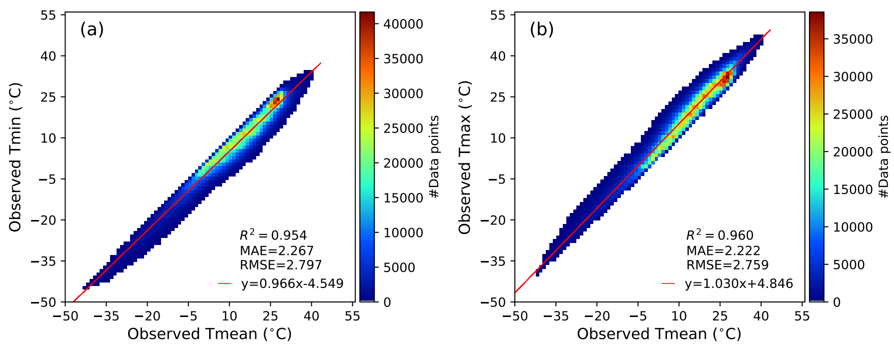

Fig. 1. Statistical relationships between Tmean, Tmin, and Tmax, which are collected from stations worldwide in 2012. (a) Statistical relationship between Tmean and Tmin. (b) Statistical relationship between Tmean and Tmax.

1. Relationships between Tmean, Tmin, and Tmax

We downloaded the Ta datasets from Global Surface Summary of the Day (GSOD, https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516), and found that the strong linear relationships could also be seen on a global scale. We show the relationships for the years 2012-2016 (Fig. 1 - Fig. 5). The reason why the strong linear relationships exist is unclear, but the linear relationships illustrate that it is possible to retrieve Tmean, Tmin, and Tmax simultaneously.

Assuming that we are retrieving Tmean y utilizing the satellite products x. That is, we find a model f

$$\hat{y} = f(x) \tag{1}$$

where y means the observed Tmean and $\hat{y}$ means the estimated Tmean. We use R^2, MAE, and RMSE to evaluate the model f, and then

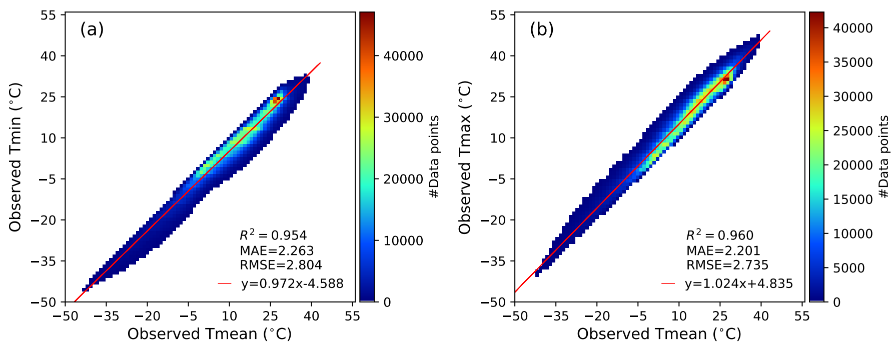

Fig. 2. Statistical relationships between Tmean, Tmin, and Tmax, which are collected from stations worldwide in 2013. (a) Statistical relationship between Tmean and Tmin. (b) Statistical relationship between Tmean and Tmax.

$$R_f^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{2}$$

$$\mathrm{MAE}_f = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{3}$$

$$\mathrm{RMSE}_f = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{4}$$

where $\bar{y}$ is the mean of y.

Taking Tmin (or Tmax) z as an example, since z and y have a strong linear relationship, we have a linear model h, and

