Text S1: machine learning models

MLP (Multilayer Perceptron) neural networks are common ANN (Artificial Neuro Network) algorithms that consist of an input layer, an output layer and one or more hidden layers(Heidari et al., 2019; Jiang, 2008; Premalatha and Valan Arasu, 2016). All neurons in an MLP are similar, each with several input (connected to the previous layer) and output (connected to the next layer) neurons that pass the same value to multiple output neurons connected to it. Support vector machines (SVM) are based on Vapnik-Chervonenkis and structural risk minimization theory and use their kernel functions to deal with small sample problems, non-linear situations and high-dimensional pattern recognition problems, where they show unique advantages(Cortes and Vapnik, 1995; Moazenzadeh et al., 2022). Multiple Linear Regression (MLR) is a regression analysis model involving two or more independent variables. MLR models allow for an intuitive and rapid analysis of the correlation between multiple variables and the dependent variable(Zelterman, 2015). As a result, MLR models are widely used in the social sciences, economics and technology.

Text S2: 10-fold cross-validation

The step of 10-fold cross-validation is to first divide the dataset into 10 equal parts, of which 9 are used as training datasets to build the model and the remaining 1 is used as the validation dataset for testing. This process is repeated 10 times until each sample of each part is tested once, and then the average of the 10 test results (mainly contain R^2*, RMSE, and MAE*) is taken as the final cross-validation result.

Figure S1: Correlations among the DGSR and the meteorological variables

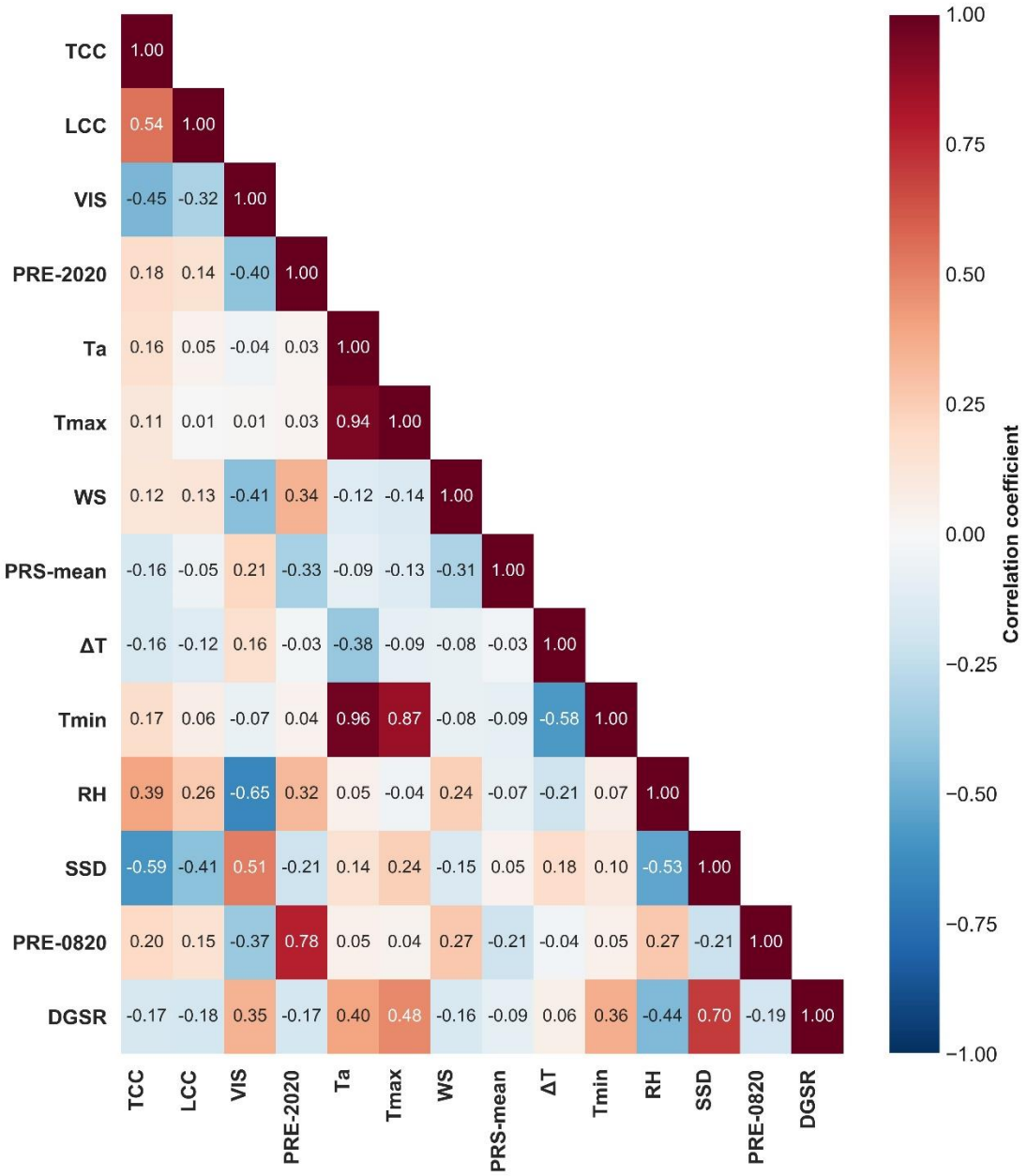

Figure S1. Correlations among the DGSR and the meteorological variables. Colors indicate the Spearman's rank correlation coefficient values.

References:

Cortes, C., Vapnik, V., 1995. Support-Vector Networks. Mach. Learn. 20, 273–297. https://doi.org/10.1023/A:1022627411411

Heidari, A.A., Faris, H., Aljarah, I., Mirjalili, S., 2019. An efficient hybrid multilayer perceptron neural network with grasshopper optimization. Soft Comput. https://doi.org/10.1007/s00500-018-3424-2

Jiang, Y., 2008. Prediction of monthly mean daily diffuse solar radiation using artificial neural networks and comparison with other empirical models. Energy Policy. https://doi.org/10.1016/j.enpol.2008.06.030

Moazenzadeh, R., Mohammadi, B., Duan, Z., Delghandi, M., 2022. Improving

generalisation capability of artificial intelligence-based solar radiation estimator models using a bio-inspired optimisation algorithm and multi-model approach. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-021-17852-1

Premalatha, N., Valan Arasu, A., 2016. Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. J. Appl. Res. Technol. https://doi.org/10.1016/j.jart.2016.05.001

Zelterman, D., 2015. Applied Multivariate Statistics with R, Applied Multivariate Statistics with R. https://doi.org/10.1007/978-3-319-14093-3