Properties of Antarctic Boundary Layer Thermodynamic Structure from Long-Term Radiosonde Observations

Justin Wu (1) , Yuekui Yang (2) , Manisha Ganeshan (3)

(1) Marriotts Ridge High School, 12100 Woodford Dr, Marriottsville, MD 21104

(2) NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771

(3) Morgan State University, 1700 E Cold Spring Ln, Baltimore, MD 21251

Acknowledgement:

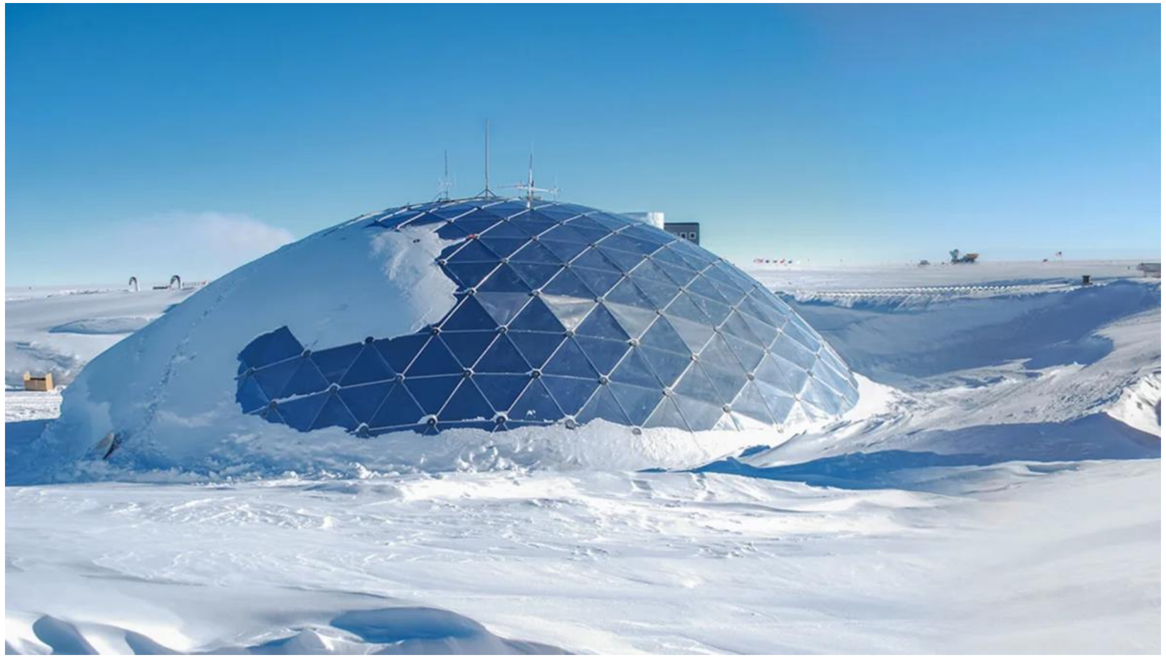
This work was supported by NASA Office of STEM engagement program in the summer of 2024

Code:

https://github.com/Justin123-Wu/AntarcticaTrend.git/

Outline

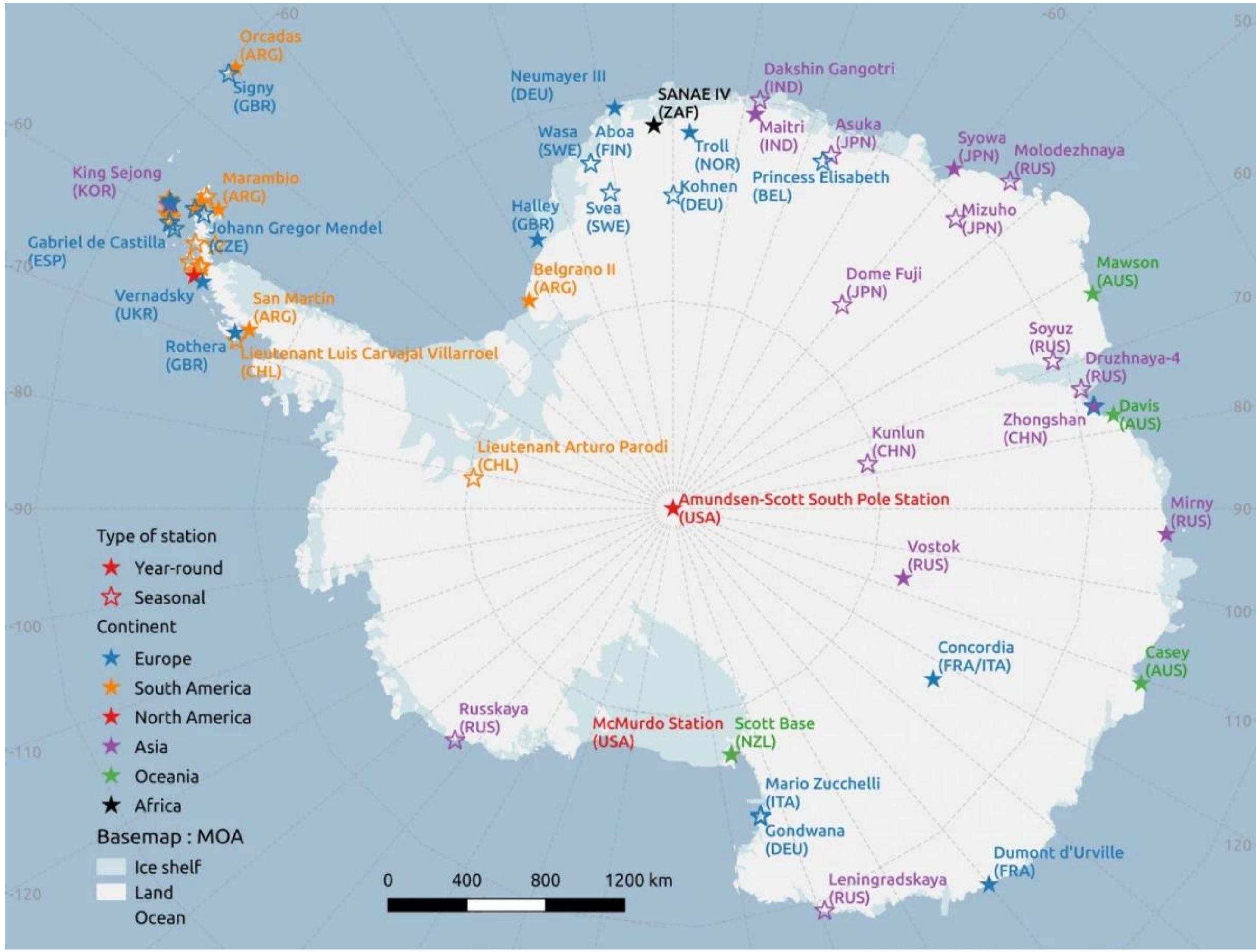
- Motivation
- Data Processing Pipeline
- Monthly Profiles of an Observation Station
  - ✓ Definition of observation
  - ✓ Temperature interpolation
  - ✓ Surface based inversion detection
  - ✓ Monthly profile format
- Monthly and Seasonal Statistics
  - ✓ Calculate monthly mean & median of boundary layer thermodynamics
  - ✓ Calculate seasonal mean & median of boundary layer thermodynamics
- Application of Trend Decomposition & Trend Analysis on Boundary Layer Thermodynamics
  - ✓ Concept of trend decomposition
  - ✓ Calculate monthly and seasonal trends
- Results
- Summary and Future Work



(Image courtesy of https://www.antarcticacruises.com/guide/research-stations-in-antarctica)

Motivation

- Why Antarctica?
  - ✓ The Antarctic Ice Sheet accounts for ~ 70% of the Earth's freshwater. It is losing mass at an accelerated pace, significantly contributing to the rise in sea level.
- Why radiosonde observations?
  - ✓ Boundary layer processes play a key role in the changes observed in Antarctica and radiosonde datasets are best suited for the analysis of boundary layer structure.
- Observation Stations in Antarctica
  - ✓ Scottish National Antarctic Expedition built 1st meteorological station in 1903
  - ✓ 70+ research stations
  - ✓ 20 countries
  - ✓ Data has geographical diversity



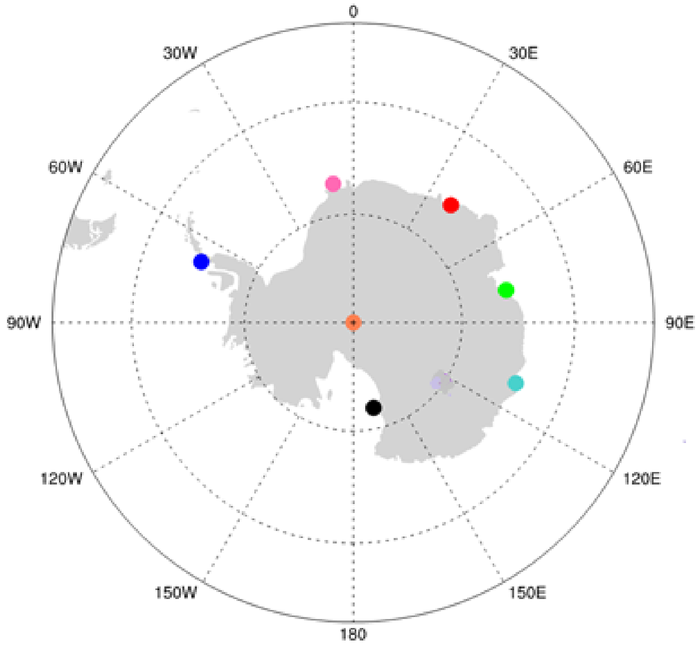
(Image courtesy of https://www.antarcticacruises.com/guide/research-stations-in-antarctica)

Radiosonde Data Sets

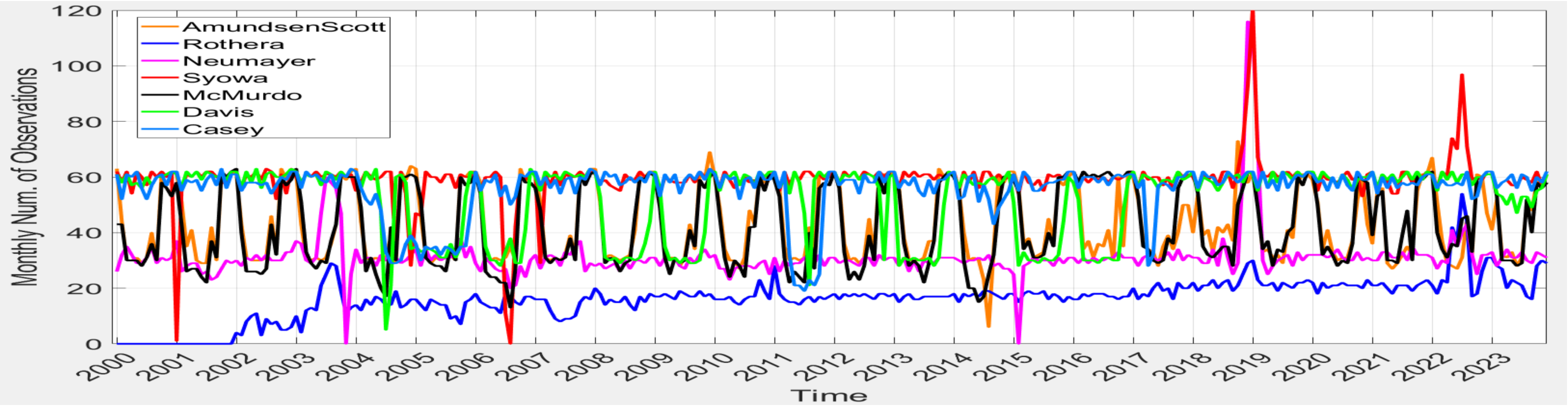
- **Source: Integrated Global Radiosonde Achieve (IGRA)**

  https://www.ncei.noaa.gov/data/integrated-global-radiosonde-archive/access/data-por/

- **Observation**
  - ✓ An observation is a set of measurements collected from the ground up to the highest point reached by a weather balloon during flight, including temperature, pressure, relative humidity, and wind speed.
  - ✓ Observations from 7 stations are selected for the analysis based on their long term data availability (2000 to present).



- **Monthly observations**



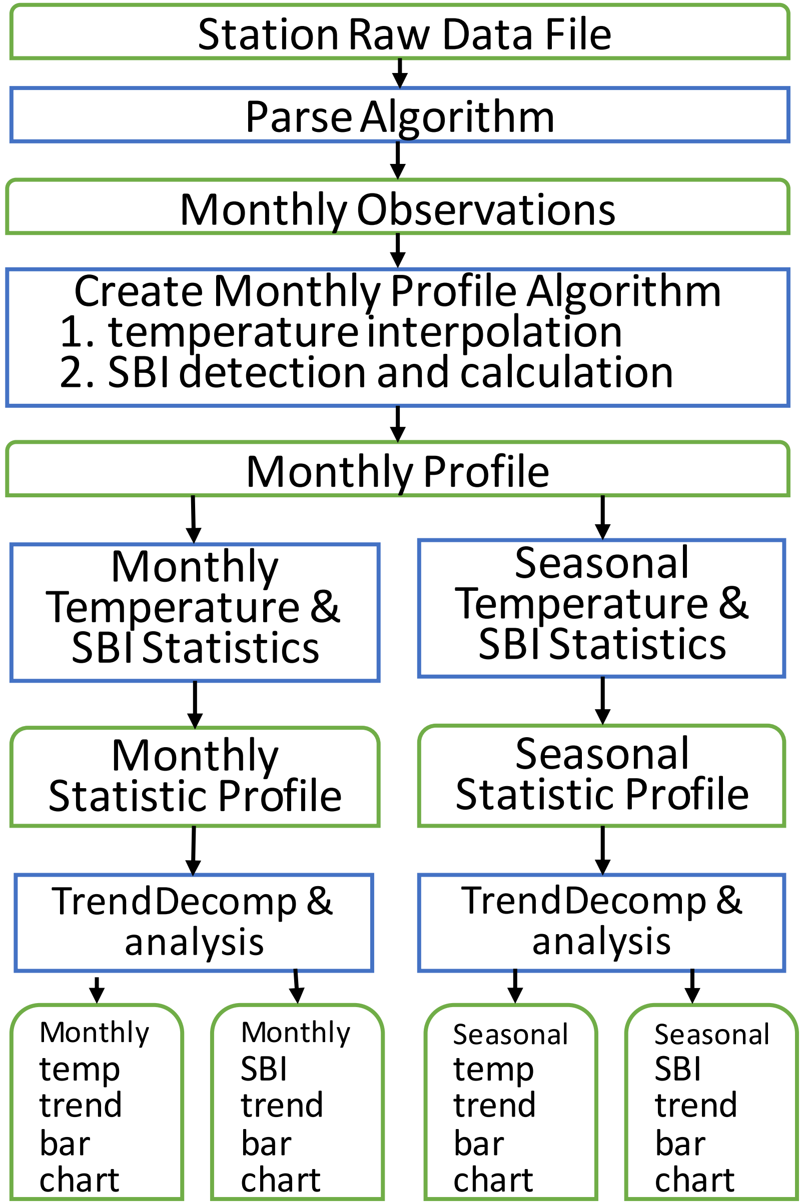
Radiosonde Data Sets (cont.)