$$\hat{z} = h(y) = ay + b \tag{5}$$

where a and b are the parameters in the linear model h, and $a > 0$. Then,

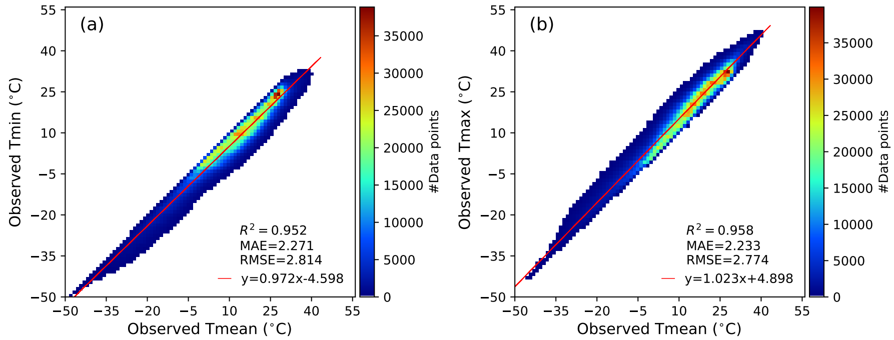

Fig. 3. Statistical relationships between Tmean, Tmin, and Tmax, which are collected from stations worldwide in 2014. (a) Statistical relationship between Tmean and Tmin. (b) Statistical relationship between Tmean and Tmax.

$$R_h^2 = 1 - \frac{\sum_{i=1}^{n}(z_i - \hat{z}_i)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} = 1 - \frac{\sum_{i=1}^{n}(z_i - ay_i - b)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} \tag{6}$$

$$\mathrm{MAE}_h = \frac{\sum_{i=1}^{n}|z_i - \hat{z}_i|}{n} = \frac{\sum_{i=1}^{n}|z_i - ay_i - b|}{n} \tag{7}$$

$$\mathrm{RMSE}_h = \sqrt{\frac{\sum_{i=1}^{n}(z_i - \hat{z}_i)^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(z_i - ay_i - b)^2}{n}} \tag{8}$$

To illustrate that it is possible to estimate Tmin $\hat{zp}$ according to the strong linear relationships between z and y, we can suppose that

$$\hat{zp} = p(x) = a\hat{y} + b \tag{9}$$

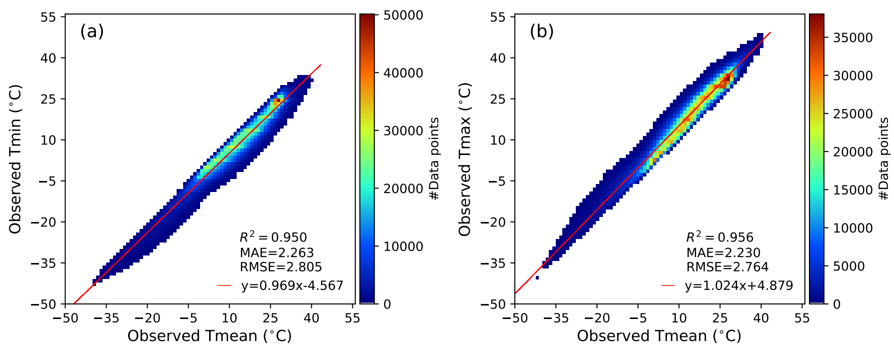

Fig. 4. Statistical relationships between Tmean, Tmin, and Tmax, which are collected from stations worldwide in 2015. (a) Statistical relationship between Tmean and Tmin. (b) Statistical relationship between Tmean and Tmax.

then,

$$\begin{aligned}
\mathrm{MAE}_p &= \frac{\sum_{i=1}^{n} |z_i - z\hat{p}_i|}{n} \\
&= \frac{\sum_{i=1}^{n} |z_i - a\hat{y}_i - b|}{n} \\
&= \frac{\sum_{i=1}^{n} |z_i - ay_i - b + ay_i - a\hat{y}_i|}{n} \\
&\leq \frac{\sum_{i=1}^{n} |z_i - ay_i - b| + \sum_{i=1}^{n} |ay_i - a\hat{y}_i|}{n} \\
&= \mathrm{MAE}_h + a\mathrm{MAE}_f
\end{aligned} \tag{10}$$

and

$$\begin{aligned}
R_p^2 &= 1 - \frac{\sum_{i=1}^{n}(z_i - z\hat{p}_i)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} \\
&= 1 - \frac{\sum_{i=1}^{n}(z_i - a\hat{y}_i - b)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} \\
&= 1 - \frac{\sum_{i=1}^{n}(z_i - ay_i - b + ay_i - a\hat{y}_i)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2}
\end{aligned} \tag{11}$$

