TU Delft

Delft University of Technology

Automated actual evapotranspiration estimation

Hybrid model of a novel attention based U-Net and metaheuristic optimization algorithms

Bafti, Alireza Ghaderi; Ahmadi, Arman; Abbasi, Ali; Kamangir, Hamid; Jamali, Sadegh; Hashemi, Hossein

DOI
10.1016/j.atmosres.2023.107107

Publication date
2023

Document Version
Final published version

Published in
Atmospheric Research

Citation (APA)
Bafti, A. G., Ahmadi, A., Abbasi, A., Kamangir, H., Jamali, S., & Hashemi, H. (2023). Automated actual evapotranspiration estimation: Hybrid model of a novel attention based U-Net and metaheuristic optimization algorithms. *Atmospheric Research*, *297*, Article 107107. https://doi.org/10.1016/j.atmosres.2023.107107

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Atmospheric Research 297 (2024) 107107



Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres



Automated actual evapotranspiration estimation: Hybrid model of a novel attention based U-Net and metaheuristic optimization algorithms

Alireza Ghaderi Bafti [a,*,1], Arman Ahmadi [b], Ali Abbasi [a,c], Hamid Kamangir [d], Sadegh Jamali [e], Hossein Hashemi [f]

[a] Department of Civil Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
[b] Department of Biological and Agricultural Engineering, University of California Davis, Davis, CA 95616, USA
[c] Faculty of Civil Engineering and Geosciences, Water Resources Section, Delft University of Technology, Delft, the Netherlands
[d] Department of Agricultural and Environmental Engineering, University of California Davis, Davis, CA 95616, USA
[e] Department of Technology and Society, Faculty of Engineering, Lund University, 221 00 Lund, Sweden
[f] Department of Water Resources Engineering & Center for Advanced Middle Eastern Studies, Lund University, Lund 221 00, Sweden

ARTICLE INFO

Keywords:
Actual evapotranspiration mapping
Deep learning
Attention mechanism
Harris Hawks Optimization
Simplified surface energy balance (SSEBop)
U-Net architecture

ABSTRACT

Actual evapotranspiration (ET_a) plays a crucial role in the water and energy cycles of the earth. An accurate estimate of the ET_a is essential for management of the water resources, agriculture, and irrigation, as well as research on atmospheric variations. Despite the importance of accurate ET_a values, estimating and mapping them remains challenging due to the physical and biological complexity of the ET process. As a novel approach for rapid and reliable estimation of ET_a, the present study develops automated deep learning (AutoDL) models that incorporate a metaheuristic optimization algorithm for image processing, architectural design, and hyperparameter tuning. The proposed AutoDL models integrate three different spatial and channel attention mechanisms, including a novel activated spatial attention mechanism (ASPAM), with the U-Net architecture. Bypassing the need for meteorological inputs, the proposed framework uses Moderate Resolution Imaging Spectrometer (MODIS) products and Digital Elevation Model (DEM) data as inputs. To evaluate the performance of the models, they are applied to three study areas in the United States with various climatic characteristics. According to the results, during the spring and summer, when the target values have higher certainty, the estimations are highly promising, with R^2 as high as 0.91 and MAPE as low as 6.40%. Furthermore, the proposed ASPAM module improves the accuracy of ET_a estimations compared to attention gate (AG) and squeeze and excitation (SE) attention modules. The results also indicate that the MODIS raw products and derived vegetation and water indices can predict ET_a within a reliable range of accuracy, with the addition of DEM data marginally enhancing the models' performance. The automatic workflow of this model makes it significantly easy to use, contributing to its applicability and generalizability for enhancing atmospheric research.

1. Introduction

Actual evapotranspiration (ET_a) is a fundamental component of the hydrological cycle (Su et al., 2022). Approximately 60 to 90% of the total precipitation returns to the atmosphere through ET_a (Dayal et al., 2021; Fisher et al., 2017). Moreover, ET_a is the primary consumer of incoming solar energy and plays a vital role in the Earth's energy cycle (McCulloch, 1976; Mu et al., 2011). Therefore, understanding the distribution of ET_a on a large scale is essential for various practices such as management of agriculture, and water resources, and assessment of atmospheric variables including relative humidity, net radiation, and wind speed (Wang et al., 2021; Ahmadi et al., 2022). Despite its importance, the accurate and reliable measurement of ET_a remains challenging (Mementi and Choudhury, 2007), primarily due to the complex interplay of factors including radiation, atmospheric evaporative demand, soil moisture, and its spatial variation at different scales (Amatya et al., 2016; Gowda et al., 2008).

