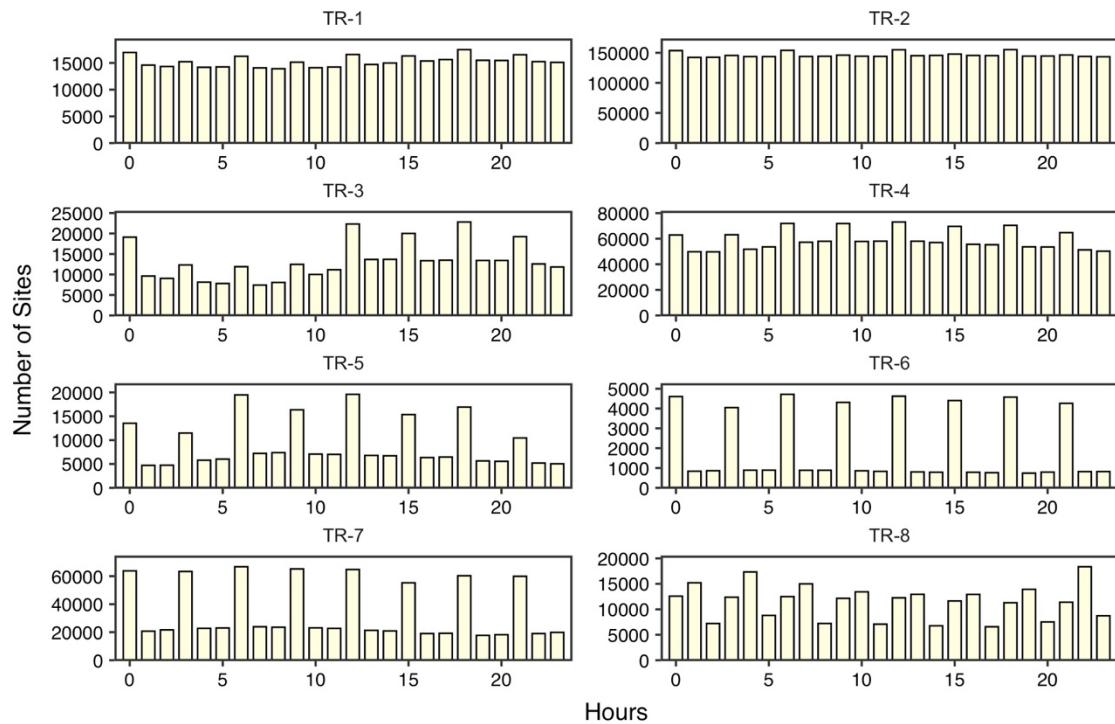
Supplement of

GHRSAT: the first global hourly dataset of all-sky remotely sensed estimates of surface air temperature

5 Zhenwei Zhang et al.

Correspondence to: Zhenwei Zhang (zhangw@nuist.edu.cn)

S1. Coverage of ground stations across different task regions

10   **Figure S1. The number of the ground sites with available SAT records in different task regions for the hours of the day. The number of sites on the hourly basis was computed using the records in ISD for January 2011. Distribution of the number of sites across the hours for each region is similar for other months in 2011–2023.**

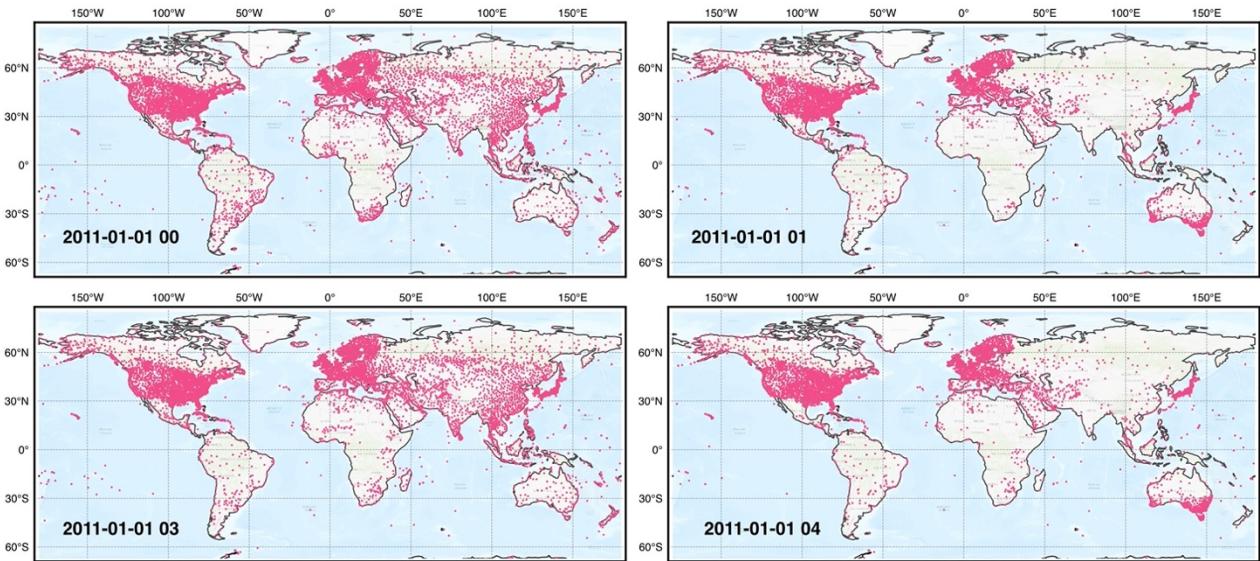



Figure S2. Spatial distribution of the ground sites with available data records for the hours of 0, 1, 3 and 4 on January 1, 2011. The land areas of US, the southern parts of Canada and the most of Europe are covered with high-density stations, and the stations are available with continuously hourly SAT records across the 24 hours of the days between 2011 and 2023. Many stations in other regions only contain records for several hours of the day. In particular, most of stations contained in ISD in China only have available records for 8 synoptic hours of 0, 3, 6, 9, 12, 15, 18, 21 in the days between 2011 and 2023.

S2. Model training and validation

Table S1. Setting of the core parameters for tuning the RF models in the first stage of hybrid modelling.

Parameter	Values
number of regression trees	60, 80, 100
number of features for splitting nodes	2, 3, 4, 5
fraction of observations to be sampled	0.8, 0.9, 1.0
type of sampling observations	with / without replacement