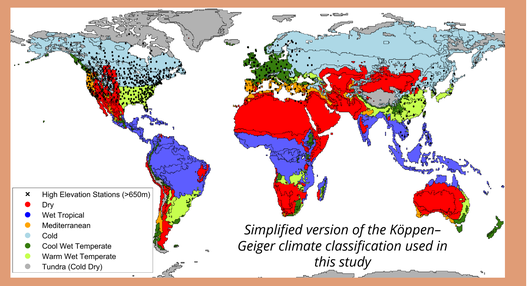
Weather Forecast Biases by Region and Climate Zone Assessed Using Hourly Observations

Cameron Gilbert[1], Sandra Yuter[1,2], Matthew Miller[1]
[1]Department of Marine, Earth, and Atmospheric Sciences, and the [2]Center for Geospatial Analytics
North Carolina State University, Raleigh, NC

Motivation

Information on where and when weather forecasts are more and less reliable is helpful to both forecast users and weather model developers.

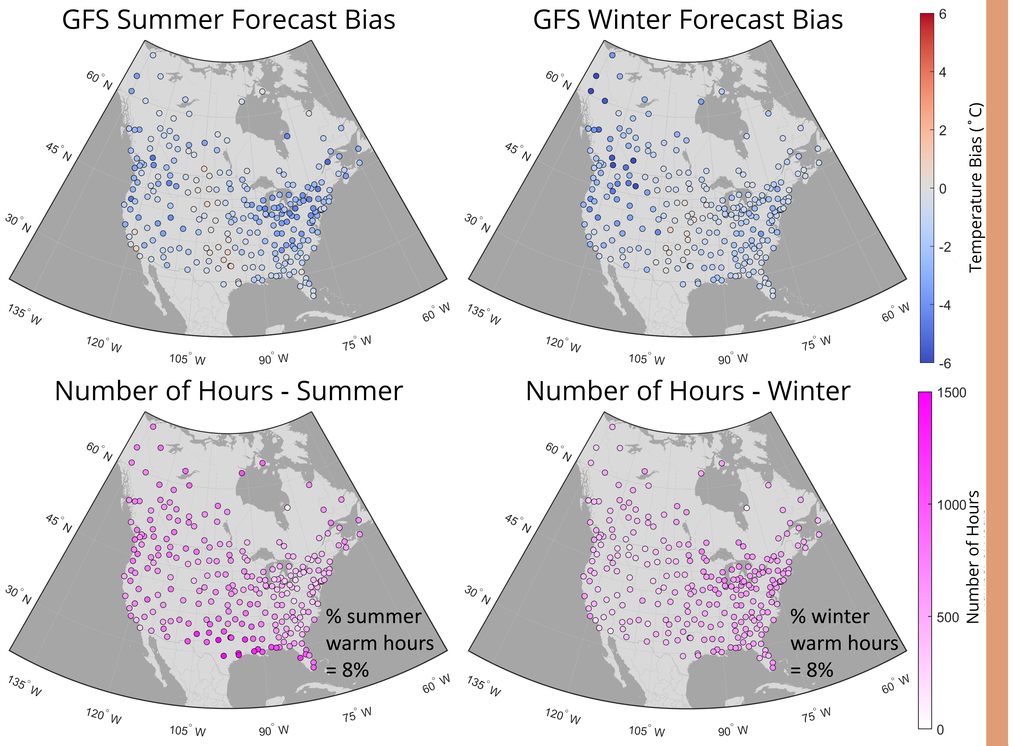


Simplified version of the Köppen–Geiger climate classification used in this study

Methods

We evaluate forecasts of heat waves and cold snaps when observed temperatures are >90th percentile and <10th percentile of the long-term climatology for that date, time, and location. We examine forecasts for two summer seasons (5/22 – 9/22 & 5/23 – 9/23) and two winter seasons (11/22 – 2/23 & 11/23 – 2/24). We also check model skill in predicting the start and end times of winter precipitation. We use ASOS (Automated Surface Observing Systems) surface observations at airport weather stations in North America to evaluate 48-hour lead time forecasts. Two models are evaluated, the U.S Navy's Regional Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) and NOAA's Global Forecasting System (GFS). Each weather station is categorized into its corresponding climate zone.