In general, the methods for estimating ET_a can be classified into two main groups: direct measurement and modeling (Chen et al., 2016). Direct measurements, including the observations from weighing

* Corresponding author.
E-mail address: a.ghaderibafti@ufl.edu (A.G. Bafti).
[1] Present address: Department of Soil, Water, and Ecosystem Sciences, University of Florida, Gainesville, FL 32611, USA.

https://doi.org/10.1016/j.atmosres.2023.107107
Received 18 July 2023; Received in revised form 21 October 2023; Accepted 8 November 2023
Available online 13 November 2023
0169-8095/© 2023 Elsevier B.V. All rights reserved.

lysimeters, large aperture scintillometers (LAS), and eddy covariance (EC) systems (Li et al., 2017), are considered the most reliable methods for estimating ET_a (Abtew and Melesse, 2013; Tanny et al., 2008). However, these approaches are costly and labor-intensive, require expert personnel, and usually offer data for limited durations. Furthermore, scaling up observational ET_a data is a complex and multidimensional process, posing challenges for direct methods (Gebremedhin et al., 2020), limiting their application to short-term observations, and validation of indirect approaches (Rahimpour and Rahimzadegan, 2021).

In contrast to measurement methods, ET_a modeling based on remote sensing (RS) technology offers a unique advantage by providing consistent and economically viable estimations at various scales, ranging from daily to yearly and from regional to global (Zhuang et al., 2021; Tang et al., 2010; Rahimpour and Rahimzadegan, 2021). In particular, the availability of numerous satellite-based sensors in recent years has significantly contributed to the development of new RS-based approaches (Doulabian et al., 2021; Sahraei et al., 2023). Among these methods, energy balance-based methods (Bastiaanssen et al., 1998a; Su, 2002) are not only more widely used (Wang et al., 2014) but are also recognized for their superior performance compared to other methods (Assouline and Mahrer, 1993; Lenters et al., 2005; Winter et al., 2003).

Among the mentioned models, the recently developed operational simplified surface energy balance (SSEBop) model (Senay et al., 2013) offers distinct benefits in terms of simplicity and operational applications for large-scale ET_a estimation (Senay, 2018). The innovative aspect of the SSEBop model is that, under clear-sky conditions, it uses a predefined boundary where each pixel has a unique differential temperature (dT), eliminating the need for manual selection of hot and cold pixels (Chen et al., 2016). Despite its simplifications, the SSEBop-ET_a product has demonstrated satisfactory performance and maintained comparable accuracy level to models such as SEBAL and METRIC across diverse landscapes (Chen et al., 2016; Weerasinghe et al., 2020; Ji et al., 2019; Chen et al., 2016; Senay, 2018). However, similar to other energy-based models, the SSEBop model has certain drawbacks. It all require a resource-intensive integration of various meteorological datasets, lacks the ability to incorporate spatial dependencies among variables, and demands expertise for implementation, further hindering their widespread adoption and convenience of use. These limitations emphasize the necessity for an expedient method for large-sale ET_a estimation based on an integrated dataset.

In recent years, as a result of the promising ability of deep learning (DL) to address complex regression problems (Rabiei et al., 2021; Yang et al., 2020), some studies have applied DL variants for ET estimation (Babaeian et al., 2022). For instance, Yan et al. (2023) utilized an energy balance method as the benchmark to train a long short-term memory on weather data from different meteorological stations for estimating reference ET. Although they achieved promising results, their model still relies on data from multiple sources and is limited to local applications. To overcome the challenges of large-scale applicability, a few studies have used global-scale data from various sources as inputs (Wang et al., 2023) and ET_a derived from an energy-based model as the target variable. However, these approaches still require data from multiple sources, demand expertise for implementation, and encounter difficulties incorporating spatial dependencies among variables.