- **Data Overview**

Station Name	Alt (m)	Observation date range	Max Altitude (m)	Number of Obs.	File name	File size (KB)
Amundsen - Scott	2835	01/01/1961 - 07/21/2024	44,710	31,223	AYM00089009-data.txt	79,016
Rothera	16.0	01/07/2002 - 07/19/2024	38,703	4,850	AYM00089062-data.txt	29,126
Neumayer	43.0	05/03/1984 – 07/172024	38,732	13,950	AYM00089002-data.txt	56,322
Syowa	29.0	01/01/1959- 07/21/2024	46,995	39,317	AYM00089532-data.txt	85,369
McMurdo	10.0	04/15/1956- 07/22/2024	52,224	31,750	AYM00089664-data.txt	82,055
Davis	27.0	02/12/1957– 07/21/2024	48,566	42,266	AYM00089571-data.txt	109,220
Casey	32.0	03/14/1957- 07/21/2024	47,906	44,448	AYM00089611-data.txt	119,114

(Image courtesy of https://en.wikipedia.org/wiki/Amundsen%E2%80%93Scott_South_Pole_Station)

Data Processing Pipeline



- Raw data file
  - ✓ All observations encoded into single file
- **Monthly observation**
  - ✓ yyyy-mm-observation.csv
  - ✓ Human-friendly to read
  - ✓ Helps debug the parse algorithm
- **Monthly profile**
  - ✓ yyyy-mm-profile-linear.csv
    yyyy-mm-profile-spline.csv
  - ✓ Temperature interpolation with respect to height
  - ✓ Surface based inversion detection and SBI feature calculation
- **Monthly/seasonal statistics**
  - ✓ monthlyStatistics-2000-2024-linear.csv
    seasonalStatistics-2000-2024-linear.csv
  - ✓ Monthly/seasonal mean, median are calculated for distinct features
  - ✓ Single file for multiple years given interpolation method
- **Final results presented as bar charts**

Parse Raw Data into Array of Observations

- **Raw Data**
  - ✓ Contains all observations from Day 1 to July 2024.
  - ✓ Text file with special format
- **Observation**
  - ✓ Head info: Id, year, month, day, hour, lat,lon,base altitude
  - ✓ Height related quantities: lvlType1, lvlType2, etime, pressure, pFlag gph, zFlag, temperature, tFlag, rh, dpdp, windir, wspd, altitude
- **The Parse Instructions & Algorithm**
  - ✓ igra2-data-format.txt (https://www.ncei.noaa.gov)
  - ✓ Hypsometric equation

$$z_2 = z_1 + \frac{R\overline{T}_v}{g}\ln\left(\frac{p_1}{p_2}\right)$$

z_1, z_2 – height at previous and current points in meters (m)
p_1, p_2 – pressure $(\mathrm{P_a})$at z_1, z_2
$R = 8.3144, (\mathrm{JK^{-1}mol^{-1}})$, specific gas constant for dry air
$\overline{T}_v$ = mean temperature in Kelvin (K) between z_1 and z_2
g = gravitational acceleration $(\mathrm{m/s^2})$

(Image courtesy of https://www.nasa.gov/centers-and-facilities/wallops/balloons-on-ice-final-flight-launches-in-antarctica/)

Temperature Interpolations

- **Interpolation goal:** Given an observation with *n*-point of measurements, (h_i, T_i) $(i = 1, \dots, n)$, calculate temperature of any height, i.e. T(h), where $h \in [h_1, h_n]$

- **Linear interpolation**

$$T(h) = \frac{T_1(h_2 - h) + T_2(h - h_1)}{h_2 - h_1}, h \in [h_1, h_2]$$

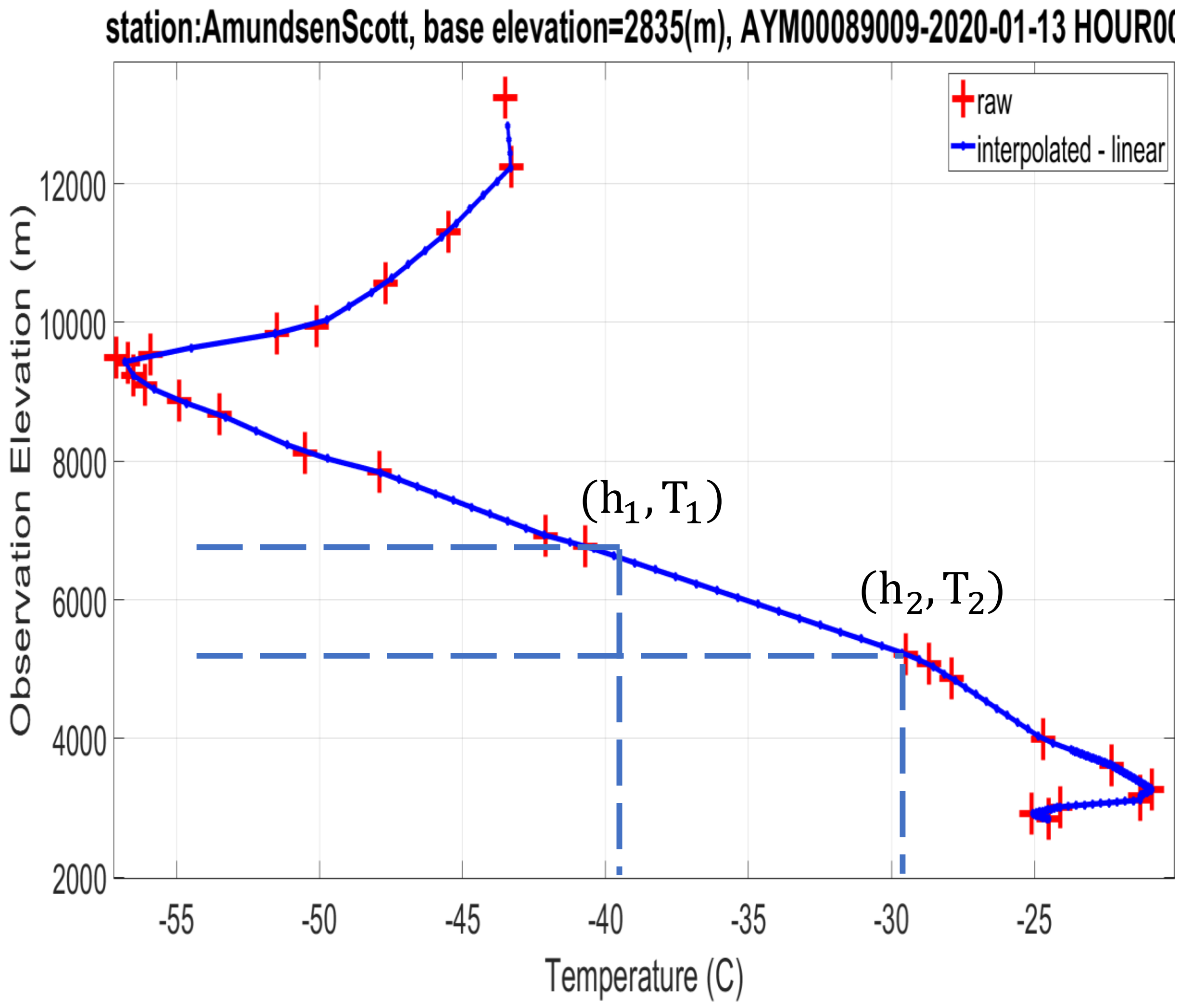
- **Multiple-observation scenario**

  If one day has *m* observations, we independently interpolate them and get $T^{\{k\}}(h), (k = 1, m)$ first, then average them to represent that day's interpolated temperature, i.e.,

$$T(h) = \frac{1}{m} \sum_{k=1}^{m} T^{\{k\}}(h)$$

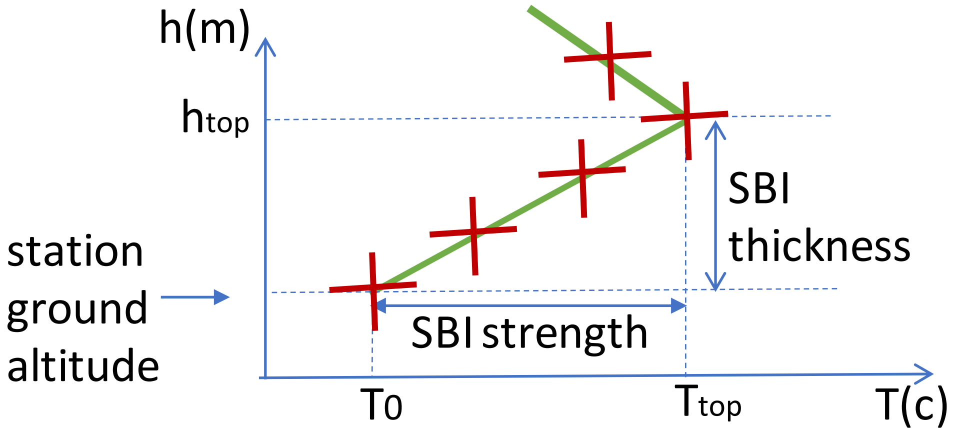
- **Non-uniform height steps**

$$\Delta h = \begin{cases} 10 & h \in [0,1000]\ m \\ 100 & h \in [1100,5000]\ m \\ 200 & h \in [5200,10000]\ m \end{cases}$$



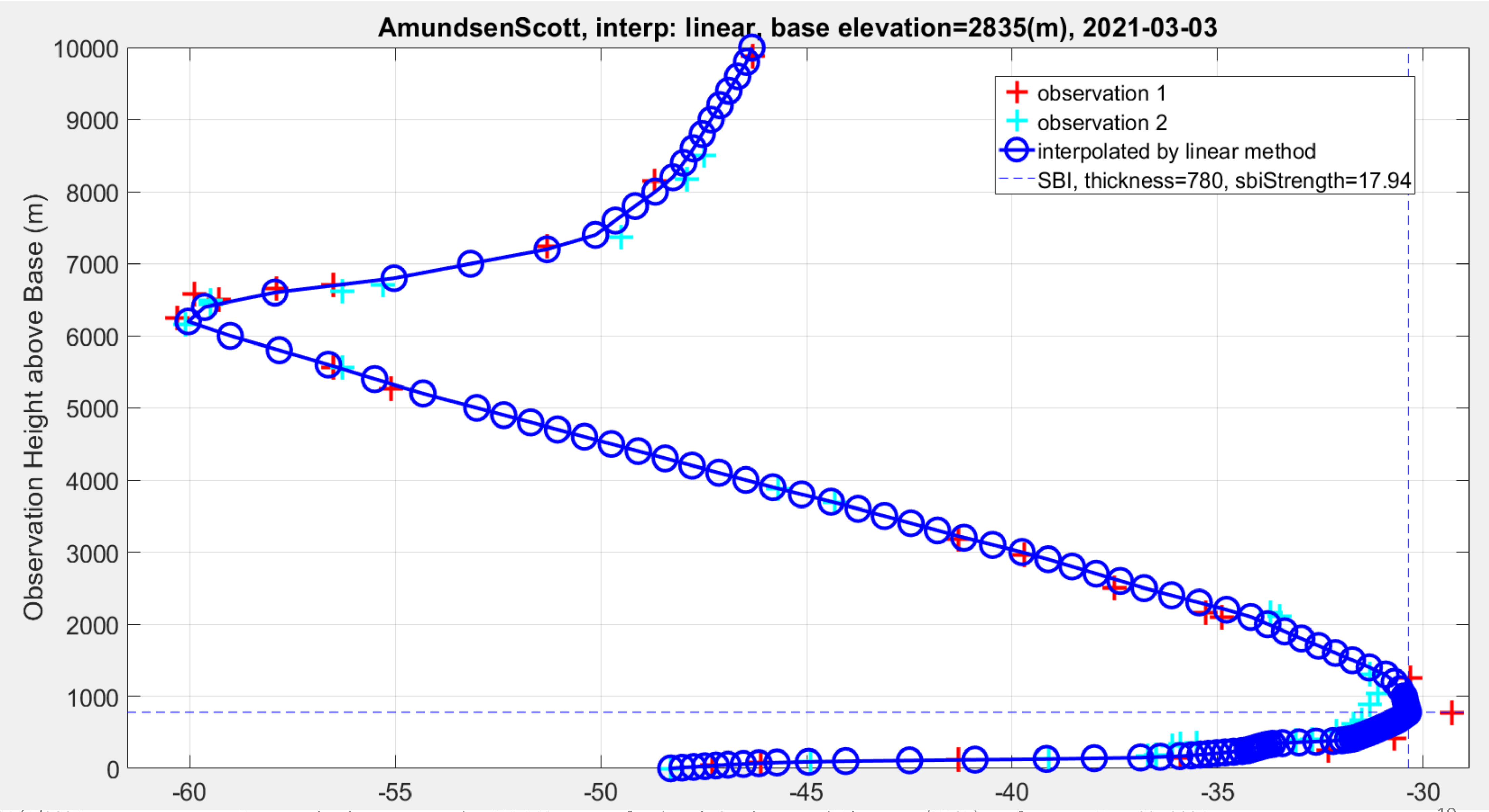
Surface based inversion

- **SBI Concept**
  A temperature inversion is a layer in the atmosphere in which air temperature increases with height.

