A novel multi-scale parameter estimation approach to the Hargreaves-Samani equation for estimation of Penman-Monteith reference evapotranspiration

Ho-Jun Kim [a], Sewwandhi Chandrasekara [b], Hyun-Han Kwon [a,*], Carlos Lima [c], Tae-woong Kim [d]

[a] *Department of Civil and Environmental Engineering, Sejong University, South Korea*
[b] *Department of Agricultural Engineering, Faculty of Agriculture, University of Peradeniya, Sri Lanka*
[c] *Department of Civil and Environmental Engineering, University of Brasilia, Brazil*
[d] *Department of Civil and Environmental Engineering, Hanyang University, South Korea*

ABSTRACT

The main focus of this study is to develop a multi-scale surrogate model for the FAO-56 Penman-Monteith (PM) evapotranspiration (ET_o) using Hargreaves-Samani (HS) equation, which uses only temperature as a hydrometeorological variable to estimate ET. This feature is particularly useful for scarce data regions and climate change impact assessment studies, where the direct estimation of ET_o from the PM equation can be problematic. As the parameters of the HS equation may vary across space, a Bayesian approach was adopted to estimate (or recalibrate) them rather than relying on the fixed values as suggested in the traditional model. The Bayesian approach allows a sound development of our model in a multi-scale temporal framework, where the ET_o at daily, monthly and annual scales are jointly used to estimate the HS equation parameters. The proposed and reference models are applied and tested using meteorological data from 17 stations located across the Han river basin in South Korea. The results indicate that the traditional HS equation with fixed parameters and without recalibration tends to overestimate the reference ET for all stations. The locally recalibrated approach to the HS equation at a daily temporal scale can effectively reduce the systematic bias associated with the use of the traditional HS equation but fails to reproduce the underlying distribution of ET_o at different temporal scales (e.g., monthly and annual). This leads to a large systematic bias in ET_o at these scales. In contrast, the proposed multi-scale surrogate model offers a more precise estimation of the reference ET at a daily timescale as well as at the aggregated monthly and annual temporal scales. This is particularly useful to minimize the systematic bias often observed when using traditional surrogate models for the reference ET in hydrological studies such as rainfall-runoff modeling and assessment of climate change impact on water resources.

1. Introduction

The hydrological cycle involves the continuous process of water fluxes across their reservoirs: land, oceans and atmosphere. Among the hydrologic flux components, evapotranspiration (ET) responds to a substantial transport of water over in the hydrological cycle (Kim et al., 2018a; Kwon et al., 2012; McColl and Rigden, 2020; Novák, 2012; Zhao et al., 2013). In terms of the hydrological cycle, ET represents the key process of removing water from the surface to the atmosphere, which is the opposite process of precipitation. The land surface returns approximately 60% of precipitation to the atmosphere annually via the ET process (L'vovich et al., 1990). Therefore, accurate estimation of ET is

often required to establish effective management plans for water resources (e.g., demand analysis of water, irrigation management, drought monitoring, etc.) (Zhao et al., 2013).

The main concepts of ET can be divided into potential evapotranspiration (ET_p) and reference evapotranspiration (ET_o). The ET_p refers to the maximum moisture loss from the surface, which is strongly affected by meteorological conditions and the surface type, assuming conditions of unlimited moisture supply (Lhomme, 1997). On the other hand, the reference evapotranspiration (ET_o) can be defined as the hypothetical evapotranspiration rate of an extensive green grass surface (Allen et al., 1998).

ET_o is routinely estimated through the physically-based FAO56

Penman-Monteith (FAO56-PM) equation, which is recommended by the Food and Agriculture Organization (FAO) and has been used by various organizations (e.g., WMO and ICID). However, the FAO56-PM method requires a set of meteorological data (i.e., wind speed, radiation, air temperature, and relative humidity) to estimate ET_o that are not readily available and therefore limit the use of the FAO56-PM in data-sparse parts of the world, especially in developing countries. Such a problem arises in the study of the impact of climate change on water resources due to difficulties in obtaining reliable meteorological forcing data from climate change scenarios required for estimating ET_o through the FAO56-PM Equation.

In this context, the Hargreaves-Samani (HS) equation (Hargreaves and Samani, 1982, 1985) using daily air temperature data was proposed as a surrogate model to the FAO56-PM for the ET_o. The HS equation has been widely applied for the estimation of ET_o due to its accuracy, consistency, and parsimonious nature. The HS formula has a physical interpretation, and the temperature range (T_{max} - T_{min}) between the maximum and minimum temperature is aiming to represent atmospheric transparency. The air temperature used in the HS equation is one of the most widely observed meteorological data, permitting easy implementation with long-term temperature data (Mendicino and Senatore, 2013). However, the HS equation relies on model coefficients, which in turn can be linked to the local climate conditions and topographical characteristics. Therefore, recalibration of the HS equation is required to minimize systematic error.

Several studies have been proposed to improve the accuracy of the estimated ET_o by recalibrating the parameters of the HS equation. In particular, the pattern of ET_o appears to be different due to various meteorological factors affecting evapotranspiration. Thus, the HS equation has been evaluated under various climate regimes (e.g., arid and semiarid) and regions (e.g., inland and coast). Most studies focused on the assessment of ET_o estimated from the HS equation in a particular climate zone (Gavilán et al., 2006; Martínez-Cob and Tejero-Juste, 2004; Razei and Pereira, 2013; Todorovic et al., 2013; Yang et al., 2021). Almorox et al. (2015) evaluated 11 representative temperature-based models for estimating ET_o in different climate zones across the world. More specifically, the applicability of the different temperature-based models was explored in the context of the Koppen climate classification as a benchmark for the FAO56-PM method on a monthly timescale. The study concluded that the HS equation could provide satisfactory performance in different climate regimes (i.e., arid, semiarid, temperate, cold, and polar climates). An alternative approach was proposed for regional adjustment of the Hargreaves-Samani coefficient by examining its relationship with temperature. Vanderlinden et al. (2004) presented the Adjusted Hargreaves-Samani coefficient (AHC) and evaluated the estimated daily ET_o against the ET_o obtained from FAO56-PM at meteorological stations located inland and in coastal areas. The adjusted Hargreaves-Samani coefficient (AHC), which is obtained through regression analysis, appears to increase at coastal stations and decrease across inland stations. Su et al. (2022) provided a comparison between several temperature- and radiation-based evapotranspiration methods for the evaluation of regional irrigation water demand. The calibrated Hargreaves-Samani and Priestley-Taylor methods can greatly increase efficiency and accuracy in estimating regional or worldwide irrigation water demand. Shirmohammadi-Aliakbarkhani and Saberali (2020) presented a reliable alternative ET_o models requiring fewer input data compared with the FAO Penman-Monteith method in Iran. In the study area, the temperature-based methods typically outperformed the radiation-based methods. More specifically, the Jensen-Haise method showed better performance for the warm growing season, while the Hargreaves-Samani method was the best performance for the cool growing season.

In recent studies, machine learning (i.e., artificial neural network (ANN), extreme learning machine (ELM), wavelet neural network (WNN)) and Bayesian theory have been used to estimate parameters of the HS equation and forecast the ET_0 (Falamarzi et al., 2014; Feng et al.,

2017, 2016; Kang et al., 2022; Traore et al., 2016). Traore et al. (2016) employed an ANN for short-term forecasting of ET_o using climate information retrieved from climate forecast products in the public domain. Feng et al. (2017) calibrated the HS equation using Bayesian theory in the Sichuan basin of southwest China. The study concluded that the calibrated HS equation still overestimated ET_o at different temporal scales, although the calibrated HS equation provided a better fit for the Penman-Monteith equation than that of the traditional HS equation. Since the HS equation was empirically developed based on data from arid to sub-humid conditions, it may not fit well in areas far different from those considered for its calibration, as is the case for humid climates (Razei and Pereira, 2013). Almorox et al. (2015) evaluated the temperature-based evapotranspiration equation for each Köppen climate zone and found that the HS equation performed best in most climatic zones, including temperate climates in which the study area (i. e., South Korea) was located.

Here a surrogate model based on the HS equation is developed, considering both daily ET_o and monthly and annual ET_o, which was referred to as "the multi-scale surrogate model". Most studies locally recalibrated parameters of the HS equation using daily temperature data as input. The existing studies consider only a single temporal scale for estimating ET_o, named "the single-scale surrogate model". Since the single-scale surrogate model solely uses temperature data as an independent variable, there may be a limitation in capturing the complex mechanism of the whole ET process using the HS equation. In this context, a modest difference in the daily estimates of ET_o may lead to a significant increase in the ET_o bias when estimated at aggregated monthly and annual scales. In this case, the estimated ET_o from a single-scale surrogate model can be difficult to use for water resource management. Therefore, our goal in this study was to develop a novel approach to the HS equation for a better fit for the ET_o obtained from the FAO56-PM method at multi-scale context as discussed above. A Bayesian modeling framework is also employed for parameter estimation. The proposed model was fit and tested using meteorological data obtained for the Han-River watershed in South Korea.

To the best of our knowledge, simultaneous parameter estimation at multiple time scales (i.e., daily, monthly and annual scales) in a surrogate model for the ET_o has not been properly explored in the literature. This study explores the following types of surrogate models in terms of estimating model parameters for improving the existing surrogate model for the ET_o:

Model–1: single-scale surrogate model using daily temperature data as a baseline model.

Model–2: multi-scale surrogate model using daily and monthly temperature data.

Model–3: multi-scale surrogate model using daily and annual temperature data.

Model–4: multi-scale surrogate model using daily, monthly, and annual temperature data.

Given the four different models, this study sought to explore the following questions by comparing the performance in terms of reproducing the underlying characteristics of the ET_o estimated from the FAO56-PM equation over the Han-river watershed in South Korea:

(1) *How effective is a locally recalibrated approach for the HS equation in terms of minimizing systematic error at different time scales?*
(2) *Can the HS model coefficients for the local daily ET_o be simultaneously estimated by considering aggregated ET_o at multiple time scales?*
(3) *Can a multi-scale surrogate modeling approach accurately reproduce the underlying distribution of the ET_o at different time scales?*

This paper is organized as follows. This section provides a brief background and the main purpose of this study. In Section 2, the study area and data used in this study are provided. The methodology for multi-scale surrogate models within a Bayesian modeling framework is
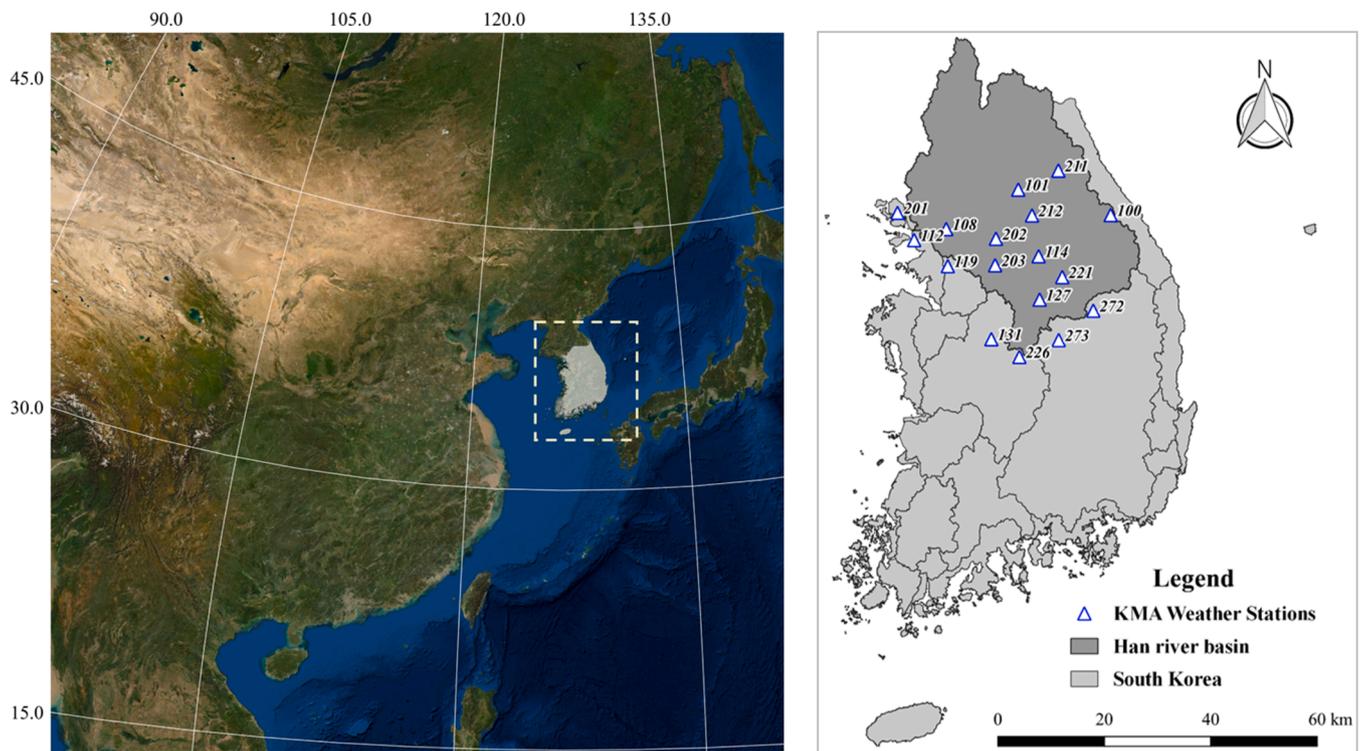
Fig. 1. Location of the Han River basin in the Korean Peninsula (left panel) and KMA weather stations (right panel).

Table 1
Weather stations in the Han River basin.

No.	Station ID	Station Name	Latitude (°)	Longitude (°)	Elevation (m)	Dataset length
1	100	*Daegwallyeong*	37.6771	128.7183	772.43	1971–2020
2	101	*Chuncheon*	37.9026	127.7357	75.82	1966–2020
3	108	*Seoul*	37.5714	126.9658	85.67	1961–2020
4	112	*Incheon*	37.4777	126.6249	68.99	1961–2020
5	114	*Wonju*	37.3375	127.9466	150.11	1973–2020
6	119	*Suwon*	37.2575	126.983	39.81	1964–2020
7	127	*Chungju*	36.9705	127.9525	114.85	1973–2020
8	131	*Cheongju*	36.6392	127.4407	58.7	1967–2020
9	201	*Ganghwa*	37.7074	126.4463	47.84	1973–2020
10	202	*Yangpyeong*	37.4886	127.4945	47.26	1973–2020
11	203	*Icheon*	37.264	127.4842	80.09	1973–2020
12	211	*Inje*	38.0599	128.1671	201.78	1973–2020
13	212	*Hongcheon*	37.6836	127.8804	140.2	1973–2020
14	221	*Jecheon*	37.1593	128.1943	264.62	1973–2020
15	226	*Boeun*	36.4876	127.7342	171.31	1973–2020
16	272	*Yeongju*	36.8718	128.5169	211.32	1973–2020
17	273	*Mungyeong*	36.6273	128.1488	173.01	1973–2020

described in Section 3. The results and discussion are presented with a comparison between models in Section 4, and the summary and conclusions for this study are finally given in Section 5.