Among the various classes of DL, the fully connected neural network (FCNN) (Long et al., 2015) offers several advantages over traditional patch-wised convolution neural networks (CNNs), including improved spatial preservation, reduced parameters, and the ability to capture multi-scale features (Xu et al., 2021). Specifically, the U-Net architecture, a specific variant of FCNN models characterized by an encoder-decoder structure (Minaee et al., 2021), demonstrates superior computational efficiency and enhanced ability to capture spatial correlations, resulting in improved accuracy compared to CNNs (Maggiori et al., 2017; Stoian et al., 2019). While the U-Net has been extensively applied for image classification and segmentation tasks, its potential for large-scale ET_a estimation in regions with diverse climatologies remains unexplored (Taccari et al., 2022).

In practice, spatial dependencies are commonly observed among atmospheric variables. As a result, incorporating an attention mechanism (Ding et al., 2019) that focuses on the most informative spatial points and disregards the less relevant ones may further holds the potential to enhance estimation from spatial feature sequences (Chen et al., 2018; Ding et al., 2019). Nonetheless, studies have yet to examine its potential in the context of hydrological forecasting and regression (Chen et al., 2020; Ding et al., 2019). We hypothesize that, inspired by the self-activated internal attention (SAINA) module (Alizadeh et al., 2021), a novel attention mechanism that uses matrix multiplication instead of more commonly used convolution blocks or Harmon multiplication in attention mechanism operations can better capture spatial dependencies by involving a greater number of arrays in matrix computations. However, to our knowledge, no such attention module has been developed or evaluated in the context of U-Net-based models.

Automated DL (AutoDL), an emerging field within the realm of DL, offers simplified DL model development and interpretation through automating workflow processes (Prasad et al., 2022; Li et al., 2022a, 2022b). This automation facilitates its implementation and expands the usage of DL models among a diverse set of researchers (Polonskaia et al., 2021). Consequently, in this study, we design a novel approach for ET_a estimation based on AutoDL models that eliminates the need for manual intervention throughout the entire process (Li et al., 2022a, 2022b). A crucial step in this approach involves the utilization of an optimization algorithm to determine the optimal architecture and hyperparameters of the DL model (Ma et al., 2020). The present study employs a meta-heuristic algorithm known as Harris Hawks Optimization (HHO) (Heidari et al., 2019), which has demonstrated high efficiency in identifying optimal values (Setiawan et al., 2021).

According to the above, the main objective of this research is to develop an AutoDL attention-based U-Net model optimized by HHO for the integrated, efficient, and reliable estimation of daily ET_a on a large scale as an alternative to traditional energy-based models. In this regard, we propose a novel Activated Spatial Attention Mechanism-U-Net (ASPAM-U-Net) and compare its ET_a estimation performance to that of AG-U-Net, SE-U-Net, and a Simple-U-Net. Since direct measurements such as Eddy covariance are sparsely distributed and cannot fully capture the spatial features of ET_a at large scales, similar to previous studies (Li et al., 2022a, 2022b; Yan et al., 2023), we use an energy-based model (SSEBop) as the target variable of our developed models. By eliminating the need for meteorological variables, we rely solely on the optical and thermal observations of the MODIS satellite to address the problem of reliance on multiple data sources. We use several deterministic evaluation metrics to evaluate the performance of these models across different regions and months of the year to assess their performance under different types of climatology, weather, and surface conditions.

This assessment also addresses the following research questions: To what extent can an optimization algorithm enhance model performance? How effectively do spatial and channel attention mechanisms improve deterministic ET_a prediction? How are the forecast errors related to the hydro-climatic conditions in the study areas? And finally, which types of inputs have the most significant impact on the model's performance?

The structure of this paper is outlined as follows: Section 2 details the materials and methods, encompassing the data collection procedures and descriptions of the study regions. Section 3 presents the results and discussion on the comparative evaluation study. Section 4 summarizes the findings and offers recommendation for future studies.