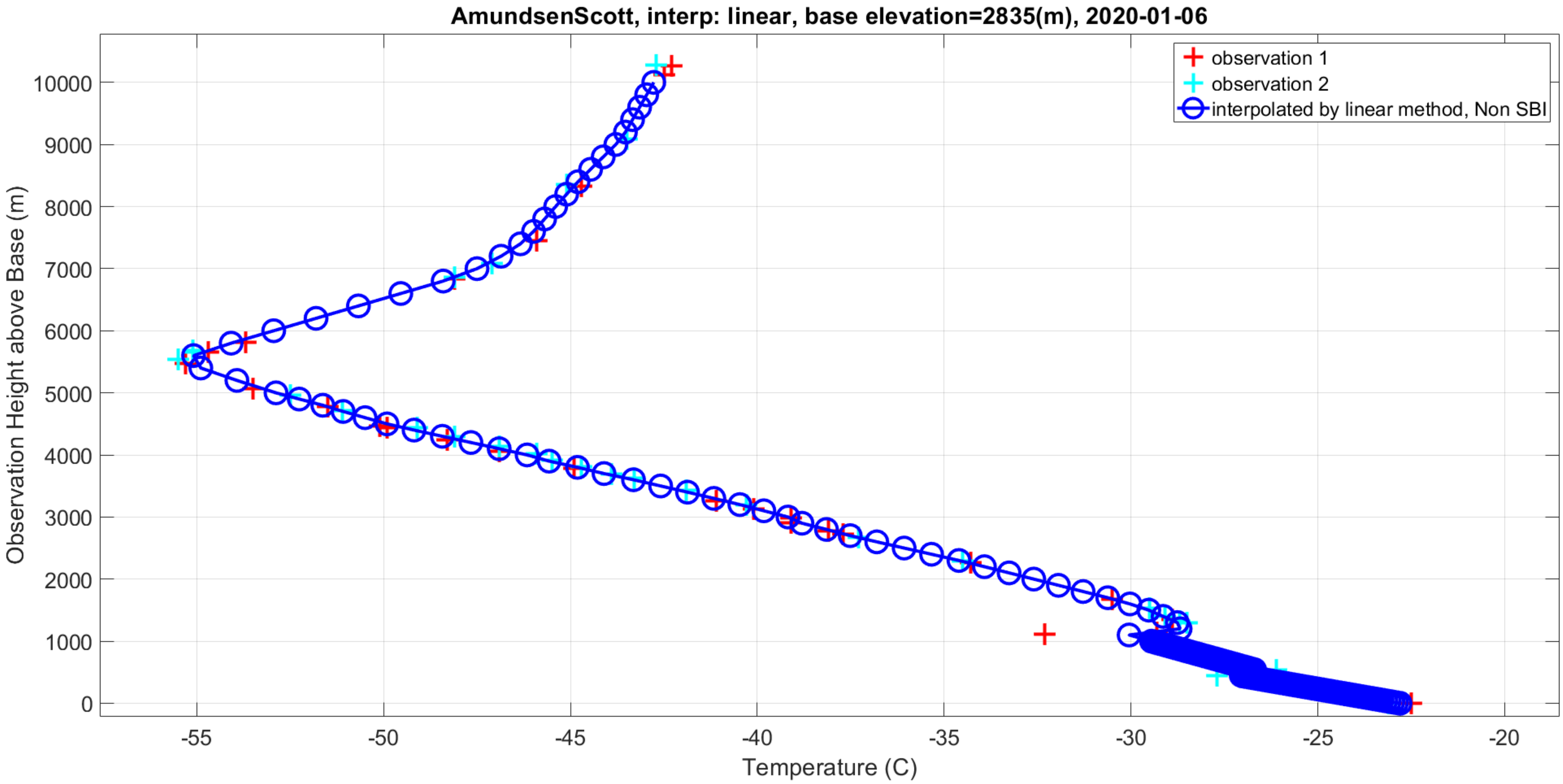


- **SBI Detection**
  - ✓ Method 1: start from the interpolated temperature $T(h=0)$, then check $T(h+\Delta h * i\ )$ to see if there is a turning point at which
    $$T(\Delta h * (i-1)) < T(h+\Delta h * i)\ (i = 1, \dots)$$
    If a turning point is present, there is an SBI; otherwise non-SBI. If present, we mark the turning point as $\left(h_{top}, T_{top}\right)$
  - ✓ Method 2: directly from the observation data, find the maximum temperature point $\left(h_{top}, T_{top}\right)$, and compare it with the nearby ground point $(h=0, T_0)$. if $h_{top} > 0$, it is an SBI, otherwise non-SBI.
- **SBI strength and thickness**
  $$\text{SBI Strength} = T_{top} - T_0,\ \text{SBI thickness} = h_{top} - h_0,$$
- **Implementation edge cases**
  - ✓ T0 is NaN (unknown, sbiFlag=-1)
  - ✓ Cannot locate turning point (non-SBI, sbiFlag=0)
  - ✓ Can detect turning point (has SBI, sbiFlag=1)

Temperature Interpolation and SBI Detection Example



Temperature Interpolation and Non-SBI Example



Create Monthly Profile

- Matrix format
  - ✓ An N x 32 matrix
  - ✓ Stored at ***data/processed/stationName/2024-07-profile-linear.csv***
  - ✓ Row [1:N-7]: interpolated temperature vs. height for day 1 through 31 of the month, (nan for unavailable data)
  - ✓ Row [N-6, N] : T200-T0, T500-T0, and SBI info of the corresponding date

note	row #	height(m)	day1	…	day31
T0: Temperature at ground (C)	1	2835	-59.50	…	7.445
T50: Temperature above 50m of ground (C)	2	2845	-56.67	…	5.328
…	…				
T10000: Temperature above 10km of ground (C)	N-7	12835	-75.30	…	2.207
T200-T0: Temperature difference of T200 & T0	N-6	nan	x	…	x
T500-T0: Temperature difference of T500 & T0	N-5	nan	x	…	x
HasSBIFlags (1 - sbi, 0 - nonSBI, -1 - unknown)	N-4	nan	1	…	0
SBI strength (C)	N-3	nan	x	…	x
SBI thickness(m)	N-2	nan	x	…	x
SBI inversion Temp (C) (for debug & visualization purpose)	N-1	nan	x	…	x
SBI inversion height (m) (for debug & visualization purpose)	N	nan	x	…	x

Interpolated raw data

To do analysis

Monthly Statistics

- **Matrix format**
  - ✓ The monthly statistics is an n x 19 matrix; each row represents one month of mean (μ), standard deviation (σ), and median (med) of T0, T200, …, SBI depth. It also counts the number of SBIs and non-SBIs for a given month.

Year	Month	T0			T200			T500			SBI strength			SBI depth			# SBIs	# Non-SBIs
		μ	σ	med	μ	σ	med	μ	σ	med	μ	σ	med	μ	σ	med		
1970	01																	
1970	02																	
…	…																	
2024	07																	

- **Calculation**
  - ✓ Given a year, month, and interpolation method name, load its corresponding monthly profile matrix ***A*** from *data/processed/stationName/yyyy-mm-profile-interpoMethod.csv*
  - ✓ Calculate statistics for *T0, T200,T500, SBI strength, SBI depth* from *A(1, 2:end) , A(N-6, 2:end),, A(N-5, 2:end), A(N-3, 2:end), , A(N-2, 2:end),* respectively. Counts # SBIs and non-SIBs from *A(N-4, 2:end),* where N is total # of rows of *A*.
- **Save Result**
  - ✓ The matrix of monthly statistics is saved into a file titled with year range and interpolation method, for instance: *data/processed/stationName/monthlyStatistics-1970-2024-linear.csv*

Seasonal Statistics

- **Seasonal code and its months given year *y***

Season code	season	
1	Summer	Dec of year y-1, Jan, Feb of y
2	Fall	Mar, Apr, May of y
3	Winter	Jun, Jul, Aug of year y
4	Spring	Sep, Oct, Nov of year y

- **Calculation**
  - ✓ Similar to the monthly statistics, we calculate the three months statistics given a year, season code and interpolation method, subsequently storing the results into the following n x 19 matrix.

Year	Season code	T0			T200			T500			SBI strength			SBI depth			# SBIs	# Non-SBIs
		μ	Σ	med	μ	σ	med	μ	σ	med	μ	σ	med	μ	σ	med		
1970	1																	
1970	2																	
1970	3																	
1970	4																	
…	…																	
2024	2																	

- **Result file**
  - ✓ data/processed/stationName/seasonalStatistics-1970-2024-linear.csv

TrendDecomp – Introduction

- Assumption

$$x(t) = trend(t) + s1(t) + s2(t) + \epsilon(t)$$

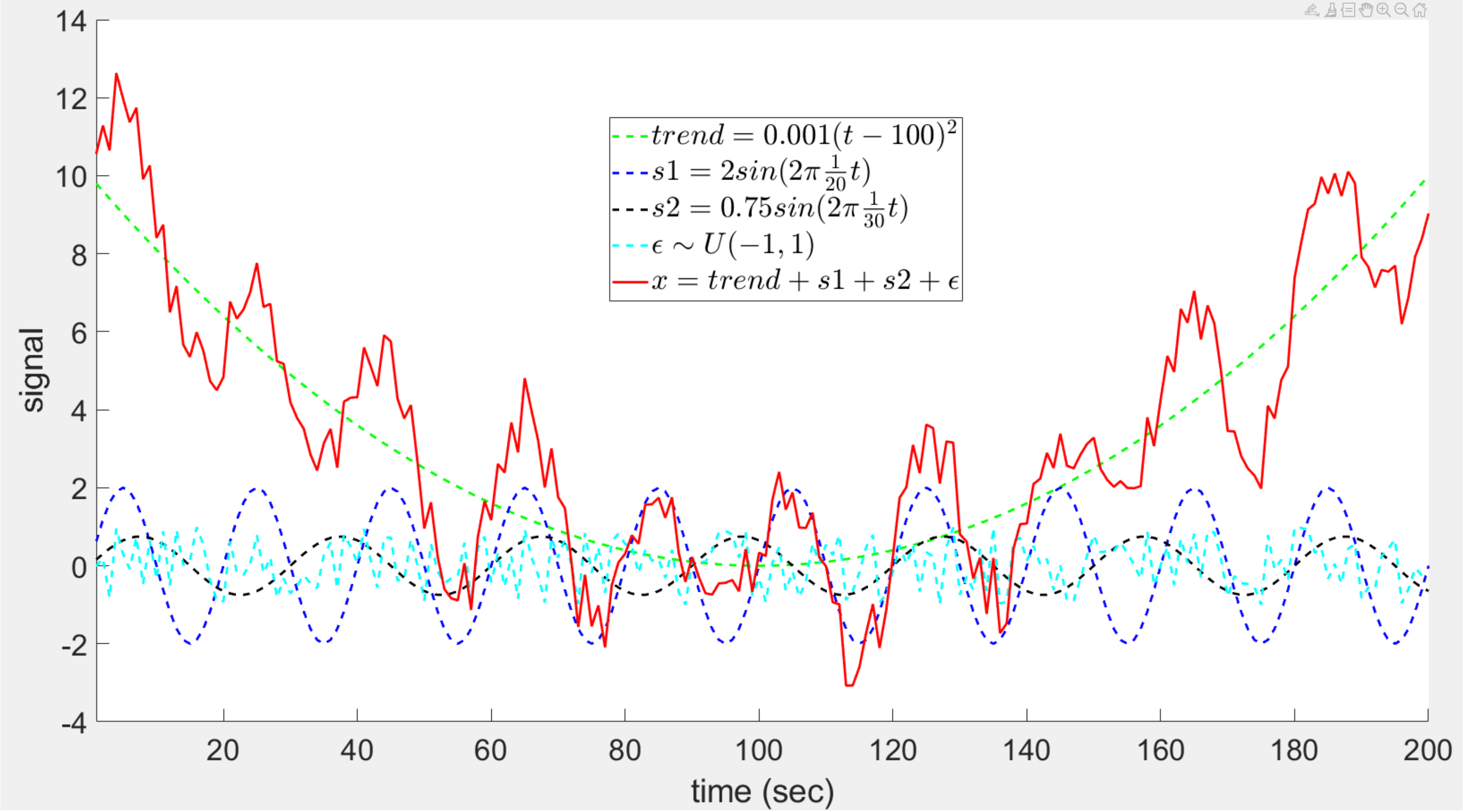
- Example

$$trend(t) = 0.001 * (t - 100)^2$$

$$s1(t) = 2\sin\left(2\pi\frac{1}{20}t\right)$$

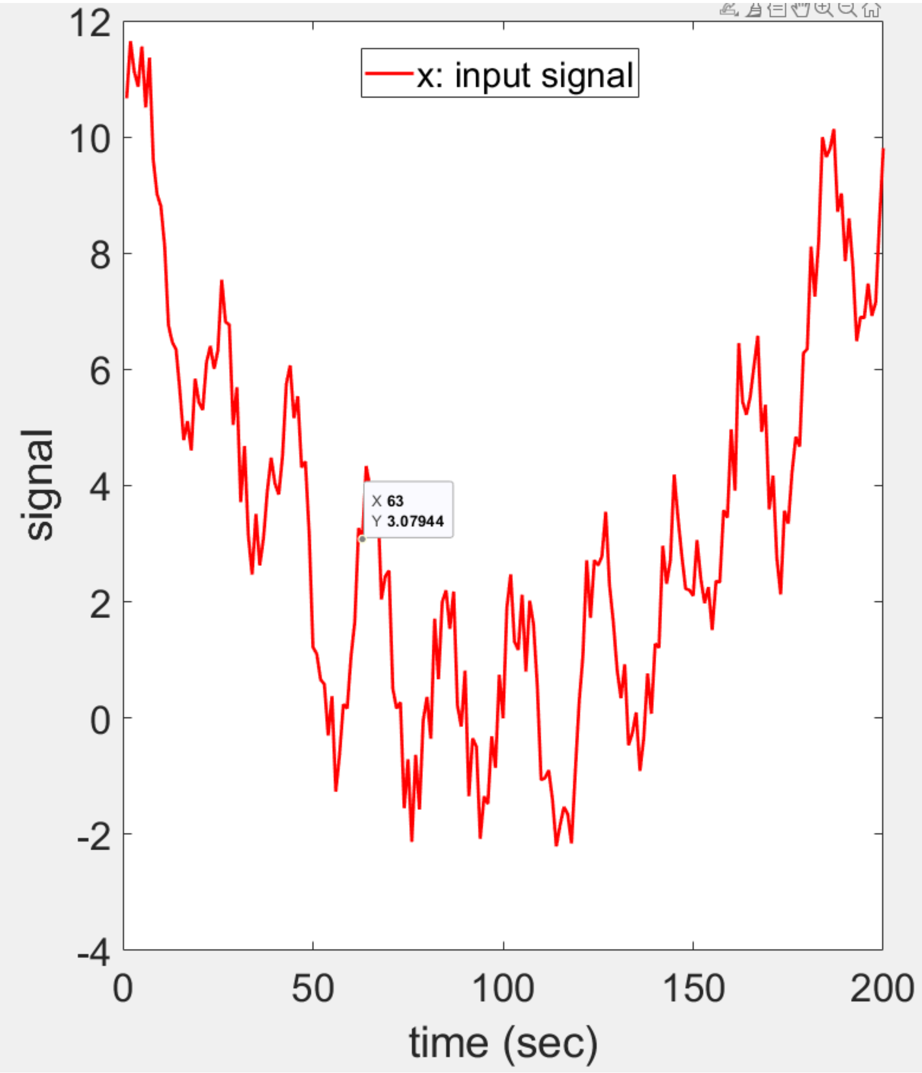
$$s2(t) = 0.75\sin\left(2\pi\frac{1}{30}t\right)$$