2. Study area and meteorological data

This study was carried out on the basis of 17 ASOS (automated synoptic observing service) stations operated by the KMA (Korea Meteorological Administration) over the Han-River basin in South Korea. The daily solar radiation, air temperature (i.e., average, maximum and minimum temperature), humidity, and wind speed data were collected from ASOS stations through the website https://data. kma.go.kr. The ET_o was then obtained from the FAO56-PM method using a set of meteorological data listed above. The Han River is a major river in South Korea, which originates from the North and South Han rivers in the central part of the Korean peninsula (latitude

36°30'–38°55'N, longitude 126°24'–129°02'E) and flows through the capital Seoul. The basin area is about 25,594 km^2, covering a quarter of the land area of South Korea. The Han River receives annual average precipitation of about 1253 mm, and approximately 71% of the annual precipitation occurs during the summer season from June to September (So et al., 2017). Fig. 1 illustrates the location of the ASOS stations over the Han River basin and the stations selected for the study are described in Table 1.

3. Methodology

Among the temperature-based models for the ET_o, the HS equation was considered as a surrogate model for the FAO56-PM method. The parameters of the HS equation were estimated through a Bayesian theory to improve inference and better handle uncertainties in model parameters. The optimal model and its performance evaluation were
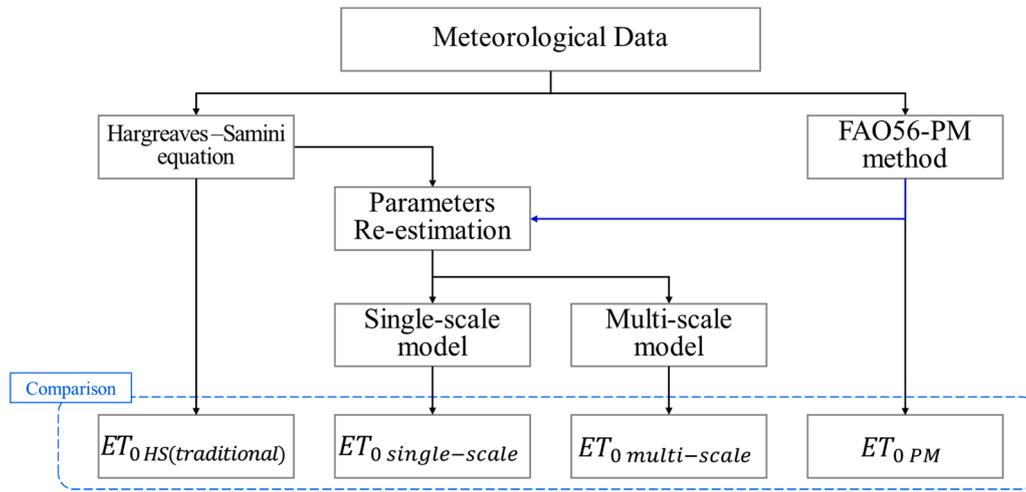
Fig. 2. A flowchart for the proposed modeling framework.

carried out using the deviance information criterion (DIC) value within a Hierarchical Bayesian modeling framework. A detailed procedure of the modeling process for the multi-scale surrogate model for ET_o is illustrated in Fig. 2. As seen in Fig. 1, the ET_o was obtained using the meteorological data in the Han River basin as input to both the FAO56-PM method and the traditional HS equation with fixed coefficients. Here, the ET_o estimated by the FAO56-PM method (ET_{oPM}) was used for locally recalibrating the HS equation. In a regression framework, the ET_{oPM} can be regarded as a dependent variable. In this study, the surrogate models can be classified as single-scale and multi-scale models in terms of temporal scale. In the final step, the results obtained from the different models are then demonstrated through goodness-of-fit measures and are compared with $ET_{o\ PM}$ and $ET_{o\ HG(traditional)}$.

3.1. Evapotranspiration equations

The evapotranspiration equations for estimating the ET (i.e., ET_o and ET_p) were empirically derived by considering physical relationships between ET and meteorological input variables. Xiang et al. (2020) provided details on the concepts, equations, and applications of reference evapotranspiration (ET_o) and potential evapotranspiration (ET_p). The ET_p equations can be categorized into four groups based on key concepts and variables: (i) Dalton, (ii) temperature, (iii) radiation, and (iv) combination. Similarly, the ET_o equations can be categorized into four types: (i) temperature, (ii) radiation, (iii) pan evaporation, and (iv) combination (Xiang et al., 2020). Various approaches to ET_o have been widely applied in the field of hydrology, including FAO56 Penman-Monteith (Allen et al., 1998), Hargreaves (Hargreaves and Samani, 1985, 1982), ASCE PM (Allen et al., 2005). This study mainly utilizes the physically-based FAO56-PM method and the temperature-based HS equation as a surrogate model for the FAO56-PM estimates.

3.1.1. FAO56-Penman Monteith methods

The FAO56-PM method has been demonstrated in many studies, and many institutions have recommended it as a representative approach for estimating the ET_o. The FAO56-PM is a physics-based ET estimation method informed by various meteorological factors (i.e., temperature, relative humidity, solar radiation, wind speed, etc.), and it can be written as follows:

$$ET_{o\ PM} = \frac{0.408\triangle(R_n - G) + \gamma \frac{900}{T+273}u_2(e_s - e_a)}{\triangle + \gamma(1 + 0.34u_2)} \quad (1)$$

where ET_o is the reference evapotranspiration (mm/day), R_n is the net radiation (MJ/m^2/day), G is the soil heat flux density (MJ/m^2/day), T is

the mean daily air temperature at 2 m height (℃), u_2 is the wind speed at 2 m height (m/s), e_s is the saturation vapor pressure (kPa), e_a is the actual vapor pressure (kPa), $e_s - e_a$ is the saturation vapor pressure deficit (kPa), Δ is the slope vapor pressure curve (kPa/℃), and finally, the γ is the psychrometric constant (kPa/℃) (Allen et al., 1998; Zotarelli et al., 2010).

3.1.2. Hargreaves-Samani equation

Hargreaves and Samani (1985) proposed the HS equation, which is based mainly on temperature, for the estimation of ET. The HS equation consists of four terms, the maximum temperature (T_{max}), minimum temperature (T_{min}), average temperature, and radiant energy:

$$ET_{o\ HS(traditional)} = 0.0023 \cdot RA\ (T℃ + 17.8)\ TR^{0.50} \quad (2)$$

where RA is the extraterrestrial radiation in mm/day, TR is the temperature range (or difference) between the maximum and minimum temperature in ℃, and T is the average temperature in ℃. There are three fixed parameters to define the relationship between the ET_o and the extraterrestrial radiation RA and air temperature T. Hargreaves (1994) presented radiant energy according to latitude and month. Allen et al. (1998) recommended the HS equation as an alternative method when the input data of the FAO56-PM method is insufficient.

Since only observed temperature data are required, the HS equation has been widely used for estimating and predicting ET_o in worldwide locations where the FAO56-PM method is not feasible due to insufficient and unreliable meteorological data (Hargreaves and Allen, 2003). It is also desirable to calibrate (or re-estimate) the associated parameters in Eq. (2) for the specific region of interest (Lee and Park, 2008; Moon et al., 2013), as the evapotranspiration might be influenced by geographical features and local characteristics (e.g., land use, altitude and so on).

The HS equation is generally calibrated using daily data, which might be insufficient to reproduce underlying characteristics at aggregated time scales (i.e., monthly and annual scales) as the parameter estimates are subject to the daily variations of the input meteorological data. Moreover, the most common approach to estimating the regression coefficients is the least square method, where parameter uncertainties may not be sufficiently addressed. In this regard, this study offers a model within a Bayesian inference modeling framework to better quantify the uncertainty associated with parameters in the HS equation. The Bayesian approach to the HS equation is presented in the following section.

3.2. Bayesian approach to the HS equation

The Bayesian inference is a statistical tool to update the model parameters with observed data from a prior distribution $p(\theta)$ to the posterior distribution $p(\theta|y)$ (Gelman et al., 2004). The Bayesian inference, which presents the parameters in the form of a probability distribution, can be used to quantitatively evaluate the uncertainty associated with the model parameters. To make probability statements about parameter θ given data y, a joint probability distribution $p(\theta, y)$ for θ and y is used. The joint probability density function can be written as a product of two probability densities that are often referred to as the prior distribution $p(\theta)$ and the likelihood $p(y|\theta)$:

$$p(\theta, y) = p(\theta)p(y|\theta) \tag{3}$$

The posterior distribution $p(\theta|y)$ can be defined as a conditional distribution by adopting Bayes' rule (Gelman et al., 2004), given a prior distribution $p(y)$ and a likelihood $p(y|\theta)$, with observations y:

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \tag{4}$$

A general form of the HS equation-based surrogated model for estimating evapotranspiration can be found in Eq. (5). There are three parameters to be estimated with the observed temperature and the ET_o obtained from the FAO-56 PM method. Several publications have presented improved parameters for HS equation in the different climatic zone (Gavilán et al., 2006; Martínez-Cob and Tejero-Juste, 2004; Raziei and Pereira, 2013; Todorovic et al., 2013; Yang et al., 2021). Among many others, the parameters suggested in the traditional HS equation (Hargreaves and Samani, 1985) are 0.0023, 17.8 and 0.5 for α, β and γ, respectively:

$$ET_{o\,HS} = \alpha \cdot RA \times (T + \beta) \times (T_{max} - T_{min})^{\gamma} \tag{5}$$

The Bayesian information criterion (BIC) is used to select the best-fitted distribution among various probability distributions. In this study, the ET_o follows a lognormal distribution (LN) with mean μ and standard deviation σ, as shown in Eq. (6). The BIC values for each distribution is illustrated in the supplementary material (Table S1). The models with lower BIC are generally preferred. In Eq. (7), $y^{day}_{obs(i,j,k)}$ denotes the ET_o value (i.e., ET_{oPM}) for the day i, month j, and year k, estimated from the FAO56-PM method, and it is assumed to be a log-normal distribution with a mean of $\mu_{day(i,j,k)}$ and standard deviation of σ. Here, the mean $\mu_{day(i,j,k)}$ is a function of the temperature range between the maximum and minimum temperature and the average temperature T. In Bayesian inference, it is critical to choose a reasonable prior distribution for parameters. When sufficient data is available but prior information about parameters is weak, a non-informative prior distribution can be used. The prior distributions of each parameter (i.e., $\alpha, \beta, \gamma, \sigma$) are defined in Eqs. 8 to 11. The α, β and γ represent parameters of the HS equation to be estimated with prior distributions (e.g., normal distribution (N) and gamma distribution (Γ)). Sufficient data (i.e., 40-year*365 days = 14,600 days) are available to estimate the four parameters. Furthermore, one can adopt non-informative prior distributions for the parameters, as proposed in (Gelman et al., 2004), considering our vague information about the parameters.

$$LN(\mu, \sigma) \sim \frac{1}{y\sigma\sqrt{2\pi}} exp\left\{ -\frac{(\log(y) - \mu)^2}{2\sigma^2} \right\} \tag{6}$$

$$y^{day}_{obs(i,j,k)} \sim LN(\mu_{day(i,j,k)}, \sigma^2),\ i = 1, ..., I(j),\ j = 1, ..., 12,\ k = 1, ..., N \tag{7a}$$

$$\mu_{day(i,j,k)} = \alpha \cdot RA_{(j)} \times (T_{(i,j,k)} + \beta) \times (T_{max(i,j,k)} - T_{min(i,j,k)})^{\gamma} \tag{7b}$$

$$\alpha \sim N(0, 10^6) \tag{8}$$

Table 2
Four types of experimental models considered in the study.

	Model-1	Model-2	Model-3	Model-4
Input data	Daily	Daily, Monthly	Daily, Annual	Daily, Monthly, Annual
Parameters	$\alpha, \beta, \gamma, \sigma$	$\alpha, \beta, \gamma, \sigma_{day}, \sigma_{month}$	$\alpha, \beta, \gamma, \sigma_{day}, \sigma_{year}$	$\alpha, \beta, \gamma, \sigma_{day}, \sigma_{month}, \sigma_{year}$
Number of parameters	4	5	5	6

$$\beta \sim N(0, 10^6)|(T_{min}, \sim) \tag{9}$$

$$\gamma \sim \Gamma(0.01, 0.01) \tag{10}$$

$$\sigma \sim \Gamma(0.1, 1) \tag{11}$$

where, I(j) represents the number of days in month j.

At some stations, the relatively low daily average temperature can lead to a negative term $(T_t + \beta)$ in Eq. 5, resulting in a negative ET_o. In this regard, as shown in Eq. 9, this study adopts a truncated normal distribution for the parameter β to be sampled from a constraint on the daily minimum temperature (T_{min}). The joint posterior distribution of the set of parameters θ = [$\alpha, \beta, \gamma, \sigma$] can be formulated as the product of the likelihood (Eq. 12) and priors as shown in Eq. 13:

$$p(\mathbf{Y}|\boldsymbol{\theta}) = \prod_{k=1}^{N}\prod_{j=1}^{12}\prod_{i=1}^{I(j)} LN\Big(\mu_{day(i,j,k)}; \alpha\, RA_{(j)} \times (T_{(i,j,k)} + \beta)$$
$$\times (T_{max(i,j,k)} - T_{min(i,j,k)})^{\gamma},\quad \sigma^2\Big) \tag{12}$$

$$p(\boldsymbol{\theta}|\mathbf{Y}) \propto p(\mathbf{Y}|\boldsymbol{\theta}) \cdot N(\alpha|0, 10^6) \cdot N(\beta|0, 10^6)|(T_{min}, \sim) \cdot \Gamma(\gamma|0.01, 0.01) \cdot \Gamma(\sigma|0.1, 1) \tag{13}$$

Here, the surrogate model with the use of daily ET_o is referred to as a single-scale model that is a reference model for comparison with the multi-scale models developed in this study. The set of parameters $\boldsymbol{\theta}$ are estimated by maximizing the joint posterior distribution, as illustrated in Eq. 13, through the Gibbs sampling approach (Gilkst et al., 1995; Kim et al., 2015; Kwon et al., 2008; Lima et al., 2018; So et al., 2017), which is a special case of the Markov Chain Monte Carlo (MCMC) approach. The primary competitive advantage of Gibbs sampling is that it can provide an efficient way to deal with multidimensional data of the complex problem compared to other approaches such as the Metropolis-Hastings (MH) algorithm.

3.3. Multi-scale Surrogate model

The main purpose of this study is to improve the accuracy of ET_o estimation at multiple temporal scales (i.e., daily, monthly, and annual) by recalibrating the single-scale model and its extension to the multi-scale model. More specifically, the multi-scale model is designed to better reproduce the ET_o across different temporal scales by simultaneously maximizing the likelihood function associated with the ET_o at multiple temporal scales. The proposed surrogate models for the ET_o can be classified into 4 models based on the different combinations of temporal scales, as summarized in Table 2.