2. Materials and methods

2.1. Data collection and study regions

In this study, three groups of data were utilized for mapping to the

target dataset of SSEBop-ET_a, including (1) optical and thermal bands collected by the MODIS sensor, (2) a set of derived vegetation- and water-based indices, and (3) topographical parameters derived from a digital elevation model (DEM). To ensure a more comprehensive evaluation, the model was applied to three different regions covering all months of the year to capture the impact of temporal and spatial climate dynamics on the performance of the proposed methodology. The details of the collected data and study regions are provided in the following subsections.

2.1.1. Input variables

A suite of MODIS satellite image products was acquired from the Land Processes Distributed Active Archive Center website as the main inputs of the models. These inputs were classified into three groups, as follows:

The first group of data consisted of 12 thermal and optical bands, including Terra MODIS surface reflectance bands 1 through 7, as well as three bands providing information on azimuth angle, solar zenith angle, and view zenith angles. These bands were collected from the MOD09A1 Version 6 product. This dataset is an 8-day composite, where each pixel comprises the best observation chosen over eight days. Two additional bands in this group are thermal bands 31 and 32, collected from the MOD11A2 Version 6 product. Each pixel value in these bands in this product represents a simple average of the corresponding pixels collected within the eight days.

The second group contained a series of indices derived from the first group of data and a MODIS-based land use map incldluing: the leaf area index (LAI) and the fraction of photosynthetically active radiation (FPAR) from the MOD15A2H Version 6 product; the yearly FAO-LCCS2 land use layer (all at 500 m resolution); percent of pixels not vegetated, and percent of pixels non-tree vegetated, all three from the MOD44B Version 6 Vegetation Continuous Fields yearly product; and land surface temperature (LST) from the MOD11A2 product. Additionally, 15 vegetation- and water-based indices (Table 1) were computed using surface reflectance bands 1 through 7 to further enhance the accuracy of the DL models, where NIR denotes near infrared.

Table 1
Equations of vegetation- and water-based indices used in the second data group.

Index	Description	Formula
NDVI	Normalized Difference Vegetation Index	$\frac{(NIR - RED)}{(NIR + RED)}$
RDVI	Renormalized Difference Vegetation Index	$\frac{(NIR - RED)}{\sqrt{\lvert NIR + RED \rvert}}$
DVI	Difference Vegetation Index	$(NIR - RED)$
IPVI	Infrared Percentage Vegetation Index	$\frac{(NIR)}{(NIR + RED)}$
NLI	Nonlinear vegetation index	$\frac{(NIR^2 - RED)}{(NIR^2 + RED)}$
OSAVI	Optimized Soil Adjusted Vegetation Index	$\frac{1.5 \times NIR - RED)}{(NIR + RED + 0.16)}$
TDVI	Transformed Difference Vegetation Index	$\sqrt{0.5 + \frac{(NIR - RED)}{(NIR + RED)}}$
VARI	Visible Atmospherically Resistant Index	$\frac{(NIR - GREEN)}{(NIR + GREEN - BLUE)}$
EVI	Enhanced Vegetation Index	$\frac{2.5 \times (NIR - GREEN)}{NIR + 6 \times RED - 7.5 \times BLUE + 1}$
Cig	Chlorophyll Index – Green	$\frac{(NIR)}{(GREEN)} - 1$
RVI	Ratio Vegetation Index	$\frac{(NIR)}{(RED)}$
SIWSI	Shortwave Infrared Water Stress Index	$\frac{(NIR - BAND6)}{(NIR + BAND6)}$
NDI7	Normalized Difference Index 7	$\frac{(NIR - BAND7)}{(NIR + BAND7)}$
NDWI	Normalized Difference Water Index	$\frac{(GREEN - NIR)}{(GREEN + NIR)}$
SAVI	Soil-Adjusted Vegetation Index	$1.5 \times \frac{(NIR - RED)}{(NIR + RED + 0.5)}$

As the final data group, five parameters based on elevation were derived from the Shuttle Radar Topographic Mission (SRTM) DEM product (SRTMGL3 DEM), including slope, flow direction, flow accumulation, flow length, and elevation.