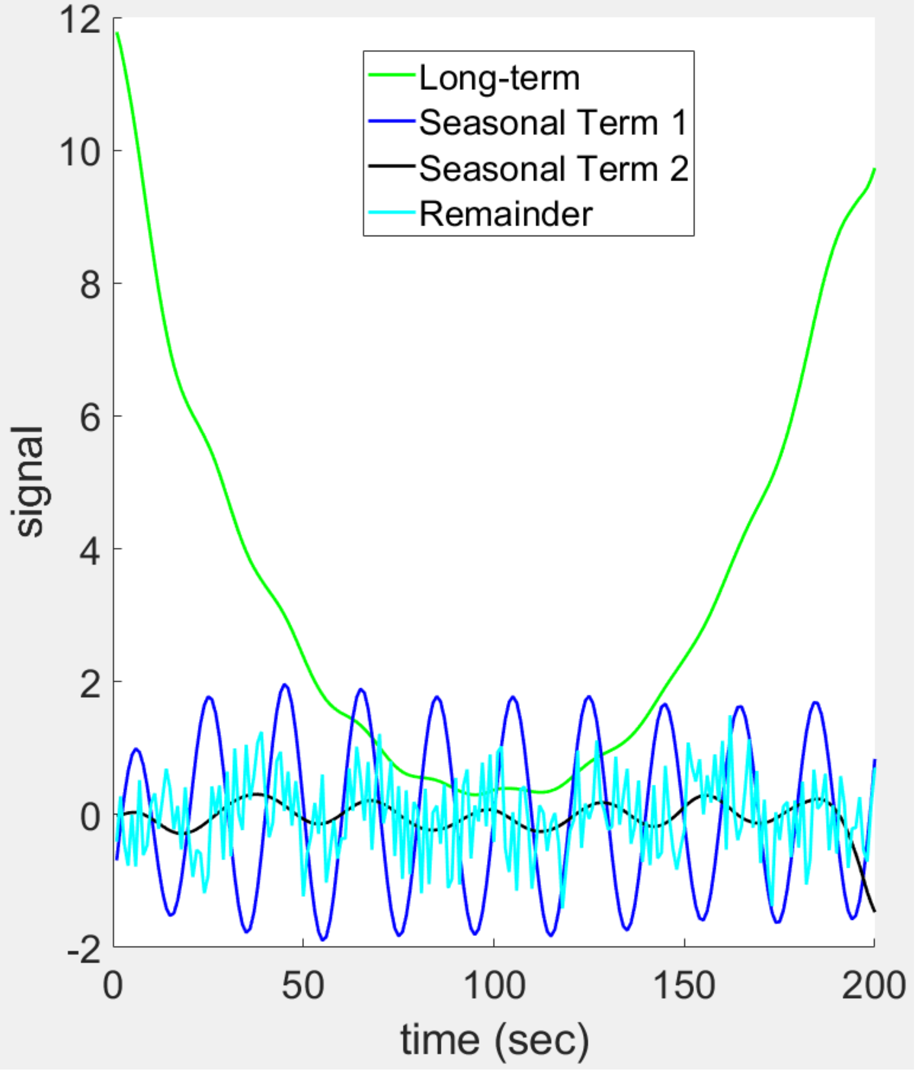
$$\epsilon\,(t) \sim \text{uniform}(-1{,}1)$$



TrendDecomp

- Solve the inverse problem: given x(t) what does its trend and seasonal components look like?
- MATLAB function: [lt,st1,st2,r] = *trenddecomp*(x);

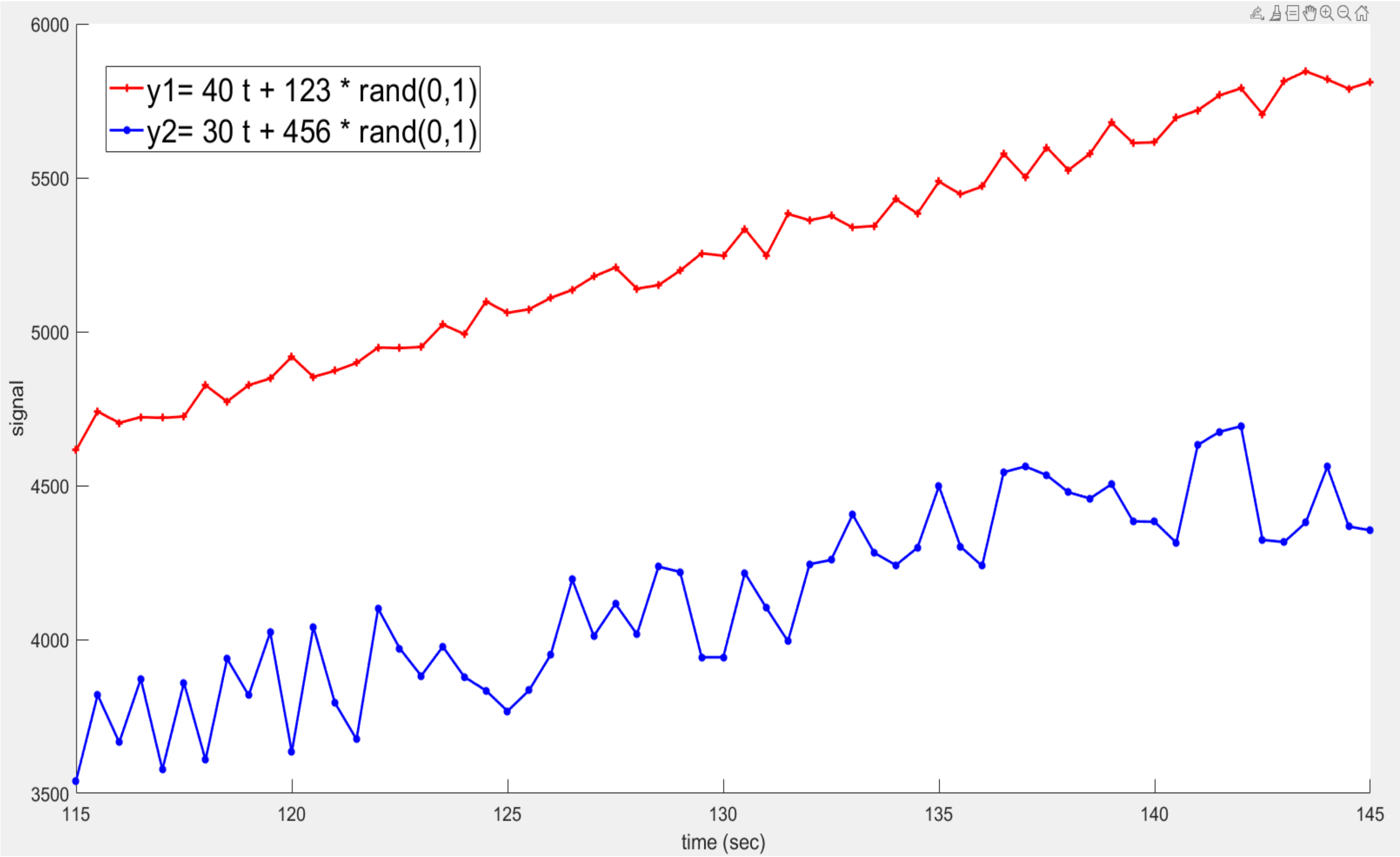




ChadGreeneTrend()

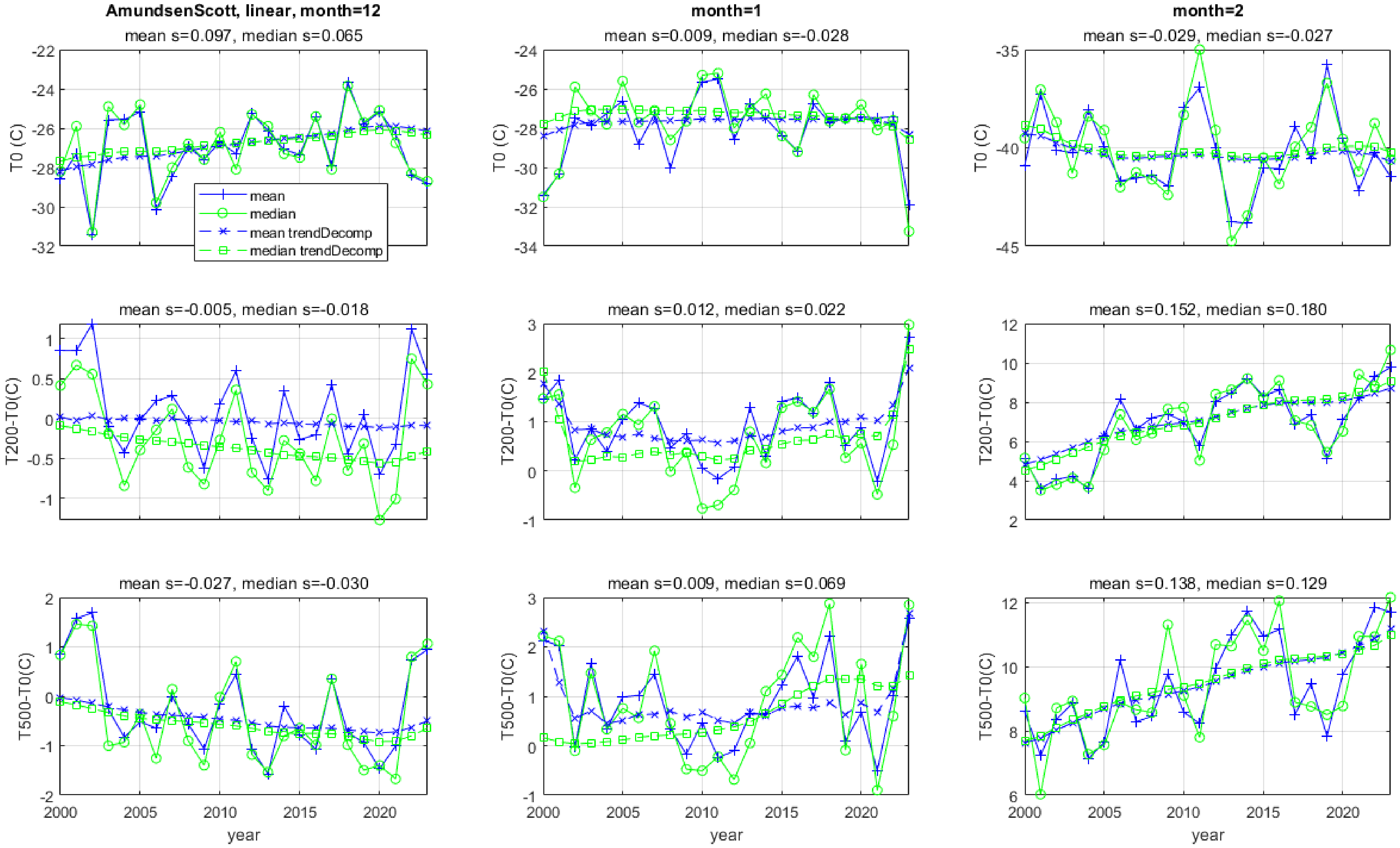
- An algorithm which estimates the slope of a signal using linear least-squares
  (https://www.mathworks.com/matlabcentral/fileexchange/46363-trend)

```
fs = 2;      % sampling frequency
dt = 1/fs; % sampling time interval

t = (115:dt:145)';
y1 = 40*t + 123*rand(size(t));
y2 = 30*t + 456*rand(size(t));

s = chadGreeneTrend([y1, y2], fs);
disp(s)

>> s = [39.8624   29.9249]
```