The key step of this study is to simultaneously maximize the likelihood functions of daily, monthly and annual ET_o (as written in Eqs. 14a-14c) in a model of the multi-scale model by sequentially summing the monthly and annual evapotranspiration from the estimated daily ET_o, as illustrated in Eqs. 15a-15c, respectively. More specifically, the monthly and annual ET_o (i.e., $y^{month}_{obs(j,k)}, y^{year}_{obs(k)}$) are assumed to follow a lognormal distribution with the summed ET_o as the mean $(\mu_{month(j,k)}, \mu_{year(k)})$ and standard deviation $(\sigma_{month}, \sigma_{year})$ (Eqs. 14b, 14c). The multi-scale models

have a different number of parameters according to the temporal scales considered in each model, as illustrated in Table 2. More specifically, the multi-scale models consider daily, monthly and annual ET_o when estimating a set of parameters by simultaneously maximizing likelihood functions at three different temporal scales in an integrated manner, as summarized in Eqs. 14a-14c. The prior distributions for regression parameters (α, β and γ), including the standard deviation ($\sigma_{day}, \sigma_{month}, \sigma_{year}$) of the lognormal distribution at each temporal scale can be defined as Eqs. 16 to 21. Like the single scale model represented in the previous section, there are sufficient data to estimate the regression parameters for the mean and standard deviation of the lognormal distribution. Thus, non-informative prior distributions for the parameters were considered as proposed in (Gelman et al., 2004). More specifically, normal prior distributions (N) were used for the regression parameters α, β while the Gamma prior distribution (Γ)) was adopted for the parameter γ. For the standard deviation ($\sigma_{day}, \sigma_{month}, \sigma_{year}$) of the lognormal distribution, the Gamma prior distribution (Γ) which is a conjugate distribution of the lognormal distribution, was used. The conjugate prior distribution is the case that the prior and posterior distributions are in the same probability distribution family, leading to a computational advantage in the process of the parameter estimation (Gelman et al., 2004).

$$y_{obs(i,j,k)}^{day} \sim LN\left(\mu_{day(i,j,k)}, \sigma_{day}^2\right), i=1,\dots,I(j), j=1,\dots,12, k=1,\dots,N \tag{14a}$$

$$y_{obs(j,k)}^{month} \sim LN\left(\mu_{month(j,k)}, \sigma_{month}^2\right), j=1,\dots,12, k=1,\dots,N \tag{14b}$$

$$y_{obs(k)}^{year} \sim LN\left(\mu_{year(k)}, \sigma_{year}^2\right), k=1,\dots,N \tag{14c}$$

$$\mu_{day(i,j,k)} = \alpha \cdot RA_{(j)} \times \left(T_{(i,j,k)} + \beta\right) \times \left(T_{\max(i,j,k)} - T_{\min(i,j,k)}\right)^\gamma, i=1,\dots,I(j), j=1,\dots,12, k=1,\dots,N \tag{15a}$$

$$\mu_{month(j,k)} = \sum_{k=1}^{N}\sum_{j=1}^{12}\sum_{i=1}^{I(j)}\mu_{day(i,j,k)}, j=1,\dots,12, k=1,\dots,N \tag{15b}$$

$$\mu_{year(k)} = \sum_{k=1}^{N}\sum_{j=1}^{12}\mu_{month(j,k)}, k=1,\dots,N \tag{15c}$$

$$\alpha \sim N(0, 10^6) \tag{16}$$

$$\beta \sim N(0, 10^6)|(T_{min}, \sim) \tag{17}$$

$$\gamma \sim \Gamma(0.01, 0.01) \tag{18}$$

$$\sigma_{day} \sim \Gamma(0.1, 1) \tag{19}$$

$$\sigma_{month} \sim \Gamma(0.1, 1) \tag{20}$$

$$\sigma_{year} \sim \Gamma(0.1, 1) \tag{21}$$

Among the multi-scale models, the likelihood function and joint posterior distribution for Model-4 are expressed as Eqs. 22 and 23 using Bayes' rule, respectively.

$$p(y|\theta) = \prod_{k=1}^{N}\prod_{j=1}^{12}\prod_{i=1}^{I(j)} LN\left(\mu_{day(i,j,k)}; \alpha \cdot RA_{(j)} \times \left(T_{(i,j,k)} + \beta\right)\right.$$
$$\times \left(T_{\max(i,j,k)} - T_{\min(i,j,k)}\right)^\gamma, \sigma_{day}^2\right) \times LN\left(\mu_{month(j,k)}; \sum_{k=1}^{N}\sum_{j=1}^{12}\right.$$
$$\left.\times \sum_{i=1}^{I(j)}\mu_{day(i,j,k)}, \sigma_{month}^2\right) \times LN\left(\mu_{year(k)}; \sum_{k=1}^{N}\sum_{j=1}^{12}\mu_{month(j,k)}, \sigma_{year}^2\right) \tag{22}$$

$$p(\theta|y) \propto \prod_{k=1}^{N}\prod_{j=1}^{12}\prod_{i=1}^{I(j)} LN\left(\mu_{day(i,j,k)}; \alpha \cdot RA_{(j)} \times \left(T_{(i,j,k)} + \beta\right)\right.$$
$$\times \left(T_{\max(i,j,k)} - T_{\min(i,j,k)}\right)^\gamma, \sigma_{day}^2\right) \times LN\left(\mu_{month(j,k)}; \sum_{k=1}^{N}\sum_{j=1}^{12}\right.$$
$$\left.\times \sum_{i=1}^{I(j)}\mu_{day(i,j,k)}, \sigma_{month}^2\right) \times LN\left(\mu_{year(k)}; \sum_{k=1}^{N}\right.$$
$$\left.\times \sum_{j=1}^{12}\mu_{month(j,k)}, \sigma_{year}^2\right) \cdot N(\alpha|0, 10^6) \cdot N(\beta|0, 10^6)|(T_{min},$$
$$\sim) \cdot \Gamma(\gamma|0.01, 0.01) \cdot \Gamma(\sigma_{day}|0.1, 1) \cdot \Gamma(\sigma_{month}|0.1, 1) \cdot \Gamma(\sigma_{year}|0.1, 1) \tag{23}$$

Here, the deviance information criterion (DIC) was used to explore the model performance. The DIC value was computed to measure the fitness of a candidate model for a given datum by penalizing the fitness of the model for model complexity defined by the effective number of parameters. In other words, overfitting may occur as the number of parameters increases, so that DIC gives a penalty for the number of parameters. DIC is used to determine the best model among competing models in Bayesian inference, where the posterior distribution of the model is obtained through the MCMC method.

$$D(\theta) = -2\log(p(y|\theta)) \tag{24}$$

$$p_D = \overline{D(\theta)} - D(\overline{\theta}) \tag{25}$$

$$DIC = p_D + \overline{D(\theta)} \tag{26}$$

The deviance $D(\theta)$ expressed in Eq. 24 is calculated by multiplying the log-likelihood by -2 and using the posterior mean of the parameters. p_D means the effective number of parameters that are calculated as Eq. 25. DIC is calculated as the sum of $D(\theta)$ and p_D as in Eq. 26. The smaller value in the DIC value confirmed the improved performance.

3.4. Goodness-of-fit measures for model validation

Statistical goodness-of-fit measures (GoF) were used to evaluate the ET_o estimated through the models during the calibration and validation processes. Here, the estimated ET_{oPM} obtained from the FAO56-PM method approach is used to validate the estimated $ET_{oHS(traditional)}$ (i.e., traditional regression coefficients in HS equation) and simulated ET_{oModel} (i.e., recalibrated regression coefficients). The model performance during the calibration and validation phase was measured using the correlation coefficient (CC) – representing the degree of a linear relationship between observed (O) and simulated (S) values. A stronger relationship was obtained as CC approaches 1. Model performance was also measured using the root mean square error (RMSE) – a value representing the difference between the observed and the estimated values. RMSE values closer to zero indicate the good fitness of the model. The third metric to gage model performance was the index of agreement (IoA) – a standardized value of the degree of error in model predictions. An IoA of 1 indicates perfect agreement. The fourth GoF indicator is the percentage bias. The percentage bias can show under- or over-estimation for the model results. Eqs. 27 to 30 show the statistical indicators considered in this study.

$$CC = \frac{\sum_{i=1}^{N}(S_i - \overline{S})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{N}(S_i - \overline{S})^2 \sum_{i=1}^{N}(O_i - \overline{O})^2}} \tag{27}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - O_i)^2} \tag{28}$$
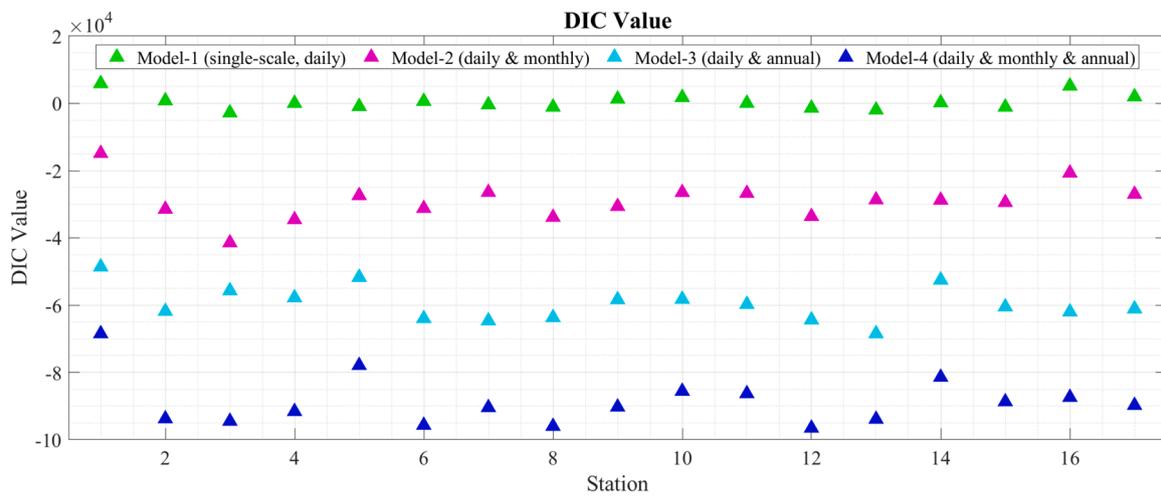
Fig. 3. Comparison of deviance information criterion (DIC) values for four models defined in Table 2 over all the stations. The DIC values are shown along the y-axis, while the number on the x-axis shows the gauge station ID, as summarized in Table 1. The green-filled triangles are DIC values of the single-scale model (Model-1), while the magenta, cyan, and blue-filled triangles represent those of multi-scale models (Model-2, Model-3 and Model-4), respectively. The DIC value for Model-4, including evapotranspiration at three temporal scales (i.e., daily, monthly and annual ET_o), is the smallest for all gauge stations.

$$\text{IoA} = 1 - \frac{\sum_{i=1}^{N}|O_i - S_i|^i}{\sum_{i=1}^{N}|S_i - \overline{O}| + |O_i - S_i|^i} \qquad (29)$$

$$\text{Percentage } Bias = \frac{\sum_{i=1}^{N}(S_i - O_i)}{\sum_{i=1}^{N}(O_i)} \times 100 \qquad (30)$$

The daily, monthly, and annual ET_o obtained from both the traditional HS equation and calibrated models through the re-parameterization process were compared with the ET_o data obtained from the FAO56-PM method. The periods selected for calibration and validation were 1973–2010 and 2011–2020, respectively. The model validation was conducted in the out-of-sample period where data were not used in the calibration process.

4. Results

In this section, results from the multi-scale model are compared with ET_{oPM} calculated by the FAO56-PM method at different time scales. As indicated, the DIC values for each model can be the benchmark for choosing the optimal model. The multi-scale models are fully compared to both the single-scale model and the traditional HS equation for the calibration and validation phase in the context of cross-validation.

4.1. Model comparison based on DIC values

To statistically examine the effectiveness of multi-scale models considering multiple temporal scales accounting for model performance and model complexity, DIC values are computed for the four types of models fit in each gauge station (Fig. 3). The proposed multi-scale models (Model-2, Model-3 and Model-4) outperform the single-scale model in terms of model fitness based on the DIC value. Model-3, which simultaneously includes the evapotranspiration series at daily and annual scales in estimating model parameters, can be more effective than Model-2 considering the evapotranspiration series at daily and monthly scales. The DIC value for Model-4, including evapotranspiration at three temporal scales (i.e., daily, monthly and annual ET_o), for all stations, is the smallest, making it the superior model compared to the others.

Table 3
The mean of estimated regression parameters and standard deviation for Model-4, including evapotranspiration series at three temporal scales, during the calibration process.

Station No.	Regression Parameters and Standard Deviation					
	α	β	γ	σ_{day}	σ_{month}	σ_{year}
100	0.0020	33.2	0.34	0.3085	0.1359	0.0516
101	0.0020	20.4	0.43	0.2558	0.0969	0.0414
108	0.0016	32.3	0.49	0.2229	0.0801	0.0541
112	0.0026	36.0	0.25	0.2548	0.0913	0.0466
114	0.0021	20.6	0.40	0.2435	0.1137	0.0567
119	0.0028	20.3	0.30	0.2540	0.0976	0.0393
127	0.0024	19.8	0.36	0.2455	0.1156	0.0395
131	0.0028	19.3	0.33	0.2376	0.0957	0.0417
201	0.0026	21.1	0.34	0.2577	0.0975	0.0450
202	0.0023	23.4	0.34	0.2657	0.1105	0.0442
203	0.0025	19.3	0.36	0.2479	0.1130	0.0448
211	0.0024	22.2	0.36	0.2342	0.0966	0.0412
212	0.0023	20.7	0.36	0.2340	0.1167	0.0370
221	0.0024	20.6	0.35	0.2493	0.1059	0.0542
226	0.0027	19.7	0.32	0.2364	0.1081	0.0453
272	0.0020	34.3	0.33	0.2976	0.1154	0.0372
273	0.0015	31.0	0.48	0.2692	0.1058	0.0408

4.2. Model performance comparison

The mean of estimated parameters for the superior model (i.e., Model-4) are summarized in Table 3 as a representative model. Moreover, the posterior distributions of regression parameters (Model-4) for all stations can be found in supplementary material (Table S2). The regression parameters (α, β, γ) and the standard deviations (σ_{day}, σ_{month}, σ_{year}) are in the ranges 0.0015–0.0028, 19.3–36, 0.25–0.49, 0.2229–0.3085, 0.0801–0.1359, and 0.037–0.0567, respectively, as obtained from the posterior distribution. The estimated parameters β and γ showed distinctive parameter space from that of the traditional HS equation (e.g., 17.8 and 0.50 for β and γ), while the parameter α is comparable to that of the traditional HS equation ($\alpha = 0.0023$) as proposed by (Hargreaves and Samani, 1985). In addition, the credible intervals can be presented from the posterior distributions of regression parameters. The credible interval of models was obtained by simulating the ET_o with parameters of the posterior distributions and demonstrated as the boxplots in the supplemental material (Fig. S1).

