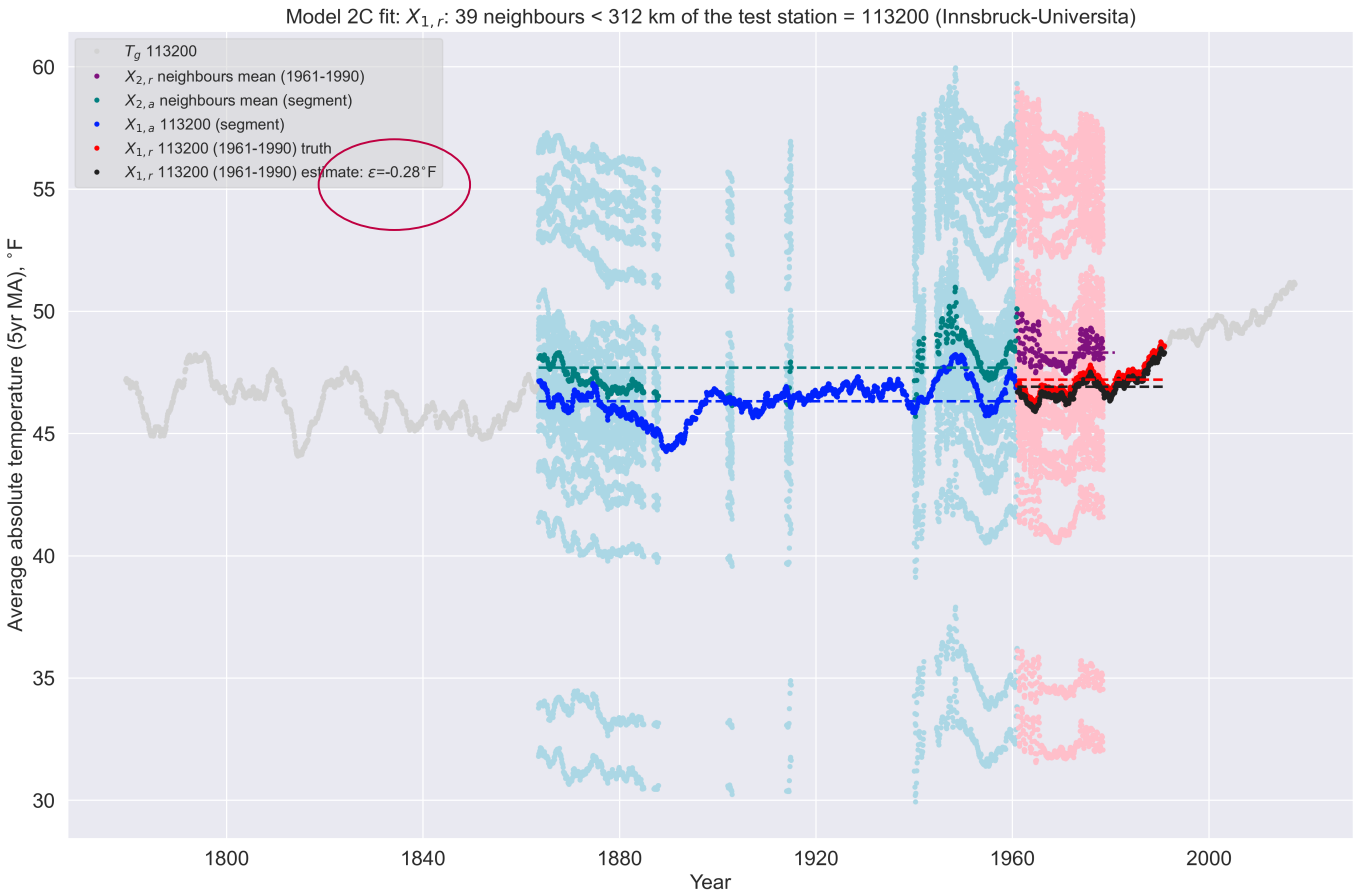
Raw ensemble





baseline





... this works without optimizing on the standard error