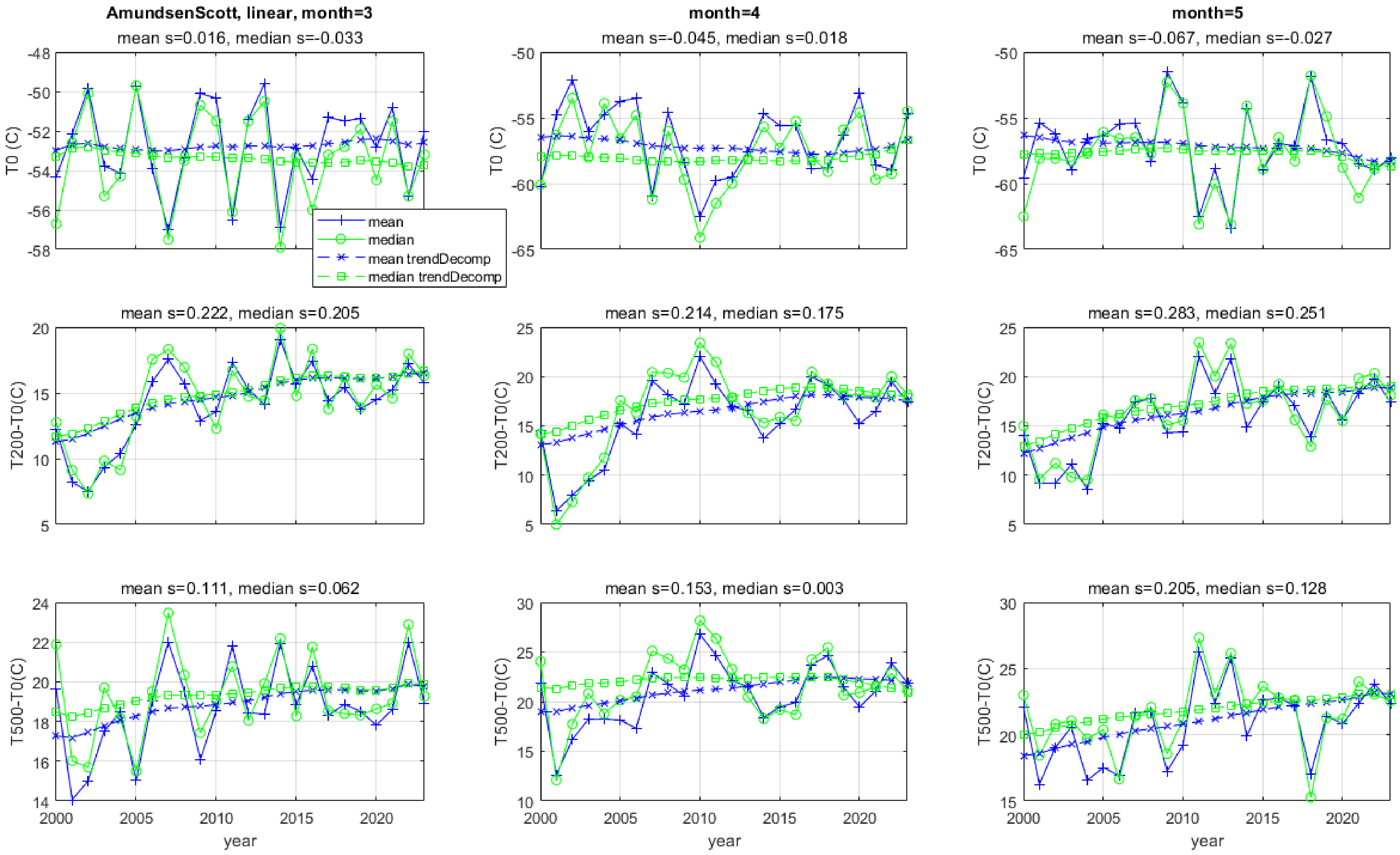


Monthly Statistics and Trends Results

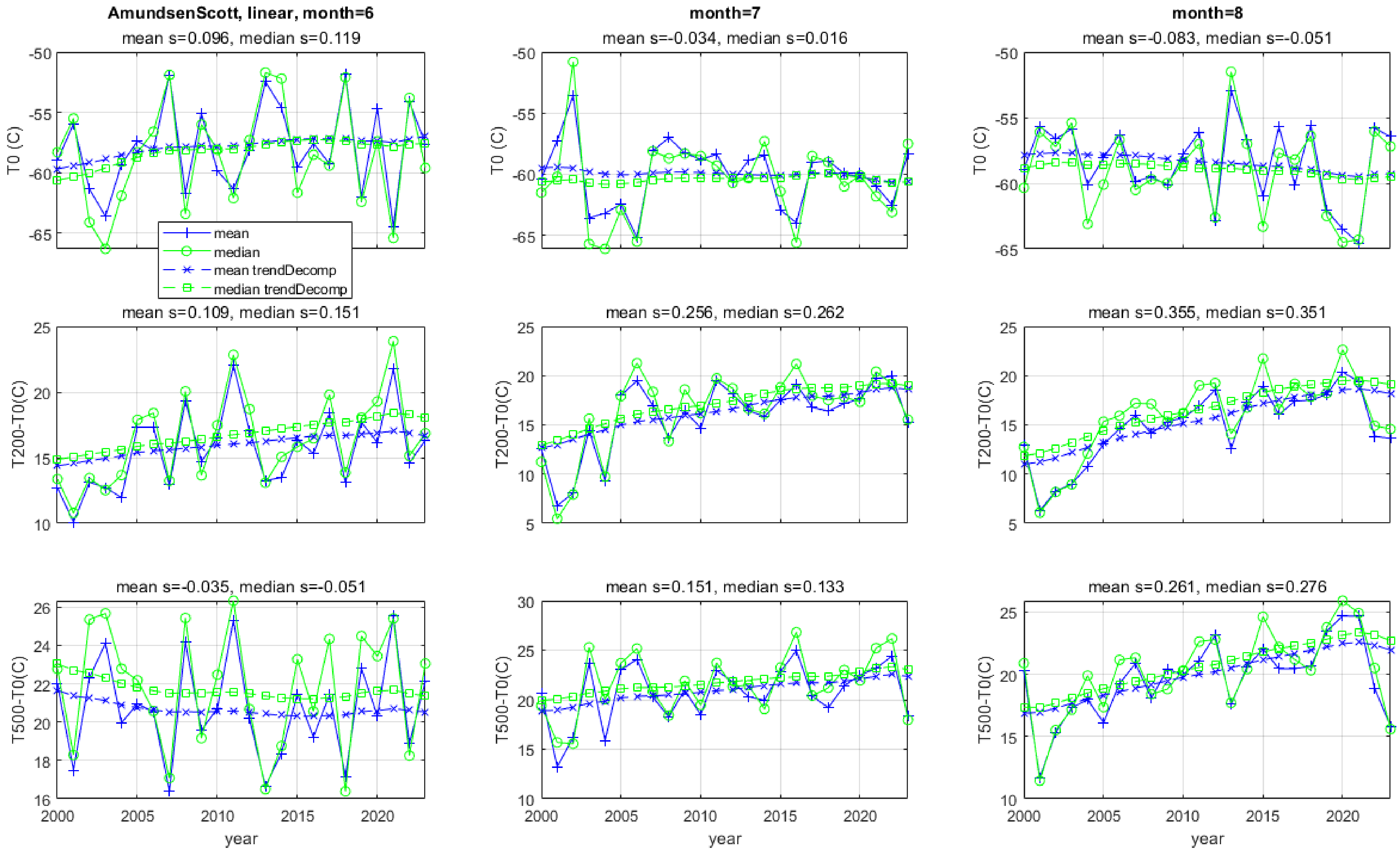
Amundsen-Scott Monthly Temperature Stats: Summer



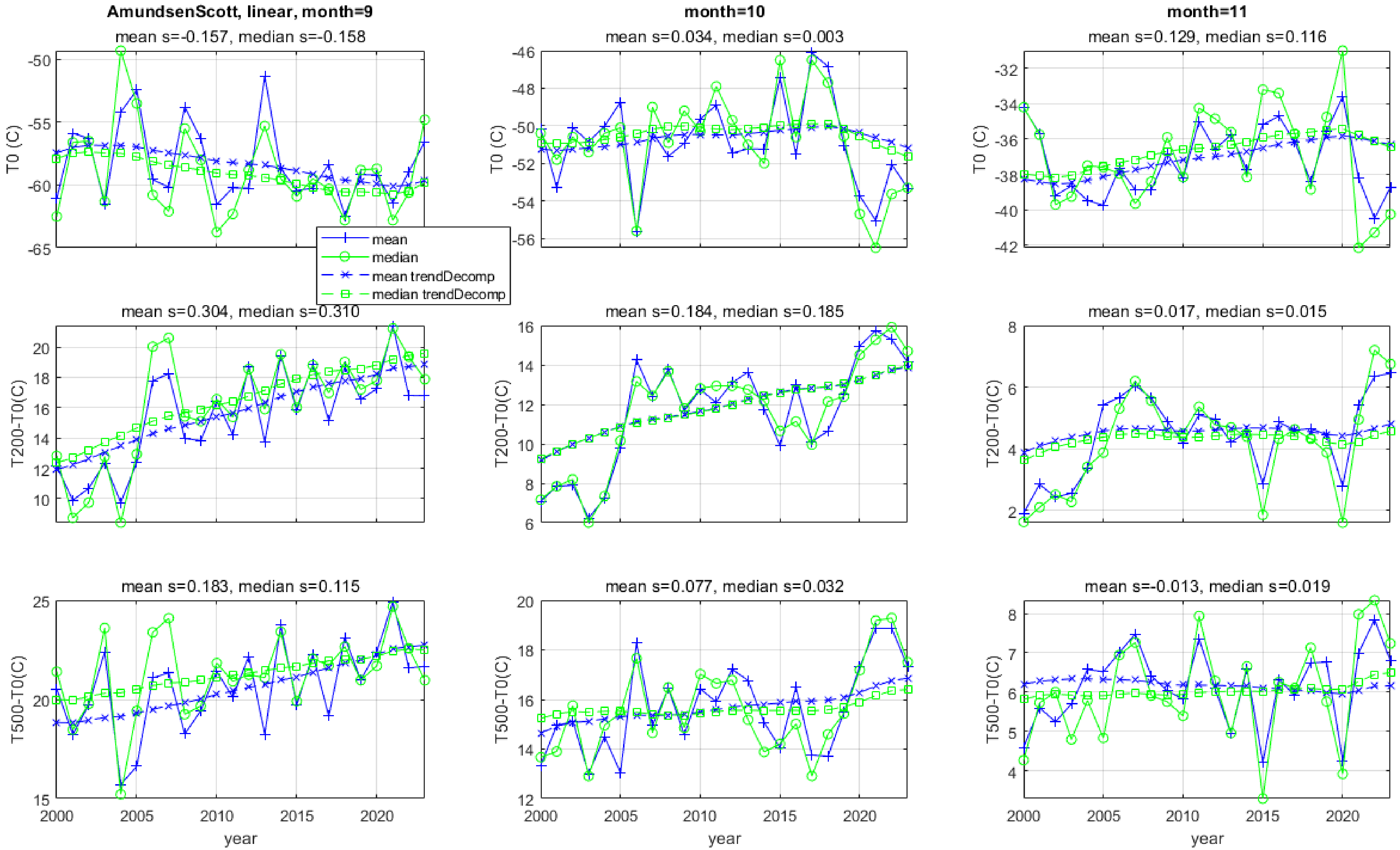
Amundsen-Scott Monthly Temperature Stats: Fall



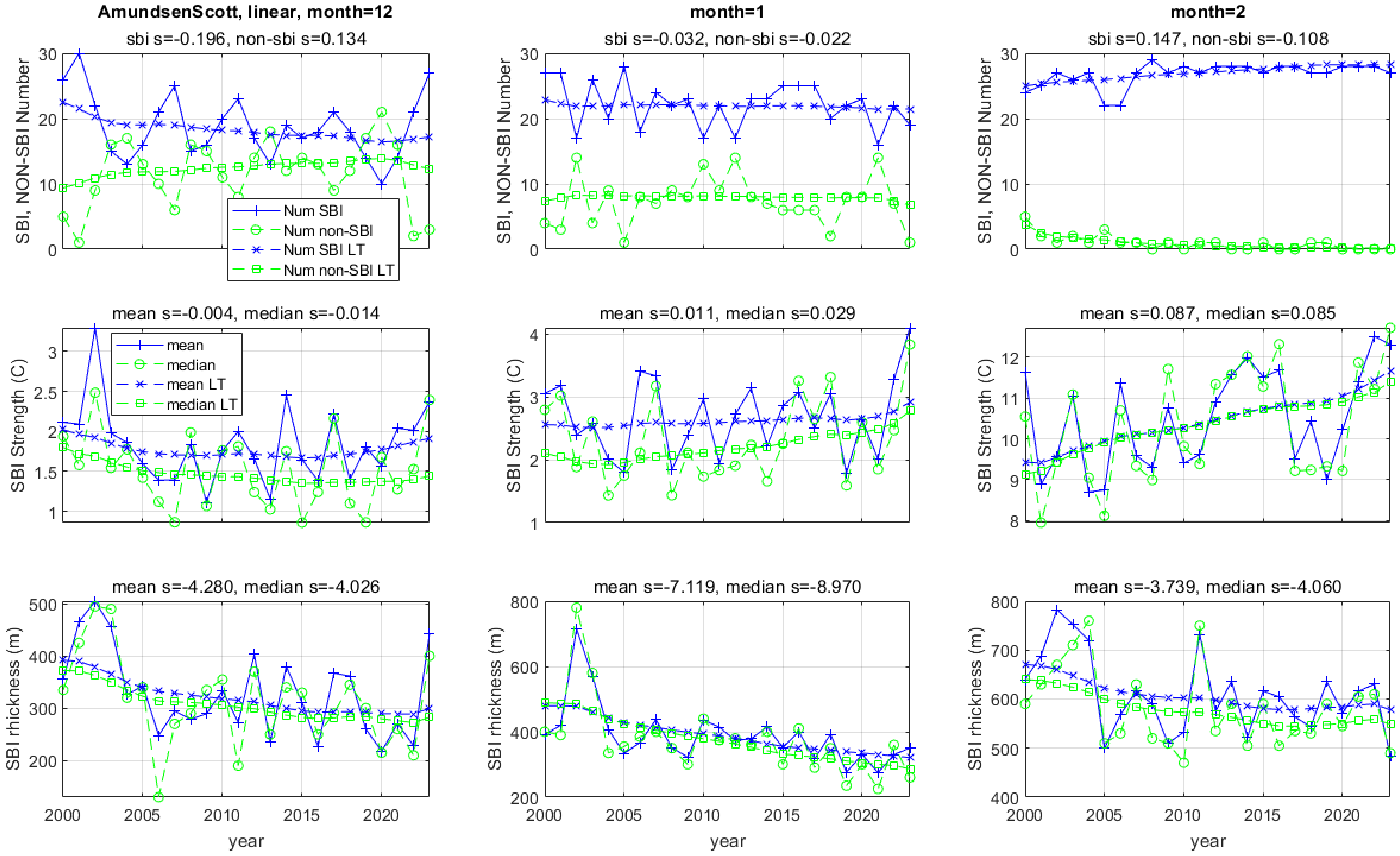
Amundsen-Scott Monthly Temperature Stats: Winter