In this section, we describe the results of the optimal multi-scale model (Model-4) with a systematic comparison of the existing tradi-
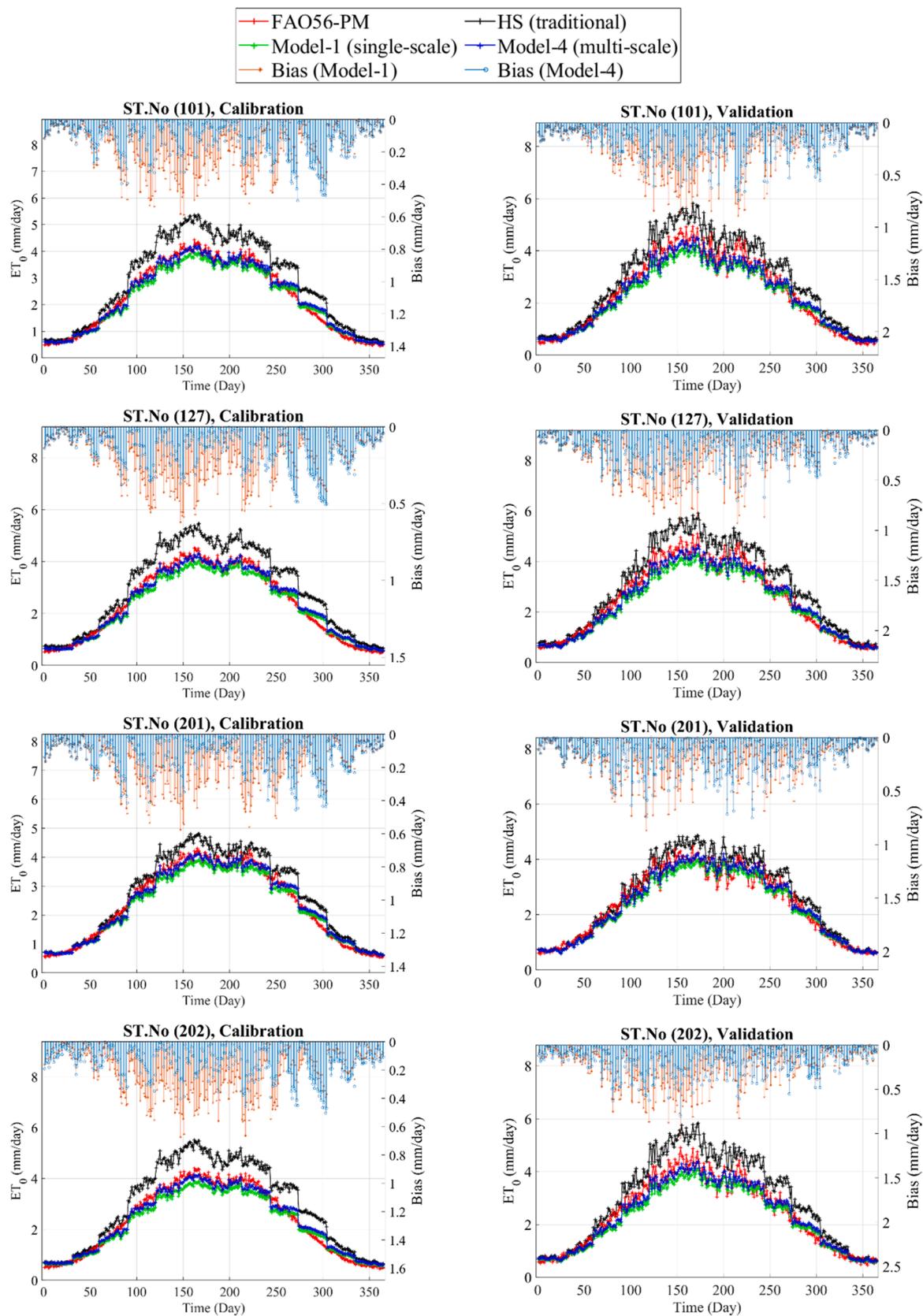
Fig. 4. Comparison of daily averaged ET_o values over the calibration (1973–2010) and validation (2011–2020) periods from different models for 8 stations (i.e., ST. Nos. 101, 127, 201, 202, 203, 212, 226 and 272). The daily averaged ET_o values from FAO56-PM (solid red line), traditional HS equation (solid black line), Model-1 (single-scale, solid green line), and Model-4 (multi-scale, blue solid line) are compared. The biases, representing difference values between evapotranspiration FAO56-PM and proposed models (Model-1 and Model-4), are shown as the orange (Model-1) and sky-blue (Model-4) bars.
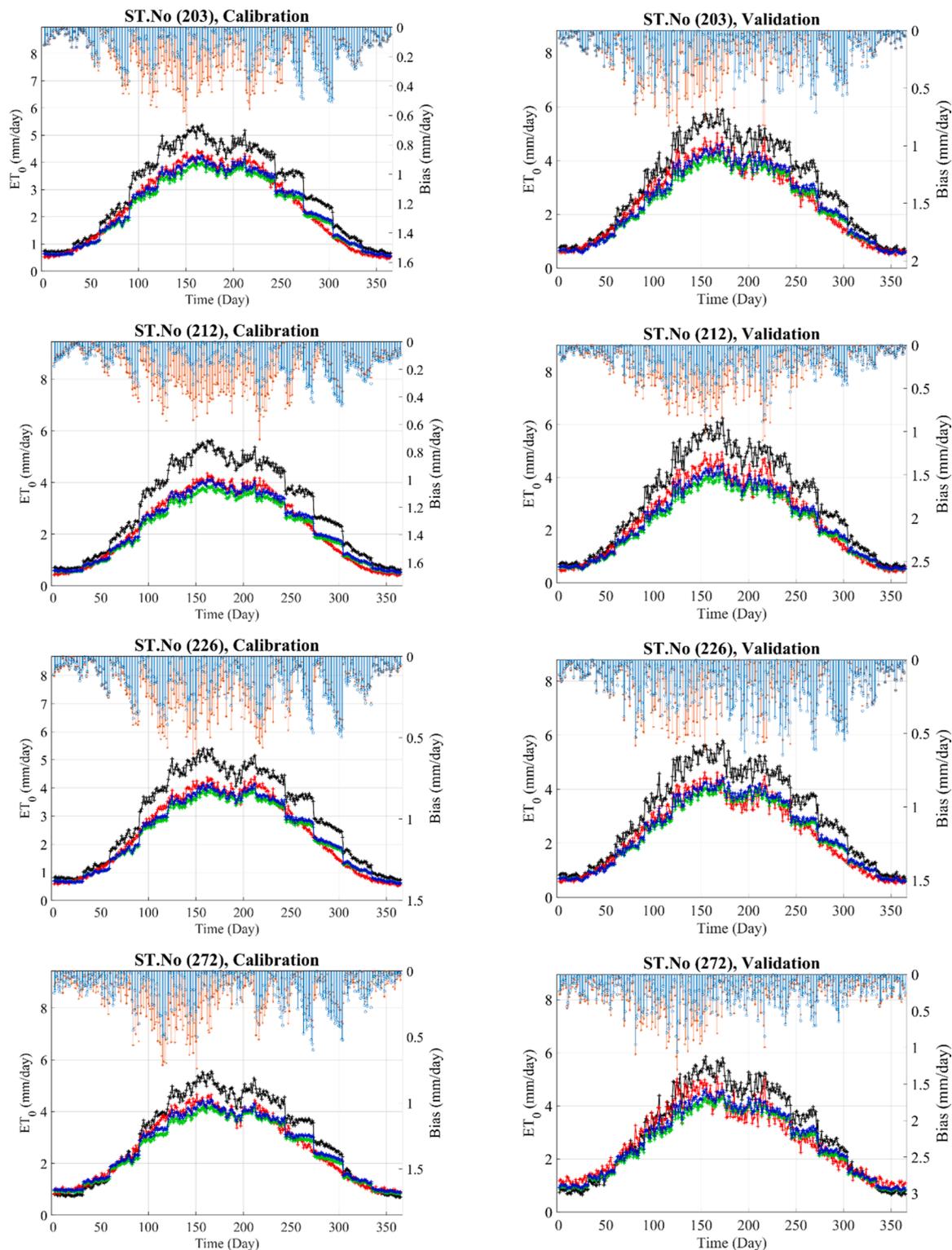
Fig. 4. (*continued*).

tional HS equation and single-scale model (Model-1). The estimated daily average ET_o series was obtained from the traditional HS equation, single-scale model, and multi-scale model in representative stations of the study area. These are presented for a comparison with the FAO56-PM in Fig. 4. The effectiveness of the recalibration process for the HS equation as an intervention was confirmed through the distinctive difference in the bias between $ET_{oHS(traditional)}$ (i.e., black solid line) and $ET_{o\ single-scale}$ (i.e., green solid line) for the calibration and validation

phases, as shown in Fig. 4. Parameter recalibration is clearly desirable to effectively capture the geographical impact on the estimation of evapotranspiration in a specific region of interest. Similarly, the effectiveness of the multi-scale model can be explained by the noticeable reduction in the bias between $ET_{o\ single-scale}$ (i.e., solid green line) and $ET_{o\ multi-scale}$ (i.e., solid blue line), especially during the summer season for both the calibration and validation phase.

To be more specific, $ET_{o\ HS(traditional)}$ (i.e., solid black line), which uses
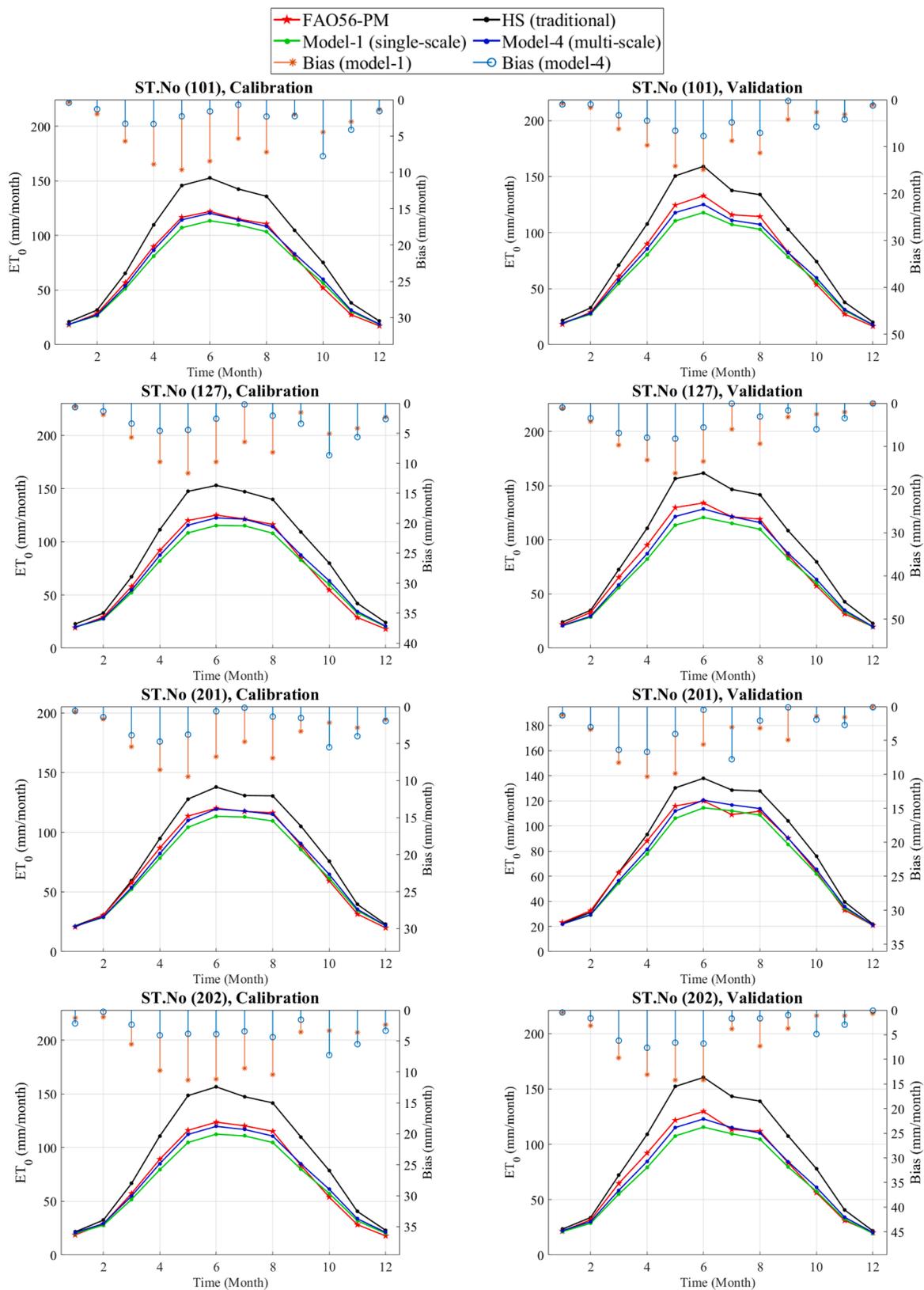
Fig. 5. Comparison of monthly aggregated ET_o values over the calibration (1973–2010) and validation (2011–2020) periods from different models for 8 stations (i.e., ST. Nos. 101, 127, 201, 202, 203, 212, 226 and 272). The monthly averaged ET_o values from FAO56-PM (red-solid line), traditional HS equation (black solid line), Model-1 (single-scale, green solid line) and Model-4 (multi-scale, blue solid line) are compared. The bias values, representing difference values between evapotranspiration FAO56-PM and proposed models (Model-1 and Model-4), are shown as the orange (Model-1) and sky-blue (Model-4) bars.
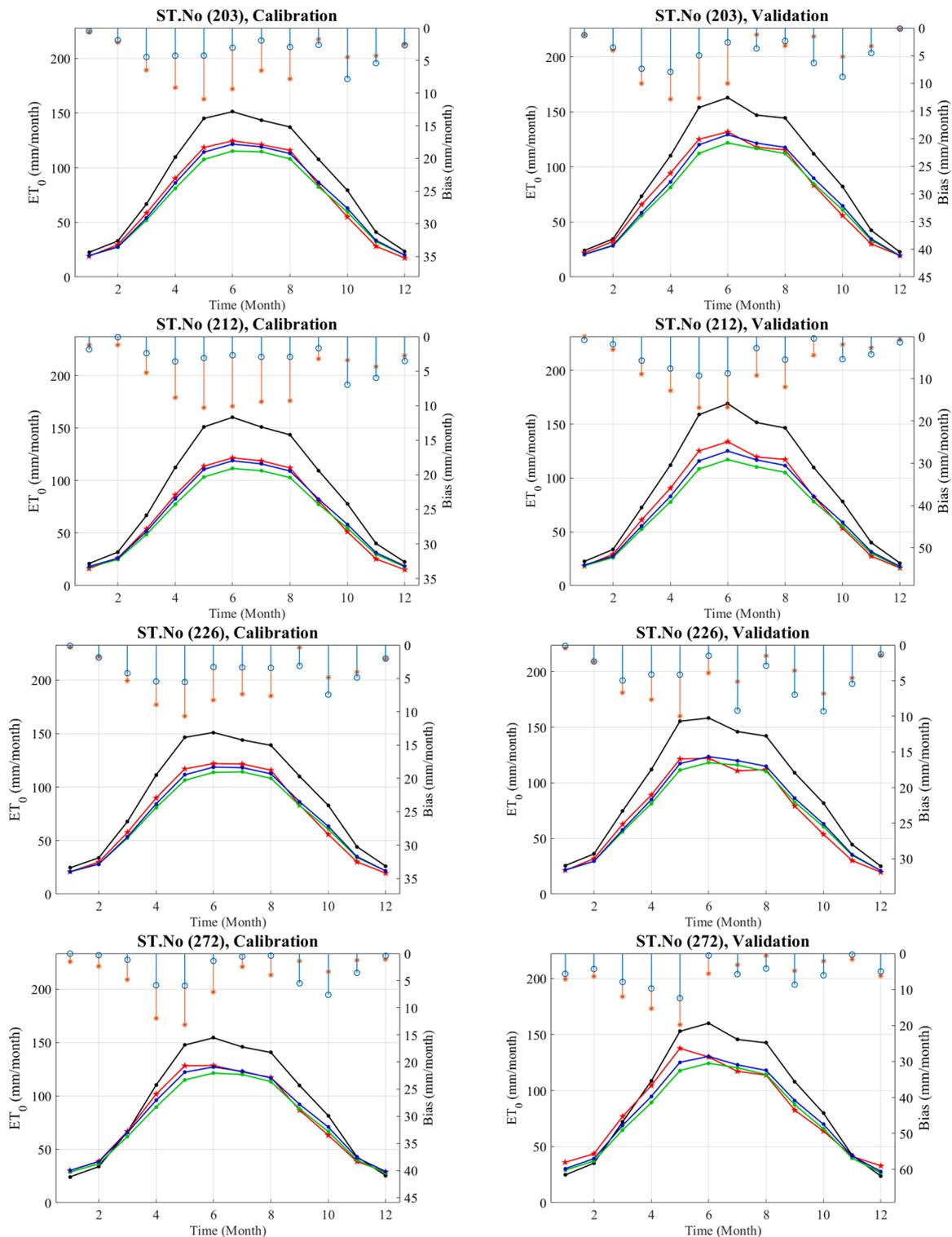
Fig. 5. (*continued*).