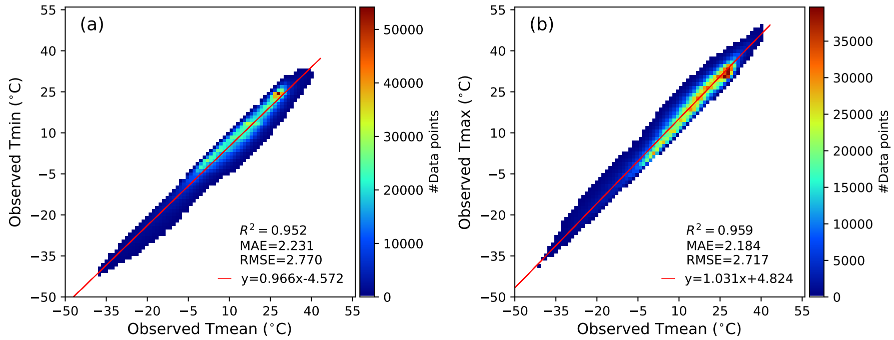

Fig. 5. Statistical relationships between Tmean, Tmin, and Tmax, which are collected from stations worldwide in 2016. (a) Statistical relationship between Tmean and Tmin. (b) Statistical relationship between Tmean and Tmax.

According to Minkowski inequality (Kreyszig, 1991)

$$
\begin{aligned}
&\sum_{i=1}^{n}(z_i - ay_i - b + ay_i - a\hat{y}_i)^2 \\
&\leq \left(\sqrt{\sum_{i=1}^{n}(z_i - ay_i - b)^2} + \sqrt{\sum_{i=1}^{n}(ay_i - a\hat{y}_i)^2}\right)^2 \\
&= \sum_{i=1}^{n}(z_i - ay_i - b)^2 + \sum_{i=1}^{n}(ay_i - a\hat{y}_i)^2 + 2\sqrt{\sum_{i=1}^{n}(z_i - ay_i - b)^2 \sum_{i=1}^{n}(ay_i - a\hat{y}_i)^2}
\end{aligned}
\tag{12}
$$

As a result,

$$
\begin{aligned}
R_p^2 &= 1 - \frac{\sum_{i=1}^{n}(z_i - ay_i - b + ay_i - a\hat{y}_i)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} \\
&\geq 1 - \frac{\sum_{i=1}^{n}(z_i - ay_i - b)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} - \frac{\sum_{i=1}^{n}(ay_i - a\hat{y}_i)^2}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} - \\
&\frac{2\sqrt{\sum_{i=1}^{n}(z_i - ay_i - b)^2 \sum_{i=1}^{n}(ay_i - a\hat{y}_i)^2}}{\sum_{i=1}^{n}(z_i - \bar{z}_i)^2} \\
&= 1 - (1 - R_h^2) - a^2(1 - R_f^2) - 2\sqrt{(1 - R_h^2)a^2(1 - R_f^2)} \\
&= R_h^2 + a^2 R_f^2 - a^2 - 2a\sqrt{(1 - R_h^2)(1 - R_f^2)}
\end{aligned}
\tag{13}
$$

Similarly,

$$
\begin{aligned}
\mathrm{RMSE}_p &= \sqrt{\frac{\sum_{i=1}^{n}(z_i - \hat{zp}_i)^2}{n}} \\
&= \sqrt{\frac{\sum_{i=1}^{n}(z_i - ay_i - b + ay_i - a\hat{y}_i)^2}{n}} \\
&\leq \sqrt{\frac{\sum_{i=1}^{n}(z_i - ay_i - b)^2}{n}} + a\sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \\
&= \mathrm{RMSE}_h + a\mathrm{RMSE}_f
\end{aligned}
\tag{14}
$$

Therefore, we can directly obtain Tmin $\hat{zp}$ utilizing estimated Tmean $\hat{y}$ and the parameters a and b that are determined by the linear relationships between z and y. The values of R_p^2, MAE_p, and RMSE_p can be used as the baselines, and other models should achieve better retrieval results than the baselines.

References

Kreyszig, E., 1991. Introductory functional analysis with applications. volume 17. John Wiley & Sons.