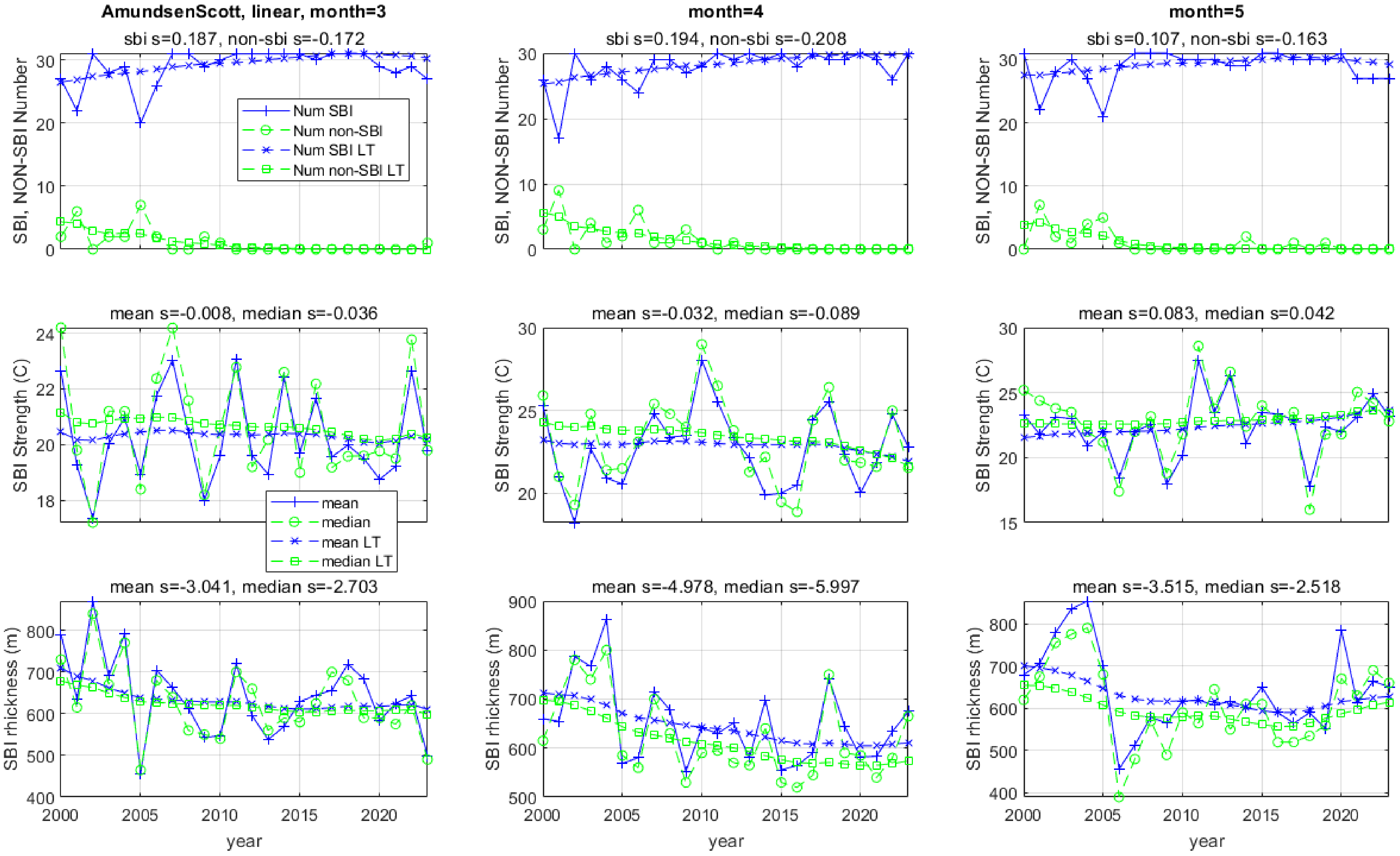
Amundsen-Scott Monthly Temperature Stats: Spring



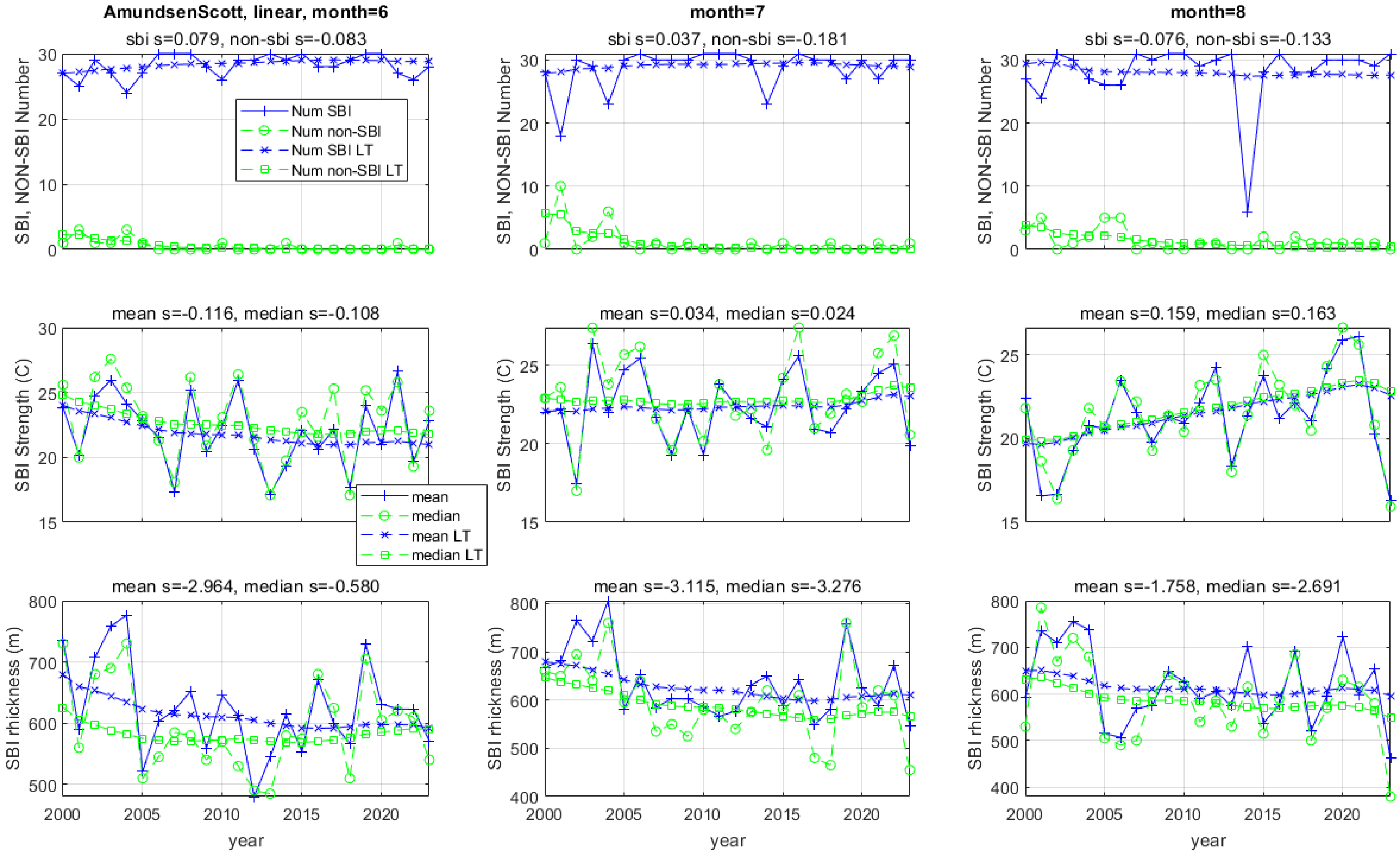
Amundsen-Scott Monthly SBI Stats: Summer



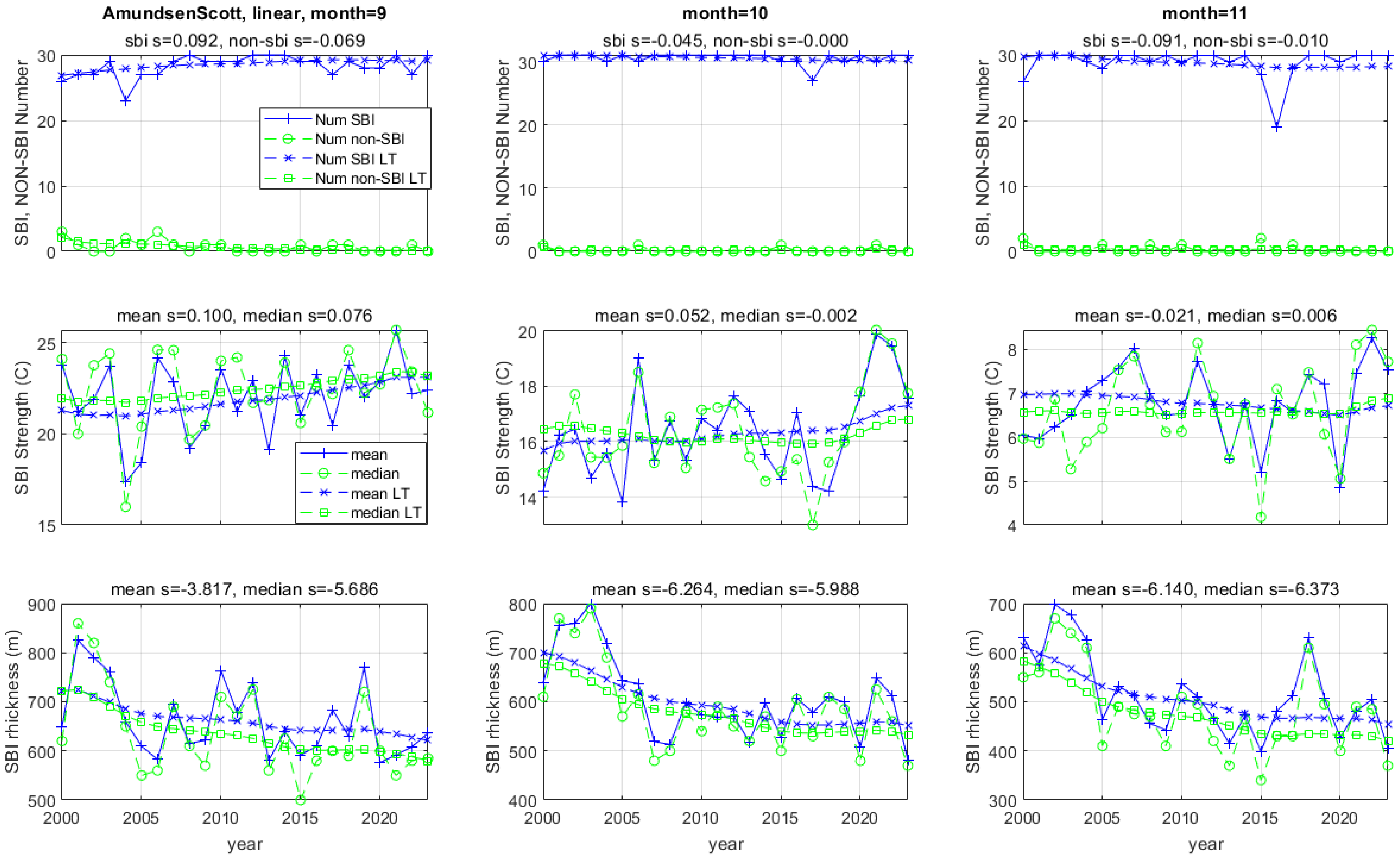
Amundsen-Scott Monthly SBI Stats: Fall



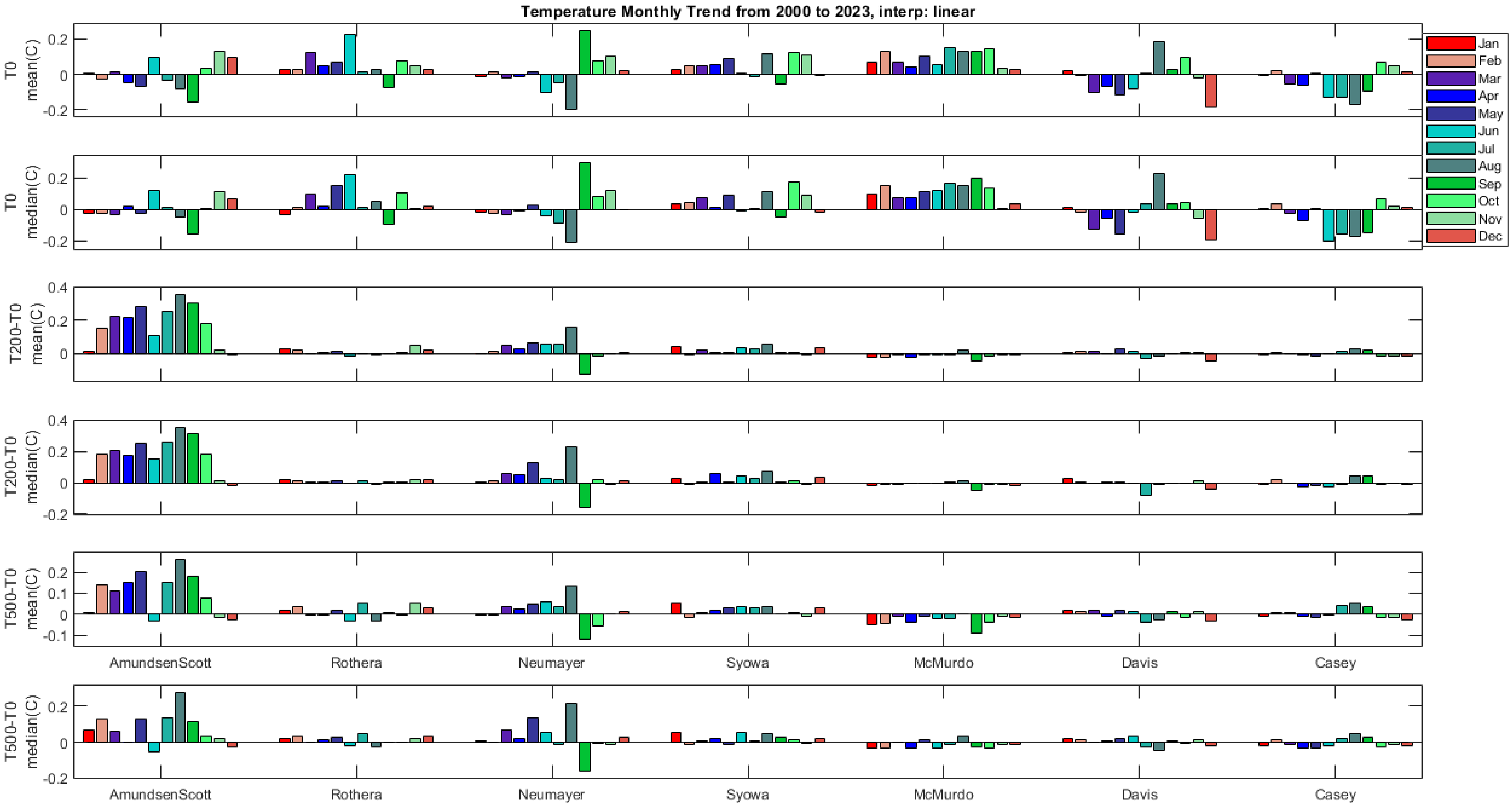
Amundsen-Scott Monthly SBI Stats: Winter



