

S1. Equation of CNN

$$y_j^l = \sum_{i=1}^m x_i^l \otimes k_{ij}^l + b_j^l \quad (S1)$$

Where y denotes the output feature vector; m denotes the length of the input feature vector; x denotes the input feature vector; k denotes the convolution kernel; b denotes the bias vector.

S2. Equations of BiLSTM

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (S2)$$

Where f_t denotes the forgetting gate, whose value ranges from 0 to 1; σ denotes the activation function; W_f denotes the weight of the forgetting gate; b_f denotes the bias of the forgetting gate; x_t denotes the input value of the t moment; h_{t-1} denotes the output value of the $t-1$ moment.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (S3)$$

Where i_t denotes the input gate, whose value ranges from 0 to 1; W_i denotes the weight of the input gate; b_i denotes the bias of the input gate.

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (S4)$$

Where $\tilde{C}_t$ denotes the t momentary temporary unit state; $\tanh$ denotes the hyperbolic tangent activation function; W_c denotes the weight of the temporary unit, b_c denotes the bias of the temporary unit.

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \quad (S4)$$

Where C_t denotes the cell state at the t moment, which takes values ranging from 0 to 1; C_{t-1} denotes the cell state at the $t-1$ moment.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (S6)$$

Where o_t denotes the output gate, whose value ranges from 0 to 1; W_o denotes the weight of the output gate; b_o denotes the bias of the output gate.

$$h_t = o_t \otimes \tanh(C_t) \quad (S7)$$

Where h_t denotes the output value of the moment, i.e. the model's prediction of the $t+1$ moment.

Equations S2–S7 present the computation process of the unidirectional LSTM. We denote the two opposing unidirectional LSTMs as $\vec{h}_t$ and $\overleftarrow{h}_t$, and the BiLSTM's final output $h_t|_{Bi}$ is given in equation S8.

$$h_t|_{Bi} = \begin{bmatrix} \vec{h}_t \\ \overleftarrow{h}_t \end{bmatrix} \quad (S8)$$

S3. Equation of AM

$$S = \sigma \left(\frac{\vec{Q}\vec{K}^T}{\sqrt{d_k}} \right) \vec{V} \quad (S9)$$

Where S denotes the final result of the attention mechanism, i.e., the weighted sum of all the hidden states of the encoder and their corresponding attention weights; σ denotes the activation function; $\vec{Q}$ denotes the output vector; $\vec{K}$ denotes the input vector; $\vec{V}$ denotes the vector of attention values; d_k denotes the dimension of the $\vec{K}$ vector.

S4. Equations of model effectiveness evaluation

$$MAE = \frac{1}{n} \sum_{t=1}^n |S_i|_t - H_i|_t| \quad (S10)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n (S_i|_t - H_i|_t)^2 \quad (S11)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|S_i|_t - H_i|_t|}{H_i|_t} \times 100\% \quad (S12)$$

$$R^2 = 1 - \frac{\sum_{t=1}^n (S_i|_t - H_i|_t)^2}{\sum_{t=1}^n (S_i|_t - \overline{H_i|_t})^2} \quad (S13)$$

$$E_i|_t = \left| \frac{S_i|_t - H_i|_t}{H_i|_t} \right| \quad (S14)$$

Where $E_i|_t$ denotes the relative error of variable i in year t ; $S_i|_t$ denotes the model predicted value of variable i in year t ; $H_i|_t$ denotes the true value of variable i in year t ; and $\overline{H_i|_t}$ denotes the average of the true values of variable i .

S5. Equation of Monte Carlo simulation

$$C_D|_T = \tilde{F}[x_{S1}|_T \cdot (1 + \omega_{x1} \cdot \frac{T - 2022}{2050 - 2022}), \dots, x_{Si}|_T \cdot (1 + \omega_{xi} \cdot \frac{T - 2022}{2050 - 2022})], \omega_{xi} \sim N(0, \sigma) \quad (S15)$$

Where, $C_D|_T$ denotes CECBS in year T under the dynamic scenario; $\tilde{F}$ denotes the mapping relationship between the independent and dependent variables in the CNN-BiLSTM-AM model; $x_{Si}|_T$ denotes the value of variable i in year T under the static scenario; ω_{xi} denotes the random parameter of variable i ; N denotes that the random parameter follows a normal distribution; σ denotes the standard deviation.

S6. Equations of FAST

$$\omega_{i,t}^{(j)} = \mu_{\omega,i,t} + \frac{\sigma_{\omega,i,t}}{2} [\cos(s_{i,t}^{(j)}) - 1], j = 1, 2, \dots, n \quad (S16)$$

Where $\omega_{i,t}^{(j)}$ denotes the disturbance coefficient of the i -th factor at time step t ; $\mu_{\omega,i,t}$ denotes the mean value of the i -th factor at time step t ; $\sigma_{\omega,i,t}$ denotes the standard deviation of the i -th factor at time step t ; $s_{i,t}^{(j)}$ represents the random parameter obtained through Fourier sampling; and n denotes the total number of samples.

$$y(s) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(ns) + b_n \sin(ns)] \quad (S17)$$

Where $y(s)$ denotes function of the input random parameter s ; a_0 denotes the DC component (constant term), which represents the baseline output without any disturbances; a_n and b_n are the coefficients of the Fourier series expansion, which denote the amplitudes corresponding to the sine

and cosine terms, reflecting the contributions of different frequencies.

$$\sigma^2(y) = a_0^2 + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \quad (\text{S18})$$

Where $\sigma^2(y)$ denotes the total variance of the output.

$$S_i = \frac{\sigma_i^2}{\sigma^2(y)} \quad (\text{S19})$$

Where S_i denotes sensitivity coefficient for the i -th factor in the model; σ_i^2 denotes the variance contribution of the Fourier component corresponding to the disturbance coefficient ω_i of the i -th factor.