2.1.2. Target variable

ET_a, as the target of the models, was obtained from the website https://earlywarning.usgs.gov/ssebop. This database covers the entire continental United States (CONUS) and provides daily ET_a estimates generated by the SSEBop model at 900 m resolution. This dataset utilizes 8-day available MOD11A2 LST data from MODIS, and ET_a is interpolated between each of the 8 days. Accordingly, the presented approach estimates daily ET_a which may further be accumulated and averaged over monthly or seasonal intervals. All the mentioned data were acquired from 09-30-2003 to 09-22-2019.

The SSEBop model (Senay et al., 2013) was developed to estimate ET_a across various temporal and spatial scales, using remotely sensed observations and climatological data. This model assumes that available net radiation (R_n) drives most of the surface energy balance process and that variations in land surface temperature (LST) can be used to quantify changes in ET_a (Senay et al., 2013; Zhuang et al., 2021). Consequently, the surface energy balance equation can be used to calculate dT by substituting R_n for the sensible heat (H) when H is assumed to be maximal, and the latent heat flux (LE) and ground heat flux (G) are assumed to be negligible on a daily time scale (Bastiaanssen et al., 1998b; Senay et al., 2013). A detailed description of the algorithm can be found in Senay et al. (2013).

2.1.3. Study areas

To ensure a comprehensive evaluation of DL models, three regions in the CONUS were selected to represent distinct climatological types according to the Köppen climate classification. The locations of all these regions are depicted in Fig. 1 (a), and their exact coordinates are presented in Fig. 1 (b). The first study area, located in the southeastern CONUS, is characterized by a hot and humid subtropical climate. The second region, located in the northeastern CONUS, features a humid continental climate with mild to hot summers. The third study area, situated in the western regions of the CONUS, is characterized by either low humidity or high temperatures and a combination of cold semi-arid, cold desert, hot/warm summer Mediterranean, and humid continental mild/hot summer. Each region spans the same land area, providing an equal and comparable number of samples for training and validating the DL models.

2.2. Image processing

A series of preprocessing steps was essential before integrating the collected data into the DL models. Therefore, all datasets were initially cropped based on the regions of interest. Then, the missing values in each raster file were filled using the average of the remaining pixels in that file for each input variable. Next, a bilinear resampling algorithm was applied to match the resolution of the inputs with the resolution of the SSEBop-ET_a product (900 m). Then, each raster file was reshaped into 1400 smaller 32 by 32 images to expedite the execution of DL models. Subsequently, all datasets were stored in the HDF5 format to facilitate efficient retrieval for future analysis.

Since the intention was to examine the effects of monthly weather dynamics, the datasets were organized into 12 different HDF5 files, each containing four daily modeled ET_a for a specific month of the year. To address the misalignment between the beginnings of the Gregorian calendar months and the start of seasons in the Northern Hemisphere (spring, summer, autumn, and winter beginning on March 20, June 21, September 23, and December 21, respectively), the seasons were divided into three-month intervals according to the aforementioned dates. Consequently, each month had data available for four days, except for the seventh and twelfth months, which contained data for three days.