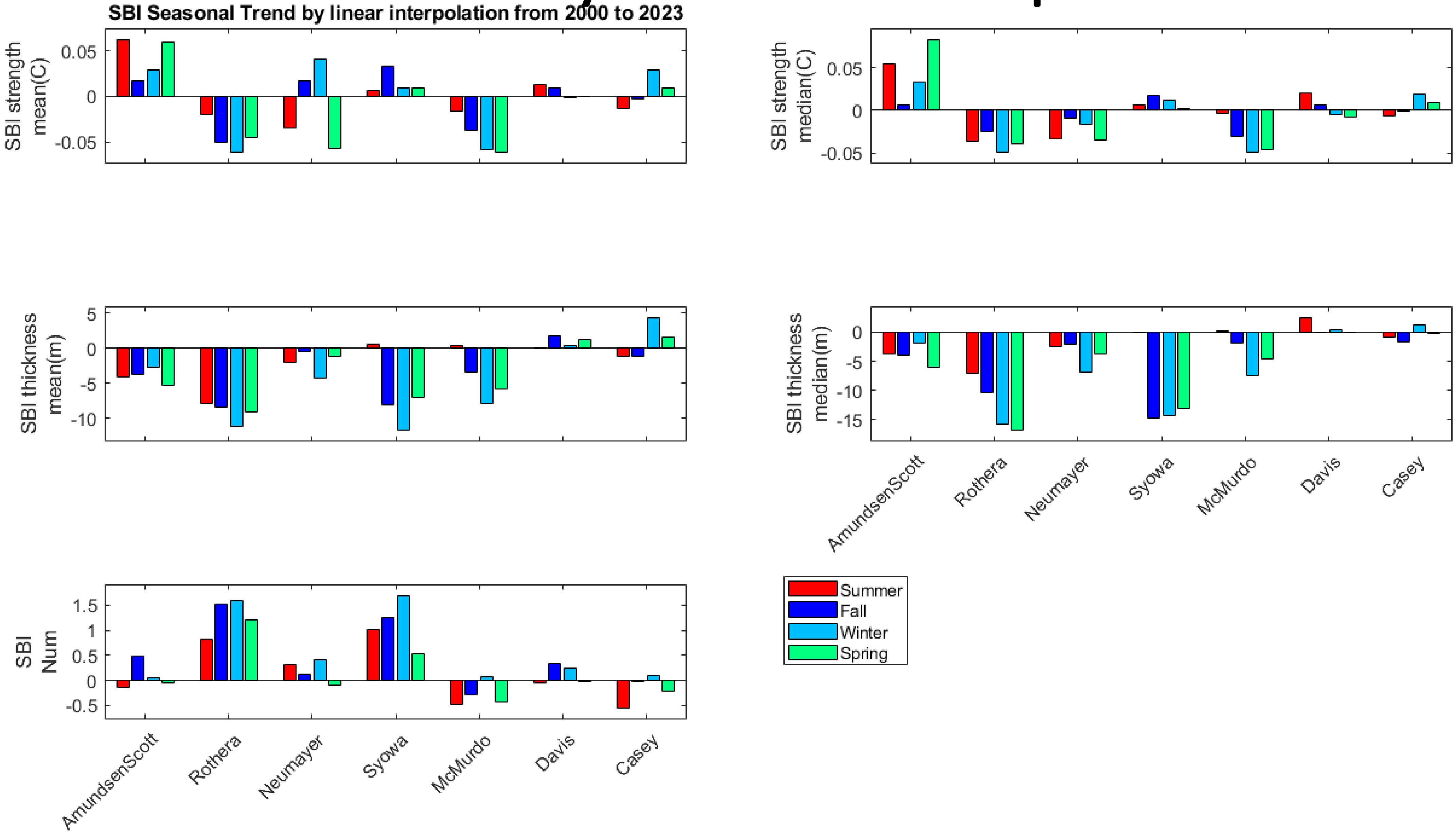
Amundsen-Scott Monthly SBI Stats: Spring



All Stations: Monthly Temperature Trend Slopes

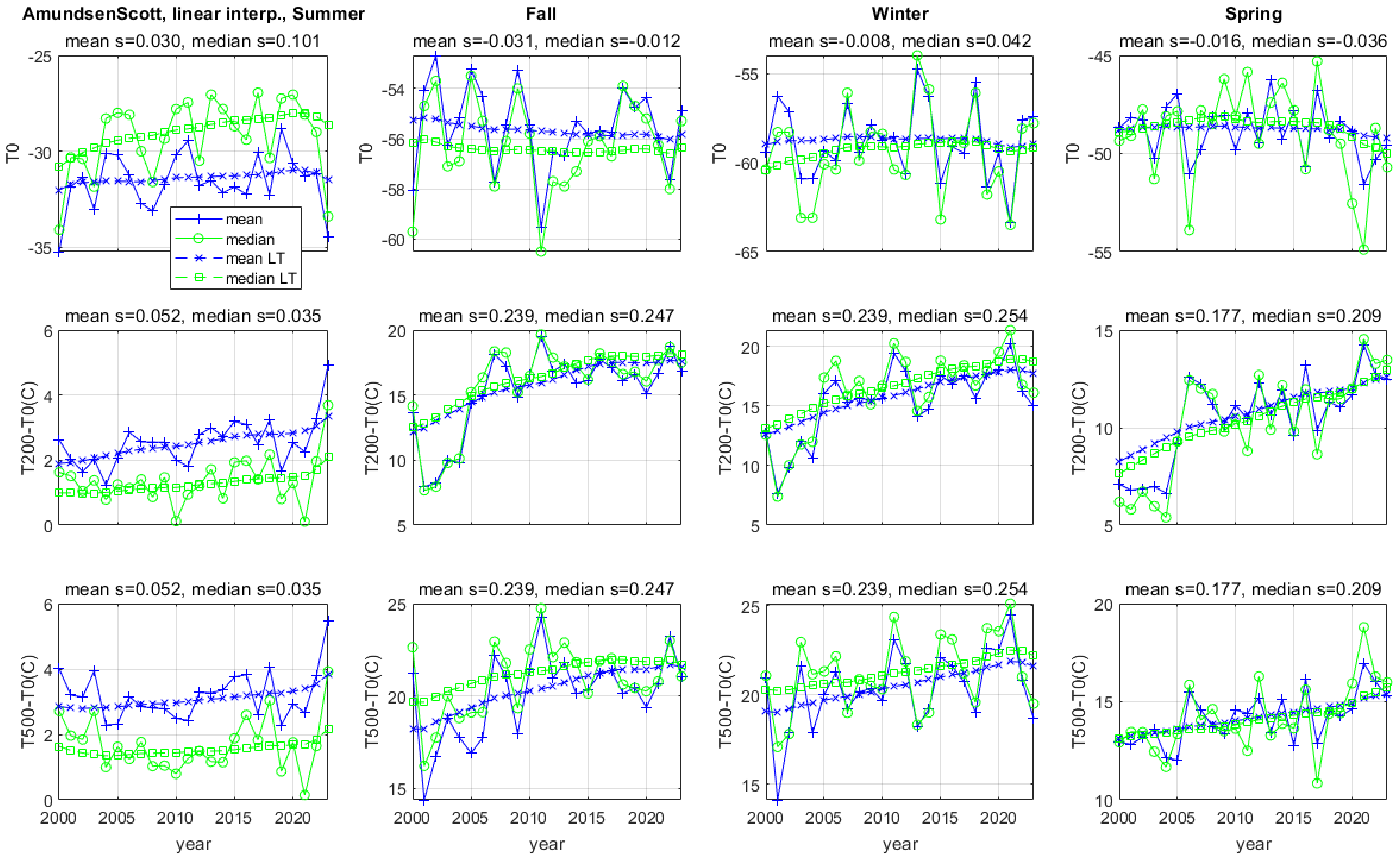


All Stations: Monthly SBI Trend Slopes

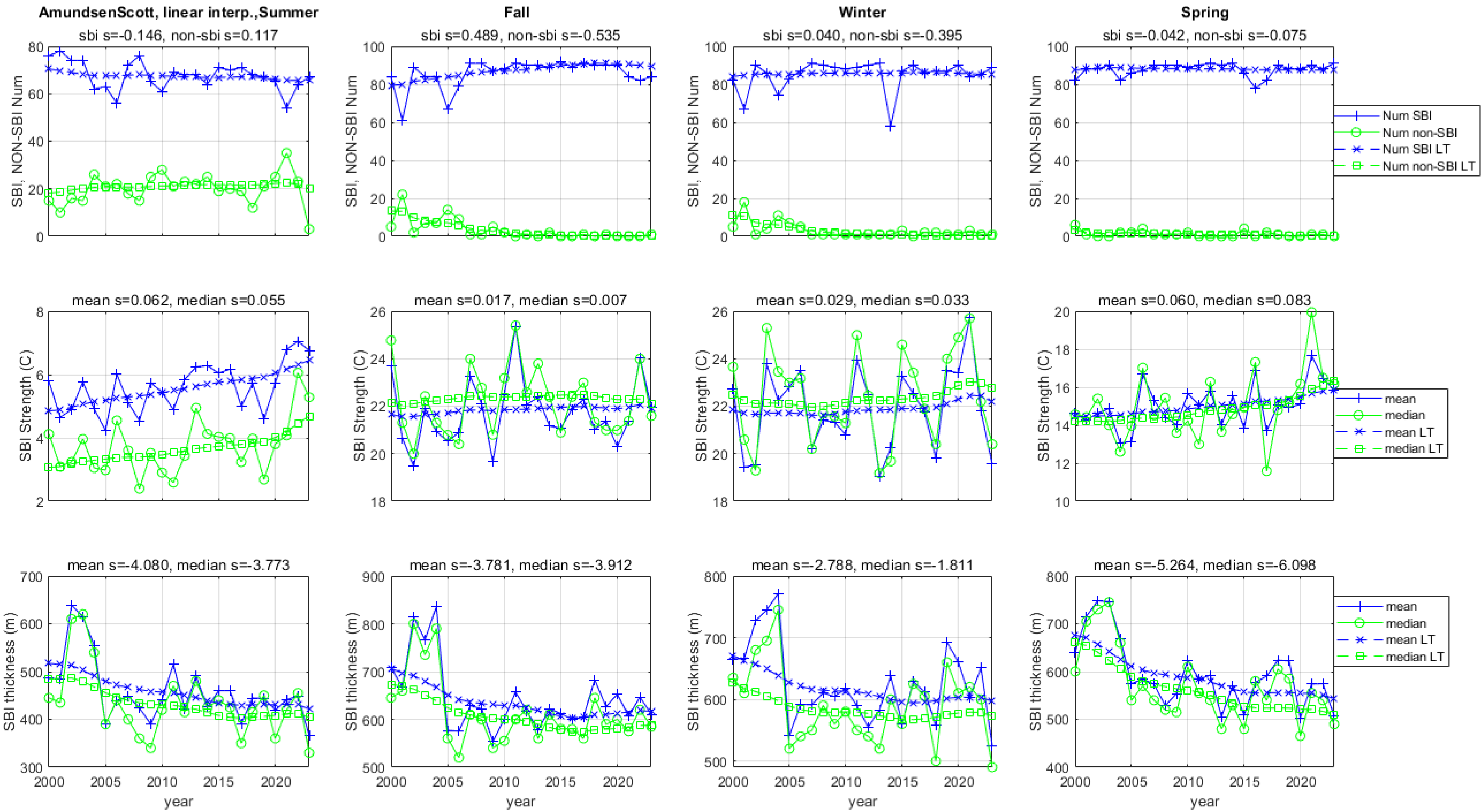


Seasonal Statistics and Trends

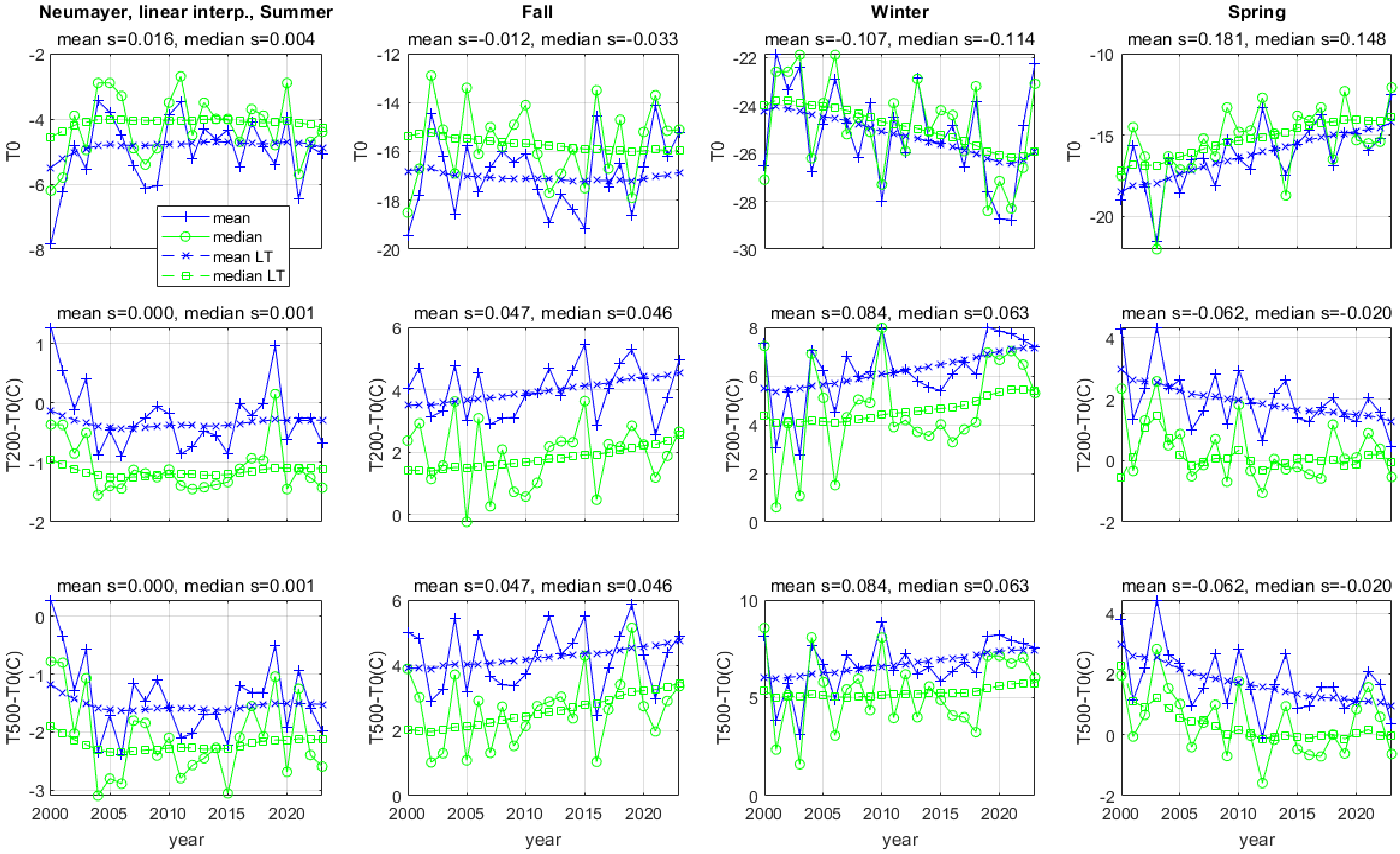
AmundsenScott: Seasonal Temperature Statistics