the fixed parameters without recalibration, clearly leads to an over-estimation at a daily timescale over all the stations. The single-scale model ($ET_{o\ single-scale}$, solid green line) reproduces the $ET_{o\ PM}$ (i.e., solid red line) fairly well during the dry season, while it fails to provide an accurate estimation during the wet season. In contrast, the multi-scale model ($ET_{o\ multi-scale}$, solid blue line) offers an accurate approximation to the $ET_{o\ PM}$ compared to the single-scale model at a daily timescale. Further, the aggregated monthly results over the entire period demonstrate a clear difference in model bias for both the calibration and

validation phase, as shown in Fig. 5. As already demonstrated in the daily timescale, the monthly average value ($ET_{o\ HS(traditional)}$, solid black line) obtained from the traditional HS equation showed a tendency to systematically overestimate at a monthly timescale for all stations and years. Overall, the multi-scale model results ($ET_{o\ multi-scale}$, solid blue line) present a more accurate estimation of the reference evapotranspiration compared to those of the single-scale. It was also observed that estimated ET_o shows a strong decreasing tendency during the monsoon period, especially in July. The boxplot in Fig. 5 represents the entire
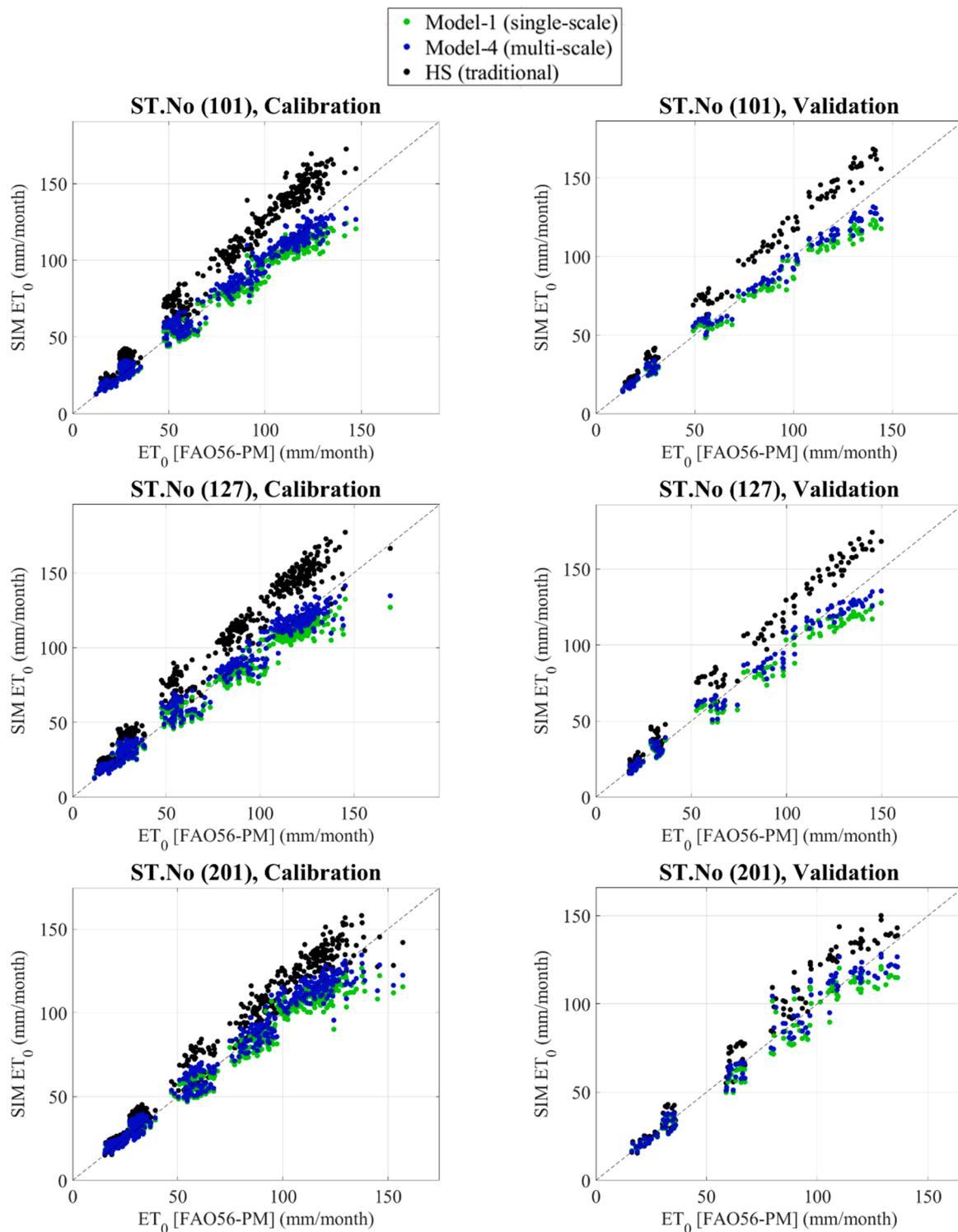
Fig. 6. Scatter plots on monthly ET_o results over the calibration (1973–2010) and validation (2011–2020) period from different models for 8 stations (i.e., ST. Nos. 101, 127, 201, 202, 203, 212, 226 and 272). The relationship between monthly $ET_{o\,PM}$ and the simulated ET_o results obtained from the traditional HS equation (black-filled circles), Model-1 (single-scale model, the green-filled circles) and Model-4 (multi-scale model, blue-filled circles) are represented.

range of $ET_{o\,PM}$ obtained from the entire period for each month, and the range becomes larger during the spring and summer season, resulting in a relatively larger bias in that season while being smaller during the winter season. More specifically, the bias of the multi-scale model can be more extensive in spring and summer due to the higher variability in temperature and precipitation. The enhancement of the efficacy of the

recalibration process at multiple temporal scales can be more easily explained and justified through a scatter plot representing the correlation with monthly $ET_{o\,PM}$, as illustrated in Fig. 6.

Finally, the aggregated annual ET_o time series over the entire period are illustrated in Fig. 7 for both the calibration and validation phase. The multi-scale model for estimating annual ET_o shows similar efficacy and
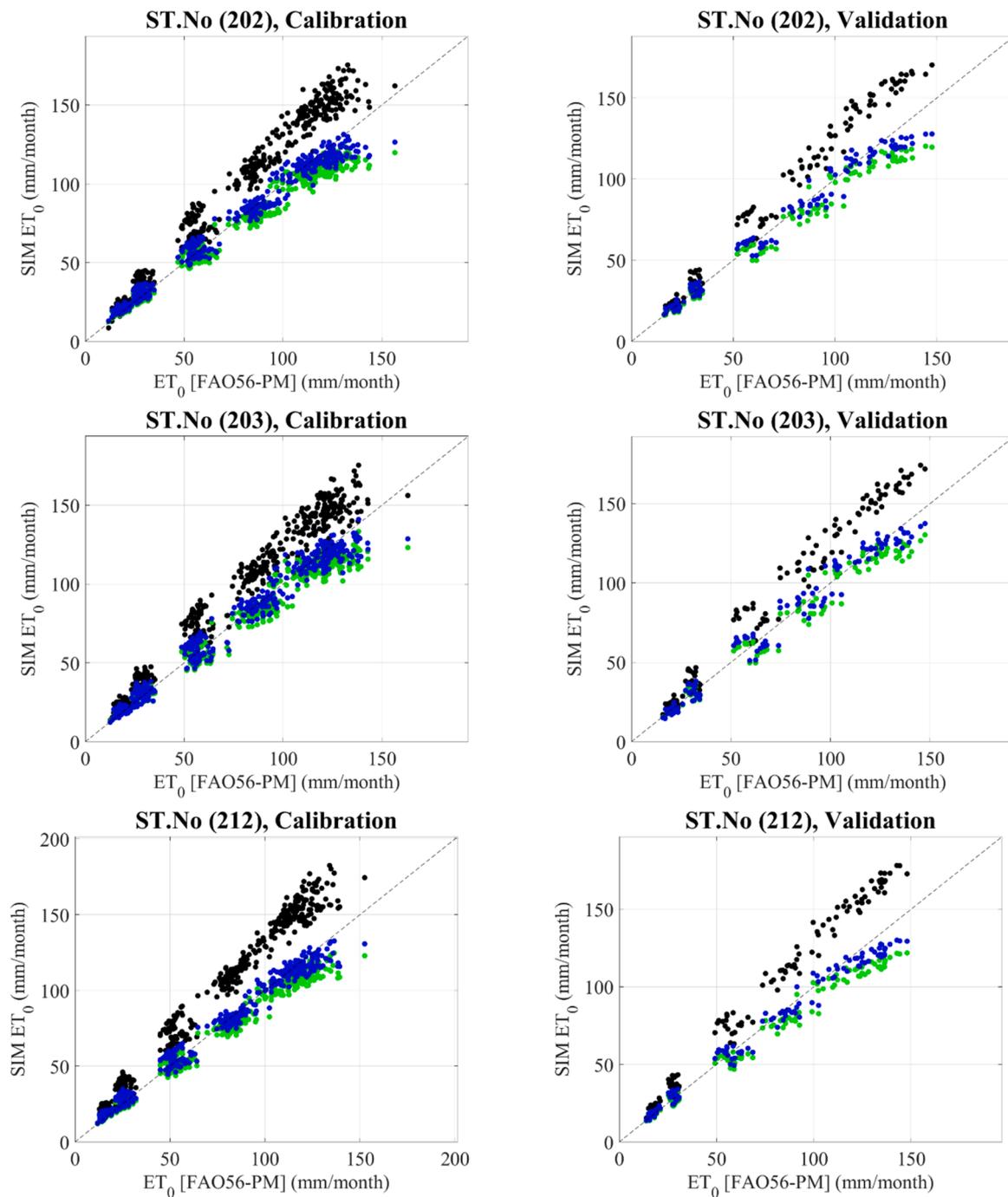
Fig. 6. (*continued*).

superior performance as its preceding temporal scales with respect to reduction in the bias but noticeable interannual variation due to climate variability. Although the temperature-based surrogate model for the estimation of the reference evapotranspiration has limitations in reproducing all the associated properties, the annual $ET_{o\ multi-scale}$ estimated from the multi-scale model shows competitive performance compared to the traditional HS equation ($ET_{o\ HS(\text{traditional})}$) and the single-scale model ($ET_{o\ single-scale}$). Moreover, we found that the multi-scale model demonstrates a high correlation between $ET_{o\ multi-scale}$ (blue filled circle) and $ET_{o\ PM}$, indicating that the recalibration of parameters at multiple temporal scales could be primarily responsible for the differences observed in the single-scale model (green filled circle) and the traditional HS equation (black filled circle), as illustrated in Fig. 8.