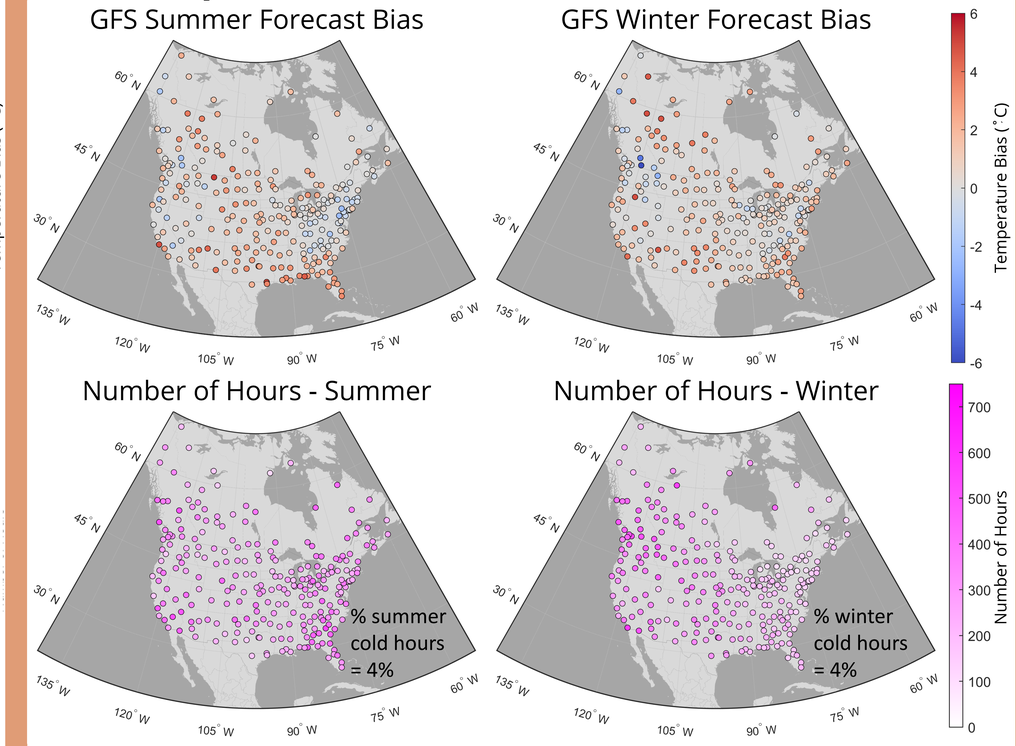
Heat Waves >90th Percentile Warm Events



Median temperature bias during events above the 90th percentile (top row) and number of hours above the 90th hourly climatological percentiles (bottom row)

Temperature (°C)	Summer '22		Summer '23		Winter '22 - '23		Winter '23 - '24	
Climate Zone	GFS	COAMPS	GFS	COAMPS	GFS	COAMPS	GFS	COAMPS
Cold	-1.90	-1.78	-2.41	-1.84	-1.40	-2.31	-1.13	-1.72
Sample Size	121	103	121	103	121	103	120	103
Dry	-0.83	N/A	-1.07	N/A	-1.59	N/A	-0.83	N/A
Sample Size	28	2	27	2	26	2	28	2
Cool Wet Temperate	-2.00	-1.82	-2.57	-1.57	-2.21	-1.74	-2.33	-1.41
Sample Size	10	9	10	9	10	9	10	9
Mediterranean	-2.05	-1.08	-2.69	-1.45	-2.92	-1.93	-2.31	-1.62
Sample Size	19	17	19	17	19	17	19	17
Warm Wet Temperate	-1.30	-1.26	-1.58	0.42	-1.45	-1.49	-1.22	-0.78
Sample Size	74	74	74	74	74	74	74	74