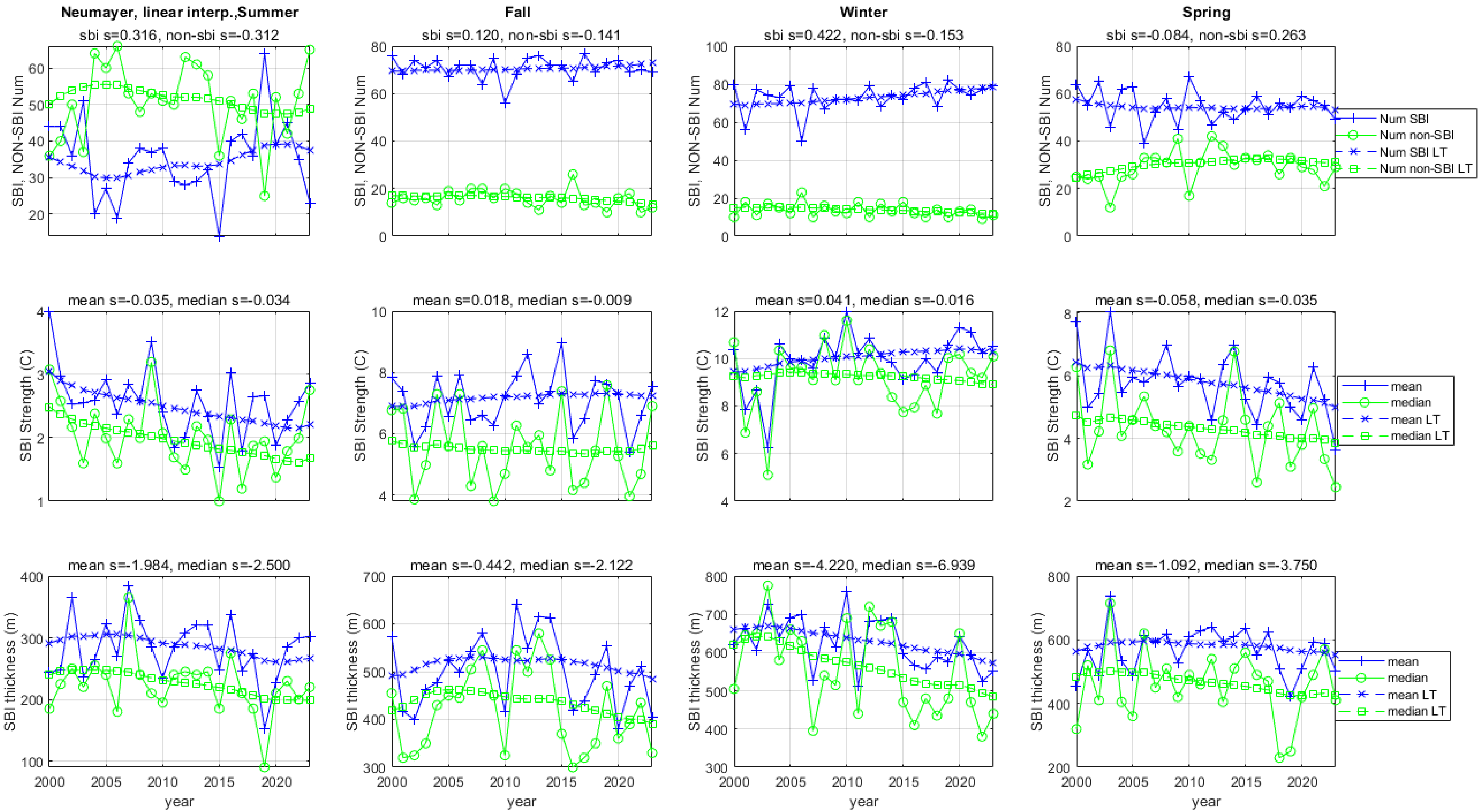
AmundsenScott Seasonal SBI Statistics



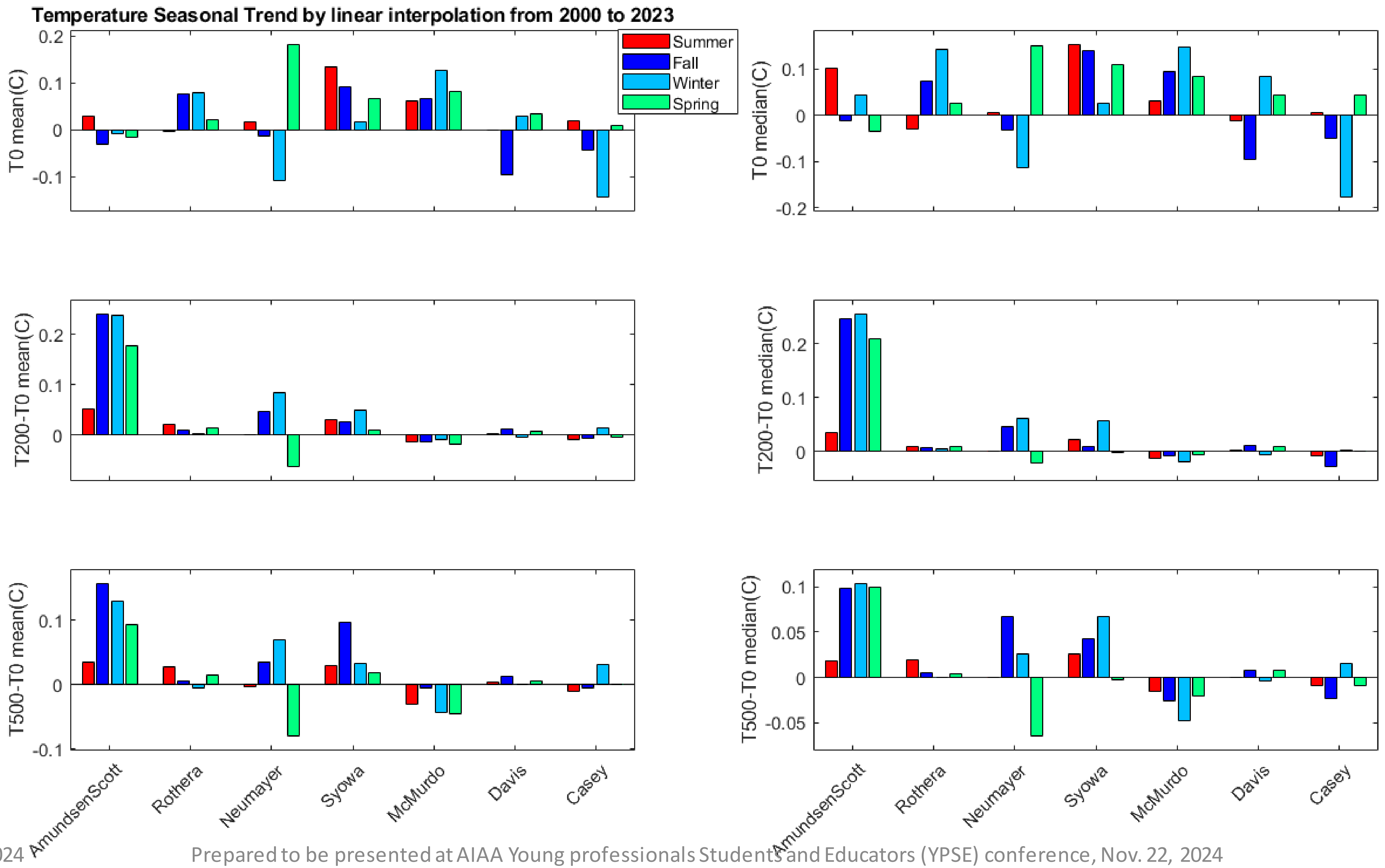
Neumayer: Seasonal Temperature Statistics



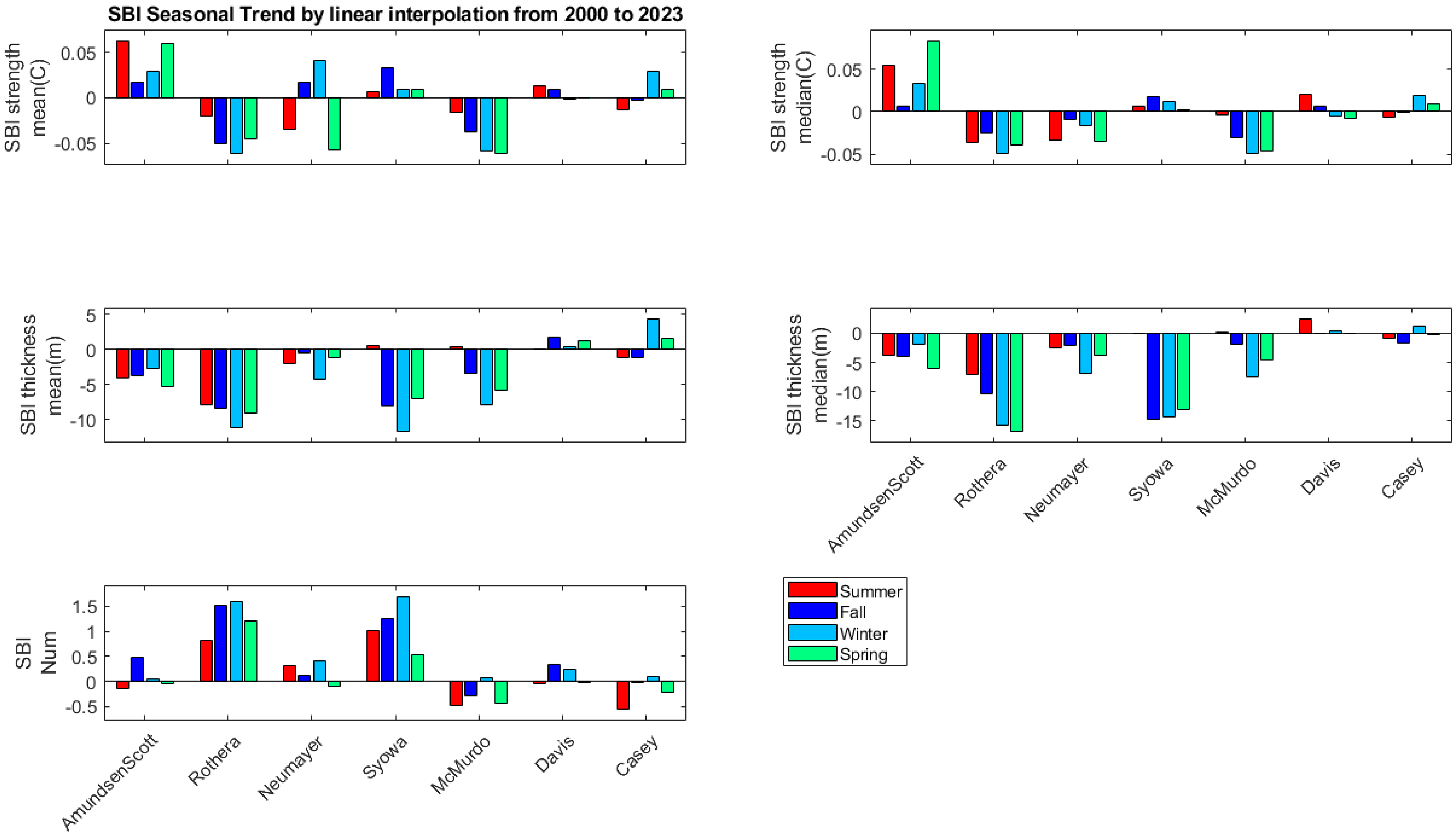
Neumaver: Seasonal SBI Statistics



All Stations: Seasonal Temperature Trend slopes



All Stations: Seasonal SBI Trend slopes



Summary

- **Pipeline Development & Algorithm Creation**
  - ✓ Built a system to parse and visualize IGRA data.
  - ✓ Organized data into human-readable monthly observation files.
  - ✓ Developed temperature interpolation and Surface-Based Inversion (SBI) algorithms.
  - ✓ Stored results into structured monthly profiles.
- **Data Analysis**
  - ✓ Calculated monthly and seasonal statistics of T0, T200–T0, T500–T0, SBI strength and thickness.
  - ✓ Applied trend decomposition to identify long-term trends (LT).
  - ✓ Calculated slopes for each LT variable.
- **Key Findings**
  - ✓ Changes of the Antarctic boundary layer structure since 2000 demonstrate regional and temporal characteristics.
  - ✓ Stations like *Syowa* and *Amundsen-Scott* show warming trends.
  - ✓ McMurdo and Casey stations exhibit mixed trends at different altitudes.
  - ✓ Stations *Rothera*, *Syowa*, and *Neumayer* show an increasing trend in SBI numbers, and SBI strength and their thickness display a decreasing trend.
  - ✓ SBI trends at other stations are not statistically significant.
- **Code Repository**
  - ✓ All code is publicly available on GitHub: *https://github.com/Justin123-Wu/AntarcticaTrend.git/*

Future Work

- **Develop a tool to help manually clean data**
  - ✓ Create a tool to visualize observations one at a time, manually deleting corrupt data.
  - ✓ Instead of automatically detecting SBIs, we can manually mark them.

- **Explore different interpolation methods**
  - ✓ Spline
  - ✓ Cubic
  - ✓ Snake

- **Compare the results from different SBI detection algorithms**
  - ✓ Searching from height = 0, find the inversion point.
  - ✓ Find maximum temperature point then check if it is greater than the ground temperature.
  - ✓ Clustering the measurements at ground and inversion points, then using their average to calculate inversion strength and thickness.