International Journal of
Geo-Information

MDPI

Supplementary Materials

Information of the additional regression models

1. Ordinary Least Squares model (OLS)

Ordinary Least Squares model (OLS) assumes that the regression coefficients have nothing to do with the spatial and temporal positions of the sample data. The independent variable coefficients calculated by OLS are not only the best unbiased estimate of the current point, but also the best unbiased estimate of the average level of the study area. Weighted Least Squares (WLS) is usually used to solve this method.

OLS model is the most commonly used and most basic statistical method to determine the regression relationship of variables. For the actual problem solving, the sample data has n sets of observation data $(GDD_i, KDD_i, VPD_i, PCPN_i; dY_i), i = 1,2,\cdots,n$, the OLS model can be expressed as (Equation S1):

$$dY_i = \beta_0 + \beta_1 GDD_i + \beta_2 KDD_i + \beta_3 VPD_i + \beta_4 PCPN_i + \varepsilon_i,\ \ i = 1,2,\cdots,n \quad \text{(S1)}$$

where dY_i is the detrend yield, GDD_i, KDD_i, VPD_i and $PCPN_i$ is the meteorology variables with sample i, ε_i is a random error.

2. Time series model (Time-LR model)

Time series model is based purely on time series data from a single point or area. By combining our research, the 40 years of simulated data can be represented to fit a time-series model as follows (Equation S2):

$$\mathrm{dY_{i,t}} = \beta_{\mathrm{i},0} + \beta_{\mathrm{i},1}\mathrm{GDD_{i,t}} + \beta_{\mathrm{i},2}\mathrm{KDD_{i,t}} + \beta_{\mathrm{i},3}\mathrm{VPD_{i,t}} + \beta_{\mathrm{i},4}\mathrm{PCPN_{i,t}} + \varepsilon_{\mathrm{i,t}} \quad \text{(S2)}$$

where i is the index for the county, t is the index for year. $dY_{i,t}$, $GDD_{i,t}$, $KDD_{i,t}$ $VPD_{i,t}$ and $PCPN_{i,t}$ are de-yield, GDD, KDD, VPD and PCPN at county i at year t, respectively. In year t, $\beta_{i,0-4}$ represent model parameters to be fit, and $\varepsilon_{i,t}$ is an error term.

3. Geographically weighted regression model (GWR model)

The GWR model is an extension of the OLR model, which embeds the spatial relationship changes caused by spatial position differentiation into the calculation of regression coefficients, expressed as (Equation S3):

$$\begin{gathered} dY_i = \beta_{(u_i,v_i),0} + \beta_{(u_i,v_i),1} GDD_i + \beta_{(u_i,v_i),2} KDD_i + \beta_{(u_i,v_i),3} VPD_{i,t} + \\ \beta_{(u_i,v_i),4} PCPN_i + \varepsilon_i \end{gathered} \quad \text{(S3)}$$

where, (u_i, v_i) represents the space coordinates of the sample county i, $\beta_{(u_i,v_i),0}$ is the regression constant, $\beta_{(u_i,v_i),1-4}$ is the regression coefficient of the independent variables separately, related to spatial position (u_i, v_i). ε_i is the error term of sample county i, subject to $\varepsilon_i \sim N(0,\sigma^2)$, and $Cov(\varepsilon_i, \varepsilon_j) = 0 (i \neq j)$.

OLS and Time-LR were implemented in R language, while GWR and GTWR model were implemented in Python.

Table S1. Summary of parameters of ordinary least square (OLS), time linear regression (LR), geographically weighted regression (GWR) with Gaussian kernel, GWR with the bi-square kernel, geographically and temporally weighted regression (GTWR) with Gaussian kernel, and GTWR with bi-square kernel models.

Parameters	OLS	Time-LR	GWR-Gaussian	GWR-Bisquare	GTWR-Gaussian	GTWR-Bisquare
Band	/	/	100	100	100	100
Space: Time	/	/	/	/	0.5:0.5	0.5:0.5