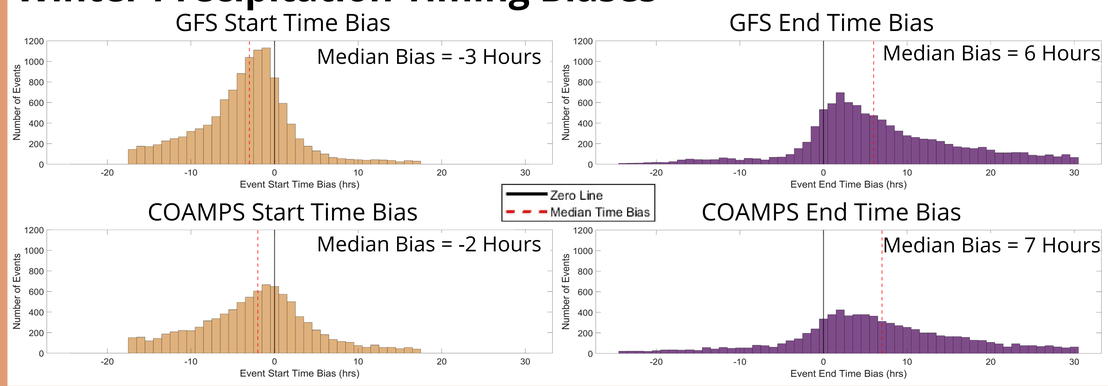
Cold Snaps <10th Percentile Cold Events



Median temperature bias during events below the 10th percentile (top row) and number of hours below the 10th hourly climatological percentiles (bottom row)

Temperature (°C)	Summer '22		Summer '23		Winter '22 - '23		Winter '23 - '24	
Climate Zone	GFS	COAMPS	GFS	COAMPS	GFS	COAMPS	GFS	COAMPS
Cold	0.09	2.16	1.64	1.80	0.60	1.35	1.50	1.68
Sample Size	121	103	121	103	121	103	120	103
Dry	0.95	N/A	2.44	N/A	1.59	N/A	2.49	N/A
Sample Size	28	2	27	2	26	2	28	2
Cool Wet Temperate	-0.39	0.52	1.74	1.24	0.20	0.00	0.91	0.53
Sample Size	10	9	10	9	10	9	10	9
Mediterranean	-0.58	0.02	1.42	1.57	0.50	-0.17	1.22	1.40
Sample Size	19	17	19	17	19	17	19	17
Warm Wet Temperate	0.80	3.15	2.29	3.46	1.04	1.90	1.34	1.83
Sample Size	74	74	74	74	74	74	74	74

Winter Precipitation Timing Biases



End Time Bias (H:mm)	Winter	
Climate Zone	GFS	COAMPS
Cold	7:00	9:38
Sample Size	88	70
Dry	6:00	N/A
Sample Size	50	8
Cool Wet Temperate	N/A	N/A
Sample Size	6	5
Mediterranean	9:15	5:00
Sample Size	35	31
Warm Wet Temperate	4:30	5:38
Sample Size	97	96

Summary (~48-h lead time forecast evaluation)

- Both weather prediction models are usually too **cold** for warm events and too **warm** for cold events across all climate zones.
- The actual number of cold hours is less than half of what is expected from the 30-year climatology.
- Weather models tend to overestimate the duration of winter-time precipitation. The **start** time is often 2+ hours too early and the **end** is often 6+ hours too late.

Acknowledgements

Thank you to Wayne Johnson, Rebecca Moore, and Nathan Faulk for their feedback. This work extends analysis by Rachel Kennedy (2023, M.S. thesis). Funded by ONR N00014-21-1-2116 and ONR N00014-24-1-2216