4.3. Model validation with statistical GoF measures

It has been acknowledged that daily ET_o from the HS equation can be strongly affected by the large variabilities in temperature, wind and solar radiation, leading to an increase in bias. Thus, the smoothed aggregation approach to the HS equation using five-day or longer climate variables has been recommended as an alternative method (Hargreaves and Allen, 2003). However, daily estimates are typically used to simulate daily long-term streamflow sequences through a continuous rainfall-runoff model. As noted, this study aims to reproduce the underlying characteristics of the ET_o across different temporal scales. In these contexts, the monthly and annual reference evapotranspiration estimates are evaluated using statistical GoF measures (i.e., CC, RMSE, IoA, percentage bias). Statistical GoF for the calibration (1973–2010)

ST.No (226), Calibration — **ST.No (226), Validation** — **ST.No (272), Calibration** — **ST.No (272), Validation**
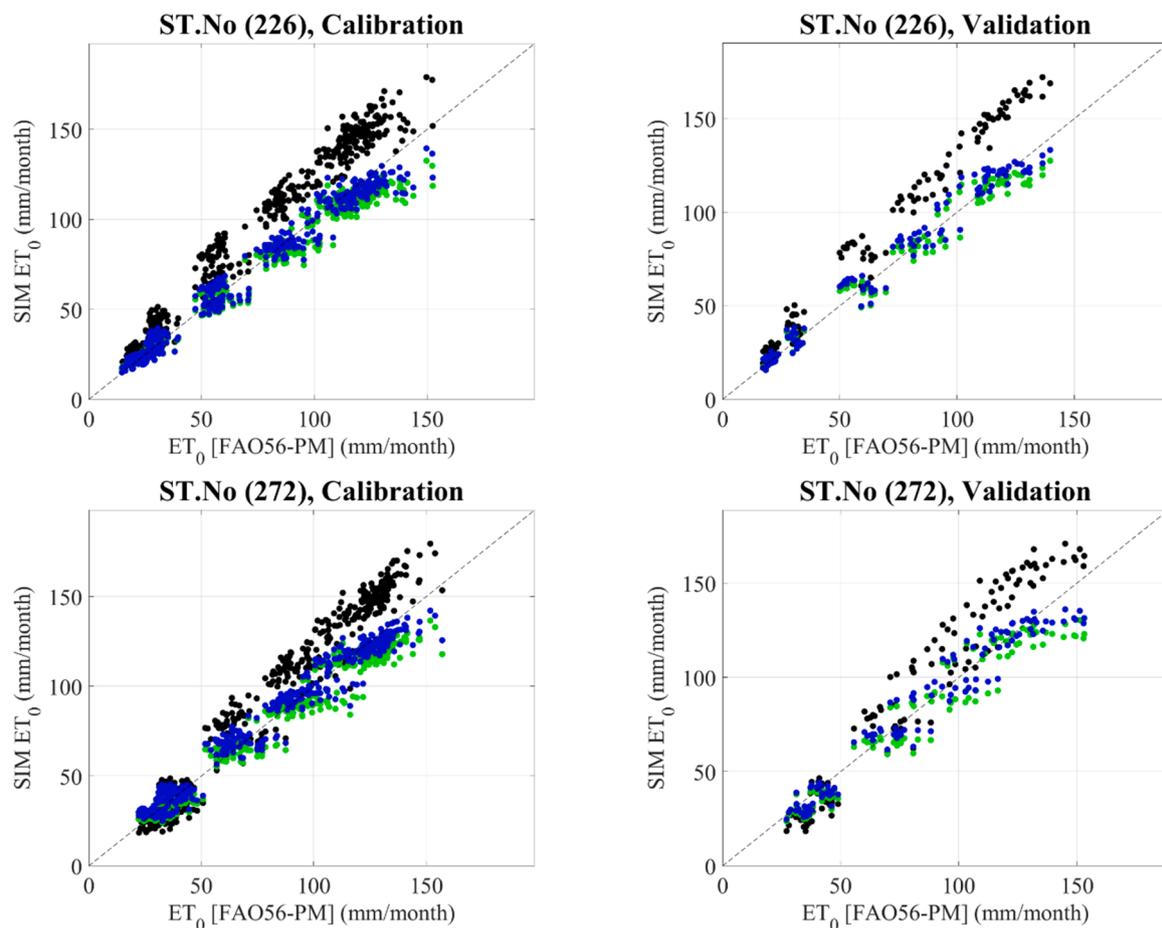
Fig. 6. (*continued*).

and validation (2011–2020) periods are evaluated in the Han River basin, as summarized in Fig. 9.

The correlation coefficient (CC) for the monthly evapotranspiration ranges from 0.967 to 0.992 (from 0.976 to 0.994) for the traditional HS equation $ET_{o\ HS(traditional)}$ (black triangles), from 0.973 to 0.993 (from 0.974 to 0.994) for the single-scale model $ET_{o\ single-scale}$ (green squares), and from 0.970 to 0.993 (from 0.972 to 0.994) for the multi-scale model $ET_{o\ multi-scale}$ (blue circles) describing the calibration period (or the validation period), respectively. As shown in Fig. 9, the range of the correlation coefficient is relatively narrow, and noticeable changes were not observed, although the traditional HS equation showed slightly better performance. It should be noted that the correlation coefficient is a statistical measure for evaluating a linear relationship, so one cannot expect an improvement within the recalibration (or bias correction) process (Kim et al., 2018b). On the other hand, both single-scale and multi-scale models demonstrate better performance in terms of the RMSE, representing a significant reduction in the RMSE. More specifically, the RMSE for the monthly evapotranspiration ranges from 9.01 to 25.65 mm/month (from 8.19 to 24.80) for the traditional HS equation $ET_{o\ HS(traditional)}$ (black triangles), from 6.13 to 9.29 mm/month (from 7.28 to 12.51) for the single-scale model $ET_{o\ single-scale}$ (green squares) and from 5.39 to 7.98 mm/month (from 5.50 to 9.32) for the multi-scale model $ET_{o\ multi-scale}$ (blue circles) describing the calibration period (or the validation period), respectively. There is a significant increase in the degree of reduction in the RMSE to between 11% and 69%. In addition, the index of agreement (IoA) was evaluated, and a similar pattern observed in the RMSE was identified. Finally, the percentage bias for models (i.e., HS model, the single-scale and multi-scale model) can effectively represent under- and over-estimation of the simulation results. The results illustrated that $ET_{o\ HS(traditional)}$ (black triangles) was overestimated and $ET_{o\ single-scale}$ (green squares) was slightly underestimated. The $ET_{o\ multi-scale}$ (blue circles) was placed almost close to zero for the calibration and validation period. Further, the model performance for the annual evapotranspiration was also explored. The correlation coefficient for the annual evapotranspiration is the range of 0.242–0.808 (0.408–0.970) for the traditional HS equation $ET_{o\ HS(traditional)}$ (black triangles), 0.134–0.805 (0.402–0.958) for the single-scale model $ET_{o\ single-scale}$ (green squares), and 0.217–0.796 (0.436–0.957) for the multi-scale model $ET_{o\ multi-scale}$ (blue circles) describing the calibration period (or the validation period), respectively. As already demonstrated in the monthly evapotranspiration, the correlation coefficient over the single- and multi-scale models was not significantly different for the annual scale. However, the RMSE for the annual evapotranspiration is in the range of 52.80–268.36 mm/year (53.38–257.76) for the traditional HS equation (black triangles), 44.86–66.01 mm/year (22.35–117.31) for the single-scale model (green squares), and 28.05–50.34 mm/year (21.51–78.40) for the multi-scale model (blue circles) describing the calibration period (or the validation period), respectively. More importantly, there was a noticeable reduction in the RMSE that led to a decrease of about 15–81% for the calibration and validation periods. Additionally, the variability of RMSE results for models is presented by the boxplot in supplemental material (Fig. S2). As seen in Figure, the model performance of the multi-scale model is outperformed in all timescales. All the GoF results over all the stations for different time scales can be found in supplementary material (Table S3).
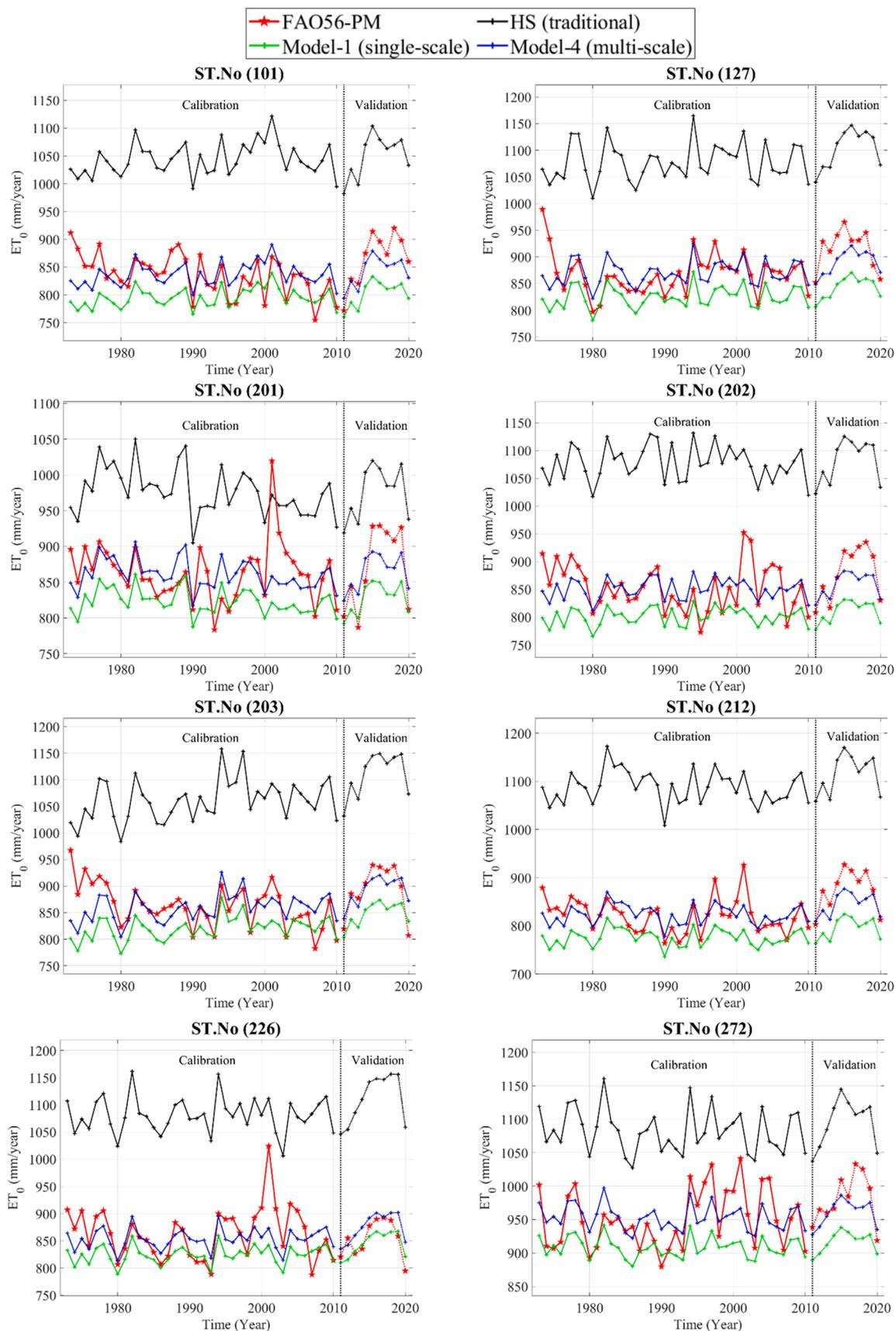
Fig. 7. Comparison of annual ET_o values over the calibration (1973–2010) and validation (2011–2020) period from different models for 8 stations (i.e., ST. Nos. 101, 127, 201, 202, 203, 212, 226 and 272). The annual ET_o values from FAO56-PM (red-solid line), traditional HS equation (solid black line), Model-1 (single-scale model, solid green line), and Model-4 (multi-scale model, solid blue line) are compared. The black dotted vertical line represents the reference year for dividing the calibration and validation period.
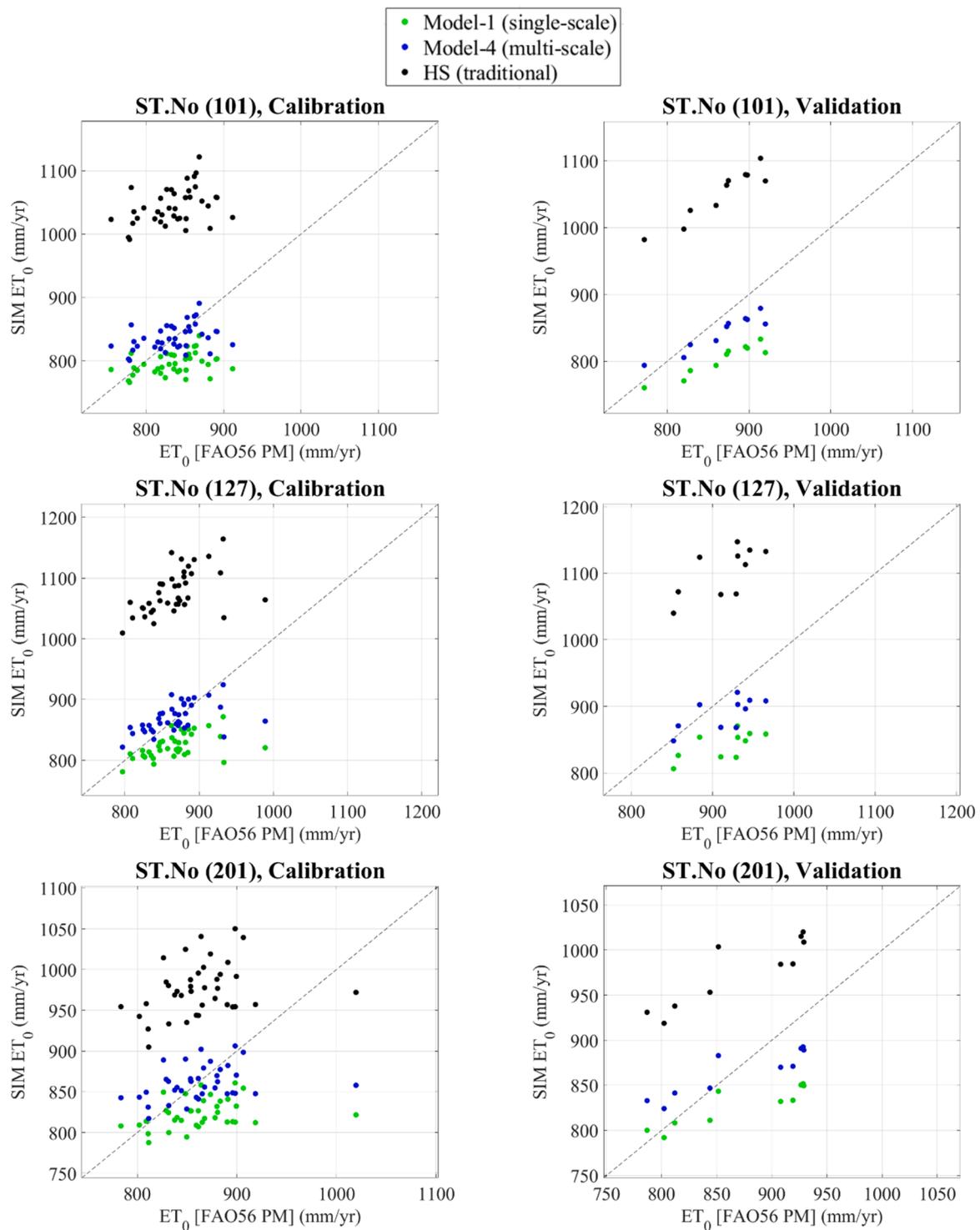
Fig. 8. Scatter plots for annual ET_o results over the calibration (1973–2010) and validation (2011–2020) period from different models for 8 stations (i.e., ST. Nos. 101, 127, 201, 202, 203, 212, 226 and 272). The relationship between annual $ET_{o\ PM}$ and the simulated ET_o results obtained from the traditional HS equation (black-filled circles), Model-1 (single-scale model, the green-filled circles) and Model-4 (multi-scale model, blue-filled circles) are represented.