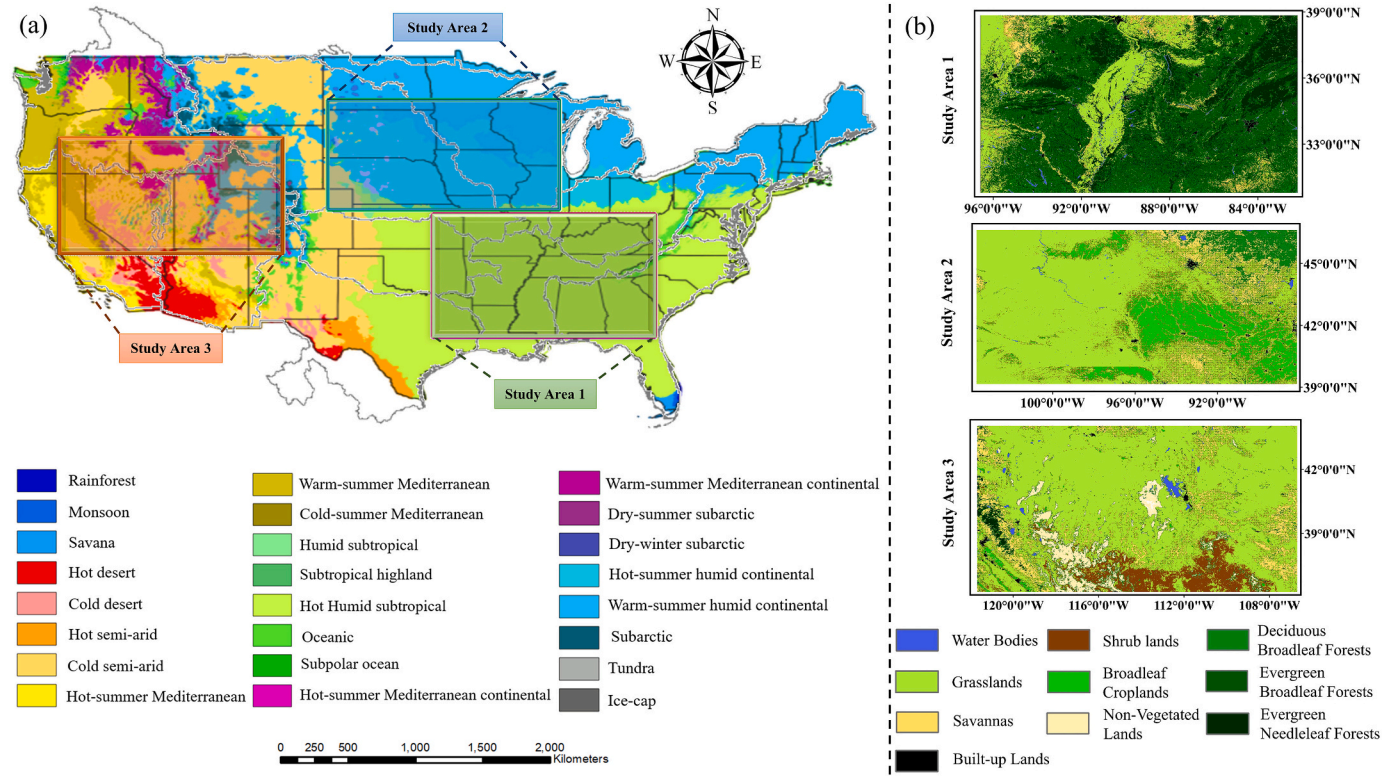


Fig. 1. Illustration of the location of three distinct study areas used in this study. Figure (a) represents the major basin boundaries and climates based on the Köppen climate classification within the CONUS. Study areas has been choosen such that reflect distinct types of climatology with the same land area. Figure (b) shows in more detail the coordinates of study areas as well as their land cover types based on the annual LAI classification.

The descriptions of the months are provided in Table 2.

In the next step, 15 indices were computed based on the data presented in Table 1 to improve the performance of DL models. Finally, the first 12 years of data were designated for training, the following two years for validation, and the final two years for testing the models. To automate the entire workflow of the present study, all the mentioned steps were implemented using the Python programming language.

2.3. Methodology

2.3.1. Deep learning attention mechanisms

In a deep CNN-based architecture, spatial information can often be lost due to cascading convolution and non-linearity in the high-level output mapping (Cui et al., 2021). To address this issue, we utilized and compared three distinct attention mechanisms to identify and extract pertinent spatial information more effectively. Attention mechanisms are effective in capturing contextual information and enhancing long-range dependency modeling (Li et al., 2021b).

In computer vision tasks, attention mechanisms are categorized into spatial attention and channel attention. Spatial attention mechanisms, such as Attention Gate (AG), are used to emphasize target locations and suppress irrelevant areas. In this study, AG was employed to capture spatial relationships and potentially improve DL models for ET_a estimation. Channel attention mechanisms, such as Squeeze and Excitation (SE), selectively weight feature map channels to improve prediction accuracy (Li et al., 2021a). The incorporation of the SE mechanism in this research may enhance the representational capacity of the U-Net model and improved ET_a estimations. The detailed mathematical equations and structure of AG and SE can be found in the respective articles (Hu et al., 2018; Schlemper et al., 2019). More importantly, in addition to implementing the AG and SE mechanisms, in this study, we developed a new attention mechanism and incorporated it into the U-Net architecture to capture spatial dependencies more effectively within the U-Net framework. The subsequent section will detail the structure of the developed attention mechanism.