5. Summary and conclusions

Existing studies have focused on recalibrating the parameters of the HS equation locally to improve the accuracy of the estimation of reference evapotranspiration using daily temperature data (Feng et al., 2017; Haslinger and Bartsch, 2016; Mehdizadeh et al., 2017; Mendicino and Senatore, 2013; Vanderlinden et al., 2004). The HS equation provides a

systematic framework for the estimation of reference evapotranspiration in the simplest, most parsimonious manner as a surrogate model. However, there is likely to be a limitation in representing the complex mechanism of the evapotranspiration process using the HS equation, which solely uses temperature data as an independent variable. Under these circumstances, this study proposed a novel, multi-scale context (i. e., daily, monthly, and annual scales) approach to the HS equation for a
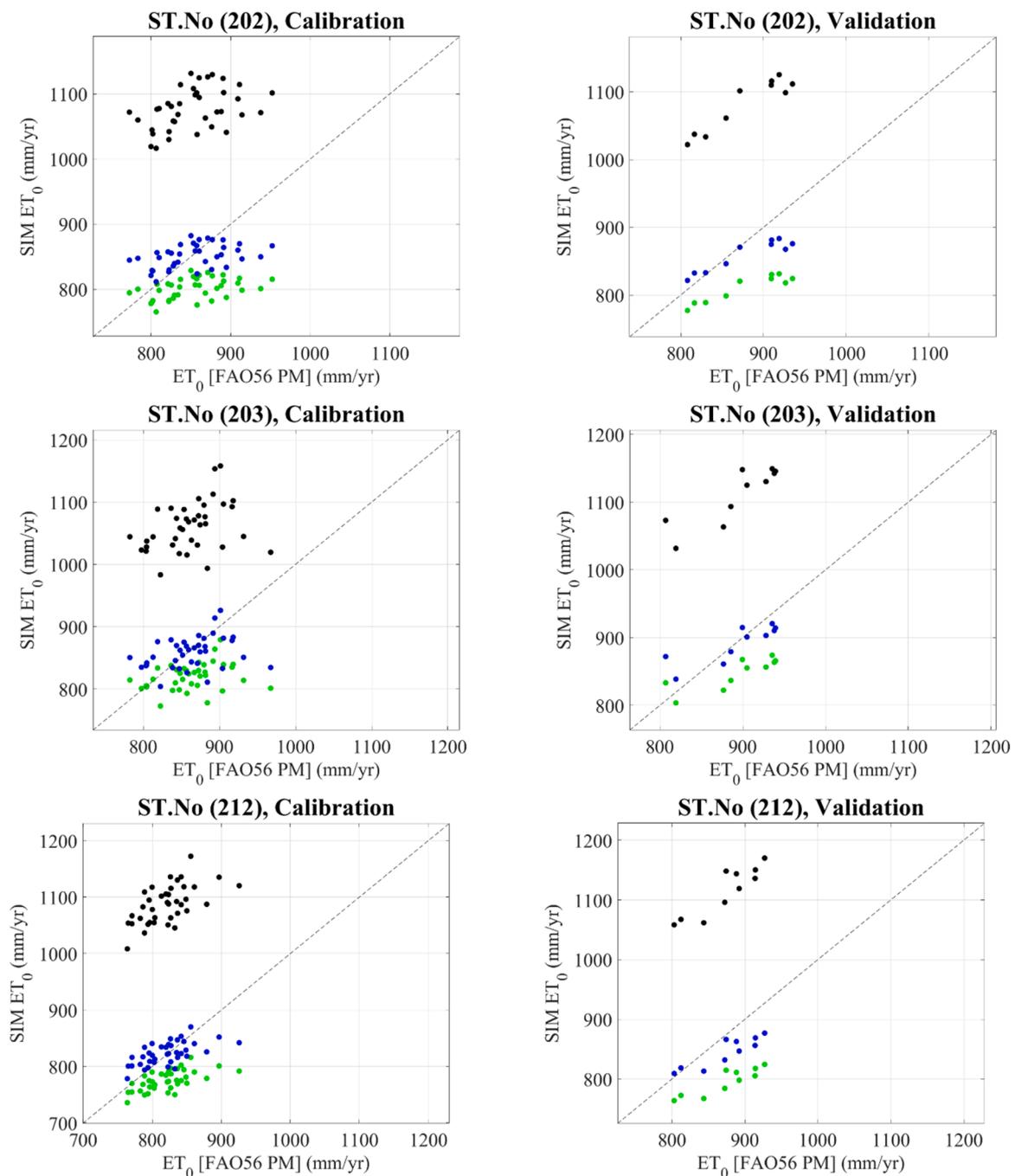
Fig. 8. (*continued*).

better fit of the reference evapotranspiration obtained from the FAO56-PM method. The proposed models were evaluated using meteorological data for the Han-River watershed in South Korea. This study explored four types of surrogate models in terms of estimating model parameters for improving the existing HS equation for reference evapotranspiration. A Bayesian modeling framework for the inference of model parameters was further proposed. Among the models, the model performances were compared through the deviance information criterion (DIC) with the single-scale model (Model-1) and the traditional HS equation that is based on fixed parameters. Finally, the models were evaluated by GoF measures during the calibration period (1973–2010) and validation period (2011–2020) in the cross-validation framework. The key findings from this study are as follows:

(1) The HS equation with the fixed parameters without recalibration clearly leads to an overestimation of the reference ET for all stations. On one hand, the locally recalibrated approach to the HS equation at a daily temporal scale (i.e., single-scale model, Model-1) can effectively reduce the systematic bias associated with the use of the traditional HS equation with fixed parameters reflecting the geographical impact on the estimation of evapotranspiration in a specific region of interest. The single-scale model can reproduce the reference evapotranspiration fairly well during the dry season while failing to provide an accurate estimation, particularly during the wet season. On the other hand, the recalibrated HS model using only daily evapotranspiration during the parameter estimation process is insufficient to reproduce the underlying distribution of evapotranspiration at
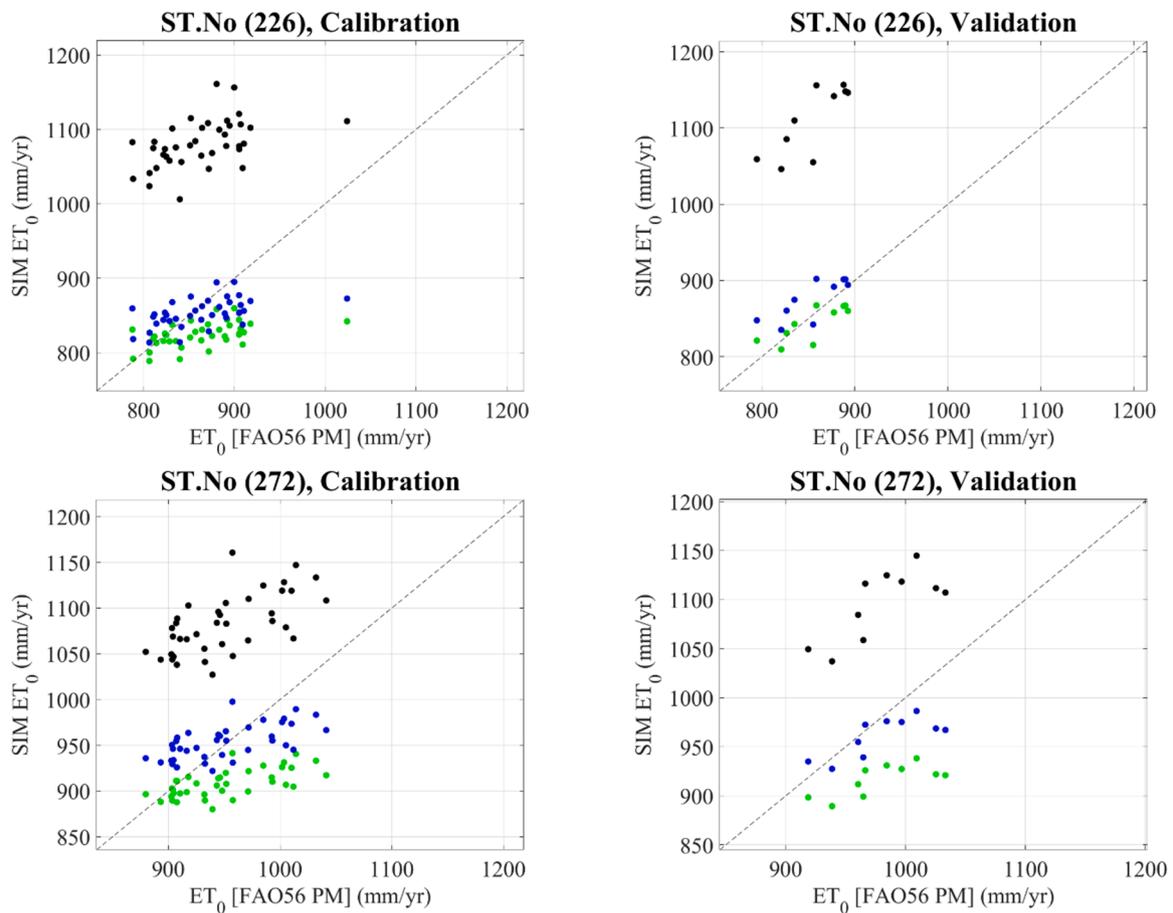
Fig. 8. (*continued*).

different temporal scales (e.g., monthly and annual), which may lead to a large systematic bias in the rainfall-runoff modeling or assessment of climate change impact on water resources.

(2) This study developed a novel approach to the HS equation to better reproduce the reference evapotranspiration across different temporal scales by simultaneously maximizing the likelihood function at multiple temporal scales within a Bayesian modeling framework. More specifically, the multi-scale model for the HS equation exploits multiple likelihood functions defined at three different temporal scales (i.e., daily, monthly and annual) in a model by sequentially summing the monthly and annual evapotranspiration from the estimated daily evapotranspiration. The regression parameters (α, β, γ) of the multi-scale model for the Model-4, including daily, monthly and annual evapotranspiration series, were in the ranges of 0.0015–0.0028, 19.3–36, 0.25–0.49, respectively, over all the stations. More importantly, the estimated parameters β and γ showed somewhat different parameter space from that of the traditional HS equation, while the parameter α was similar to that of the traditional HS equation.

(3) The multi-scale model offered an accurate approximation to the reference evapotranspiration compared to the single-scale model at a daily timescale. Overall, the multi-scale model demonstrated a more precise estimation of the reference evapotranspiration at the aggregated temporal scales (i.e., monthly and annual) compared to those of the single-scale. The enhanced efficacy of the recalibration process at multiple temporal scales can be more easily explained and justified through GoF measures such as the RMSE and IoA. To be more specific, the multi-scale models showed a significant increase in the degree of reduction in the RMSE achieved by up to about 60% and 70% compared to that of

the traditional HS equation for the monthly and annual evapotranspiration estimation, respectively. Moreover, a similar pattern observed in the RMSE was also identified for the index of agreement (IoA) measure.

(4) The main limitation of this study is that the HS equation largely relies on the standard PM model, which requires a set of meteorological data to estimate ET_o that are not readily available in many countries. Data requirements for the standard PM model can hinder the use of the HS equation in a practical application. Meanwhile, the HS equation is still effective as a surrogate model to the standard PM model due to the advantages of accuracy, consistency, and parsimoniousness. Thus, a regionalization approach to the HS parameters can be alternatively introduced by considering the local climate conditions and topographical characteristics.

To the best of our knowledge, this study is the first attempt to obtain a set of parameters of the HS equation for estimating evapotranspiration by simultaneously considering evapotranspiration at multiple time scales. The proposed model still shows limitations in reproducing the inherent characteristics of evapotranspiration during the summer season. This may be due to different responses to temperature in estimating evapotranspiration in the HS equation. This issue was not fully explored in the current study, and time-varying parameters on the basis of the seasonality of evapotranspiration will be investigated in future work. More formally, future work will focus on developing a regional surrogate model within a hierarchical Bayesian modeling framework to better characterize the spatio-temporal patterns of evapotranspiration.

(a) Monthly Timescale
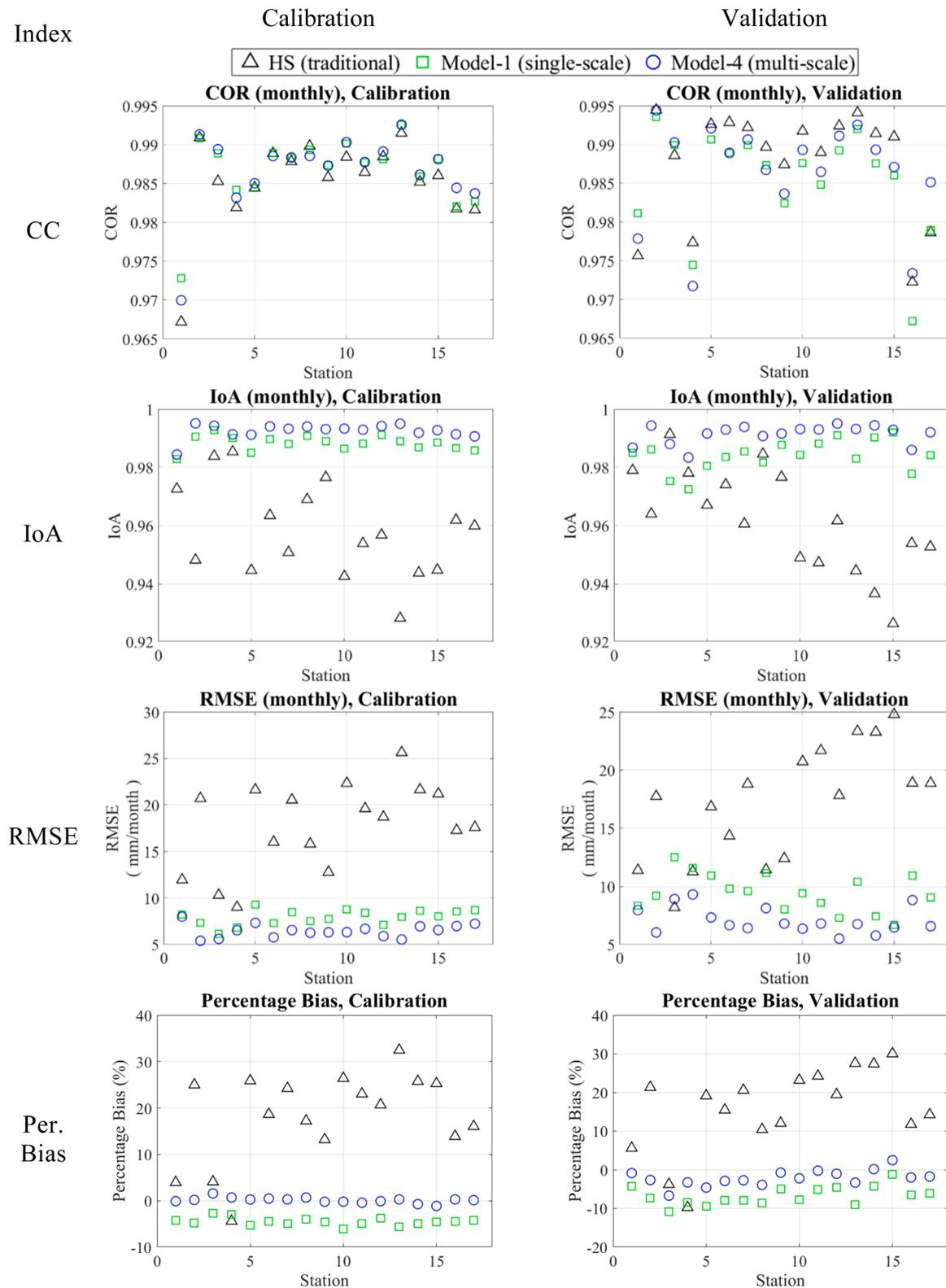
Fig. 9. Comparison of statistical GoF measures (i.e., CC, IoA, RMSE, Percentage Bias) for (a) the monthly, (b) annual scales over the calibration (1973–2010) and validation (2011–2020) period from different models for all stations. The GoF measures between $ET_{o\ PM}$ and the simulated ET_o results obtained from the traditional HS equation (black-filled circles), Model-1 (single-scale model, the green-filled circles) and Model-4 (multi-scale model, blue-filled circles) are represented. Here, a percentage bias is given only on the monthly scale because the value of the annual scale is the same as the monthly scale.

(b) Annual Timescale
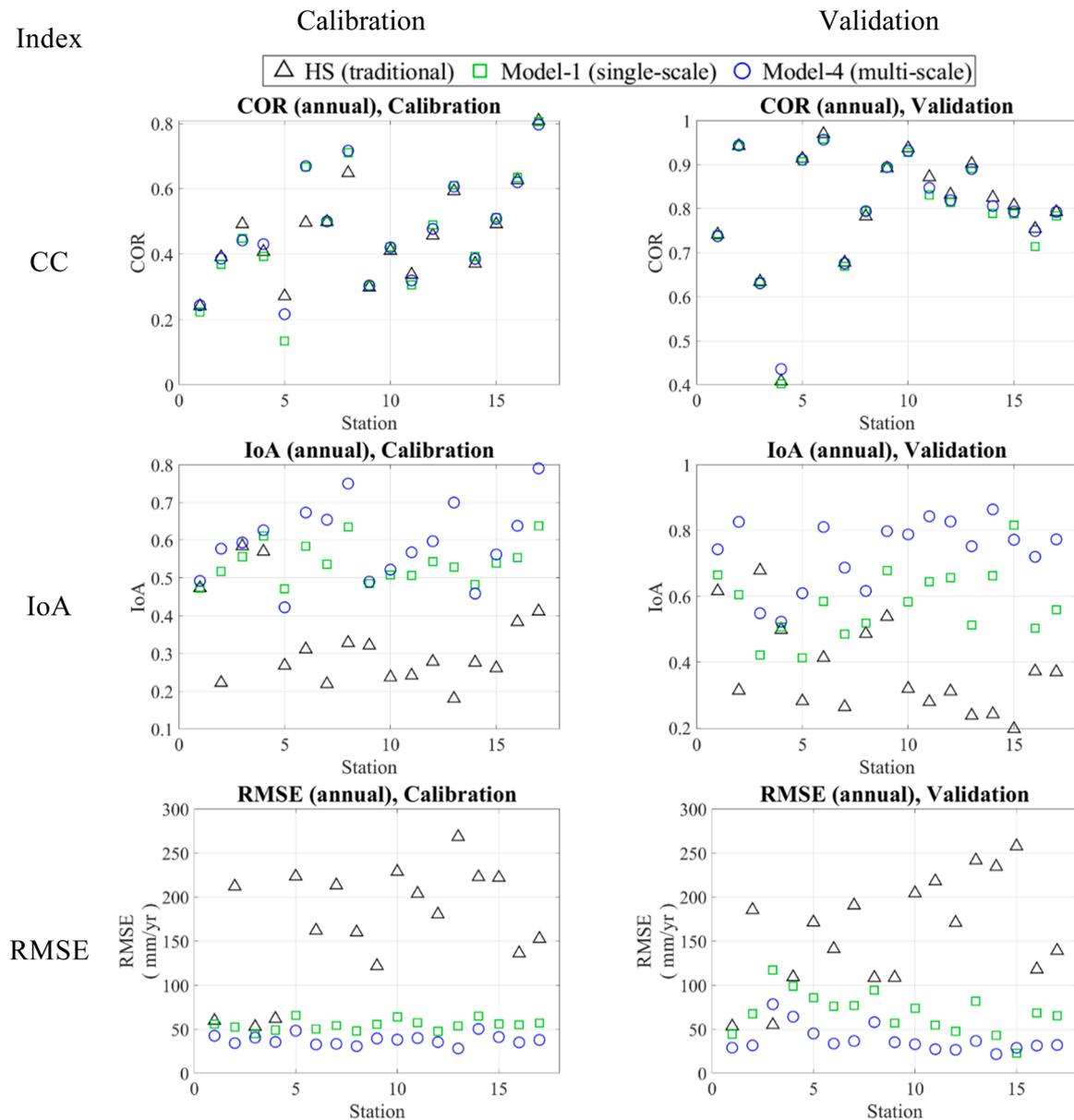
Fig. 9. (*continued*).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.agwat.2022.108038.

References

Allen, R.G., Pereira, L.S., Raes, D., 1998. Crop Evapotranspiration-guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage Paper, Rome, p. 56.

Allen, R.G., Walter, I.A., Elliott, R., Howell, T.A., Itenfisu, D., Jensen, M.E., 2005. The ASCE standardized reference evapotranspiration equation.

Almorox, J., Quej, V.H., Martí, P., 2015. Global performance ranking of temperature-based approaches for evapotranspiration estimation considering Köppen climate classes. J. Hydrol. (Amst.) 528, 514–522. https://doi.org/10.1016/j.jhydrol.2015.06.057.

Falamarzi, Y., Palizdan, N., Huang, Y.F., Lee, T.S., 2014. Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). Agric. Water Manag. 140, 26–36. https://doi.org/10.1016/j.agwat.2014.03.014.

Feng, Y., Cui, N., Zhao, L., Hu, X., Gong, D., 2016. Comparison of ELM, GANN, WNN and empirical models for estimating reference evapotranspiration in humid region of Southwest China. J. Hydrol. (Amst.) 536, 376–383. https://doi.org/10.1016/j.jhydrol.2016.02.053.

Feng, Y., Jia, Y., Cui, N., Zhao, L., Li, C., Gong, D., 2017. Calibration of Hargreaves model for reference evapotranspiration estimation in Sichuan basin of Southwest China. Agric. Water Manag 181, 1–9. https://doi.org/10.1016/j.agwat.2016.11.010.

Gavilán, P., Lorite, I.J., Tornero, S., Berengena, J., 2006. Regional calibration of Hargreaves equation for estimating reference et in a semiarid environment. Agric. Water Manag 81, 257–281. https://doi.org/10.1016/j.agwat.2005.05.001.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis. Chapman & Hall/CRC Inc., New York.

Gilkst, W.R., Best, N.G., Tan, K.K.C., 1995. Adaptive rejection metropolis sampling within gibbs sampling. Appl. Stat.

Hargreaves, G.H., 1994. Defining and using reference evapotranspiration. J. Irrig. Drain. Eng. 120, 1132–1139.

Hargreaves, G.H., Allen, R.G., 2003. History and evaluation of hargreaves evapotranspiration equation. J. Irrig. Drain. Eng. 129, 53–63. https://doi.org/10.1061/ASCE0733-94372003129:153.

Hargreaves, G.H., Samani, Z.A., 1982. Estimating potential evapotranspiration. J. Irrig. Drain. Div. 108, 225–230. https://doi.org/10.1061/JRCEA4.0001390.

Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1, 96–99.

Haslinger, K., Bartsch, A., 2016. Creating long-term gridded fields of reference evapotranspiration in Alpine terrain based on a recalibrated Hargreaves method. Hydrol. Earth Syst. Sci. 20, 1211–1223. https://doi.org/10.5194/hess-20-1211-2016.

Kang, Y., Chen, P., Cheng, X., Zhang, S., Song, S., 2022. Novel hybrid machine learning framework with decomposition–transformation and identification of key modes for estimating reference evapotranspiration. Agric. Water Manag. 273. https://doi.org/10.1016/j.agwat.2022.107882.

Kim, K.B., Kwon, H.H., Han, D., 2015. Bias correction methods for regional climate model simulations considering the distributional parametric uncertainty underlying the observations. J. Hydrol. (Amst.) 530, 568–579. https://doi.org/10.1016/j.jhydrol.2015.10.015.

Kim, K.B., Kwon, H.H., Han, D., 2018. Exploration of warm-up period in conceptual hydrological modelling. J. Hydrol. (Amst.) 556, 194–210. https://doi.org/10.1016/j.jhydrol.2017.11.015.

Kim, T.J., Kwon, H.H., Lima, C., 2018. A Bayesian partial pooling approach to mean field bias correction of weather radar rainfall estimates: application to Osungsan weather radar in South Korea. J. Hydrol. (Amst.) 565, 14–26. https://doi.org/10.1016/j.jhydrol.2018.07.082.

Kwon, H.H., Brown, C., Lall, U., 2008. Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling. Geophys Res Lett. 35. https://doi.org/10.1029/2007GL032220.

Kwon, H.H., de Souza Filho, de Assis, Block, F., Sun, P., Lall, L., Reis, U., S, D., 2012. Uncertainty assessment of hydrologic and climate forecast models in Northeastern Brazil. Hydrol. Process 26, 3875–3885. https://doi.org/10.1002/hyp.8433.

L'vovich, M.I., White, G.F., Belyaev, A. v, Kindler, J., Koronkevic, N.I., Lee, T.R., Voropaev, G. v, 1990. Use and Transformation of Terrestrial Water Systems. Cambridge University Press.

Lee, K.H., Park, J.H., 2008. Calibration of the Hargreaves equation for the reference evapotranspiration estimation on a nation-wide scale. KSCE J. Civ. Environ. Eng. Res. 28, 675–681.

Lhomme, J.-P., 1997. Towards a rational definition of potential evaporation. Hydrol. Earth Syst. Sci. 1, 257–264. https://doi.org/10.5194/hess-1-257-1997.

Lima, C.H.R., Kwon, H.H., Kim, Y.T., 2018. A local-regional scaling-invariant Bayesian GEV model for estimating rainfall IDF curves in a future climate. J. Hydrol. (Amst.) 566, 73–88. https://doi.org/10.1016/j.jhydrol.2018.08.075.

Martínez-Cob, A., Tejero-Juste, M., 2004. A wind-based qualitative calibration of the Hargreaves ET0 estimation equation in semiarid regions. Agric. Water Manag 64, 251–264. https://doi.org/10.1016/S0378-3774(03)00199-9.

McColl, K.A., Rigden, A.J., 2020. Emergent simplicity of continental evapotranspiration. Geophys Res Lett. 47. https://doi.org/10.1029/2020GL087101.

Mehdizadeh, S., Saadatnejadgharahassanlou, H., Behmanesh, J., 2017. Calibration of Hargreaves–Samani and Priestley–Taylor equations in estimating reference evapotranspiration in the Northwest of Iran. Arch. Agron. Soil Sci. 63, 942–955. https://doi.org/10.1080/03650340.2016.1249474.

Mendicino, G., Senatore, A., 2013. Regionalization of the hargreaves coefficient for the assessment of distributed reference evapotranspiration in Southern Italy. J. Irrig. Drain. Eng. 139, 349–362. https://doi.org/10.1061/(asce)ir.1943-4774.0000547.

Moon, J.W., Jung, C.G., Lee, D.R., 2013. Parameter regionalization of hargreaves equation based on climatological characteristics in Korea. J. Korea Water Resour. Assoc. 46, 933–946. https://doi.org/10.3741/jkwra.2013.46.9.933.

Novák, V., 2012. Evapotranspiration: a component of the water cycle. In: Evapotranspiration in the Soil-Plant-Atmosphere System. Springer, Netherlands, pp. 1–13. https://doi.org/10.1007/978-94-007-3840-9_1.

Razıei, T., Pereira, L.S., 2013. Estimation of ETo with Hargreaves-Samani and FAO-PM temperature methods for a wide range of climates in Iran. Agric. Water Manag 121, 1–18. https://doi.org/10.1016/j.agwat.2012.12.019.

Shirmohammadi-Aliakbarkhani, Z., Saberali, S.F., 2020. Evaluating of eight evapotranspiration estimation methods in arid regions of Iran. Agric. Water Manag 239. https://doi.org/10.1016/j.agwat.2020.106243.

So, B.J., Kim, J.Y., Kwon, H.H., Lima, C.H.R., 2017. Stochastic extreme downscaling model for an assessment of changes in rainfall intensity-duration-frequency curves over South Korea using multiple regional climate models. J. Hydrol. (Amst.) 553, 321–337. https://doi.org/10.1016/j.jhydrol.2017.07.061.

Su, Q., Singh, V.P., Karthikeyan, R., 2022. Improved reference evapotranspiration methods for regional irrigation water demand estimation. Agric. Water Manag. 274, 107979 https://doi.org/10.1016/j.agwat.2022.107979.

Todorovic, M., Karic, B., Pereira, L.S., 2013. Reference evapotranspiration estimate with limited weather data across a range of Mediterranean climates. J. Hydrol. (Amst.) 481, 166–176. https://doi.org/10.1016/j.jhydrol.2012.12.034.

Traore, S., Luo, Y., Fipps, G., 2016. Deployment of artificial neural network for short-term forecasting of evapotranspiration using public weather forecast restricted messages. Agric. Water Manag 163, 363–379. https://doi.org/10.1016/j.agwat.2015.10.009.

Vanderlinden, K., Giraldez, J., van Meirvenne, M., 2004. Assessing reference evapotranspiration by the Hargreaves method in Southern Spain. J. Irrig. Drain. Eng. 130, 184–191. https://doi.org/10.1061/ASCE0733-94372004130:3184.

Xiang, K., Li, Y., Horton, R., Feng, H., 2020. Similarity and difference of potential evapotranspiration and reference crop evapotranspiration – a review. Agric. Water Manag. https://doi.org/10.1016/j.agwat.2020.106043.

Yang, Y., Chen, R., Han, C., Liu, Z., 2021. Evaluation of 18 models for calculating potential evapotranspiration in different climatic zones of China. Agric. Water Manag 244. https://doi.org/10.1016/j.agwat.2020.106545.

Zhao, L., Xia, J., Xu, C. yu, Wang, Z., Sobkowiak, L., Long, C., 2013. Evapotranspiration estimation methods in hydrological models. J. Geogr. Sci. 23, 359–369. https://doi.org/10.1007/s11442-013-1015-9.

Zotarelli, L., Dukes, M.D., Romero, C.C., Migliaccio, K.W., Morgan, K.T., 2010. Step by Step Calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method). In: Institute of food and agricultural sciences, 1. University of Florida.