Table 2
Detail on the monthly division of available data.

Seasons	Order of Months	Day of Year (based on the Gregorian calendar)
	1	81, 89, 97, 105
Spring	2	113, 121, 129, 137
	3	145, 153, 161, 169
	4	177, 185, 193, 201
Summer	5	209, 217, 225, 233
	6	241, 249, 257, 265
	7	273, 281, 289
Autumn	8	297, 305, 313, 321
	9	329, 337, 345, 353
	10	361, 1, 9, 17
Winter	11	25, 33, 41, 49
	12	57, 65, 73

2.3.1.1. Activated spatial attention mechanism (ASPAM). We hypothesize that utilizing matrix multiplication can more effectively capture spatial dependencies among feature channels by including more arrays in matrix computations compared to the convolutional operations or Hadamard multiplication used in SE and GA. Inspired by the SAINA attention mechanism, a self-attention module initially proposed for time series analysis by Alizadeh et al. (2021), we developed a novel activated spatial attention mechanism (ASPAM) for image regression tasks within the U-Net architecture. In general terms, given a query (Q) and a set of key-value pairs (K), this attention mechanism can be characterized by computing a weighted sum of values (V) based on the query and the corresponding keys. Similar to AG, ASPAM takes two inputs: the up-sampling feature or gating signal (G) to generate both the query (Q) and value (V) matrices, and the corresponding input feature (F) to

compute the key (K) matrix. Accordingly, the proposed attention mechanism (Fig. 2) can be expressed as follows:

$$G = [G_1, G_2, \ldots, G_c]_{(c \times x \times y)}$$

$$F = [F_1, F_2, \ldots, F_c]_{(c \times x \times y)}$$

$$Q = sigmoid\left(G_{(c \times x \times y)} \otimes W_{q(c \times x \times y)}\right)$$

$$K = tanh\left(F_{(c \times x \times y)} \otimes W_{k(c \times x \times y)}\right)$$

$$V = G_{(c \times x \times y)}$$

$$S_{(c \times x \times y)} = softmax\left(sigmoid\left(Q \otimes K^T\right)\right)$$

$$O_{(c \times x \times y)} = V \otimes S \quad (1)$$

where W_q and W_k are trainable weights, S represents the attention score matrix, O is the output (or, in other words, the attended gating signal matrix), and $\otimes$ denotes the matrix multiplication operator. The indices c, x, and z represent the dimensions of matrices and are the given values for the number of channels and dimensions of images, respectively.

Similar to SAINA attention, we observed that applying the sigmoid and hyperbolic tangent functions prior to computing Q, K, and S can lead to decreased training and validation errors. We reason that these functions can enhance the calibration of the output values. It is worth noting that the x and y dimensions should be equal in these operations. Furthermore, the matrix multiplication operation is performed channel-wise for 3D matrices, resulting in a matrix with the same dimensions as its multipliers. However, in contrast to the SAINA module, we used the input feature (F) to compute the key matrix and employed matrix multiplication instead of Hammond multiplication in Eq. (1) to capture spatial dependencies more effectively.

2.3.2. ASPAM-U-Net architecture

The Fully Convolutional Neural Network (FCNN) (Tran, 2016) has been used for various image segmentation tasks, with the U-Net being a prominent variant that has demonstrated significant accuracy improvements in image analysis (Cui et al., 2021). The U-Net combines low-level features with high-level semantic representation (Asgari Taghanaki et al., 2021) through its encoder and decoder, while the skip connection between the encoder and decoder ensures that the U-Net maintains the full input image context. Thus, the U-Net architecture was employed in this study to improve the input images' robustness to disturbances, mitigate overfitting, and reduce ET_a estimation training time.

More importantly, to enhance the capability of U-Net in capturing spatial dependencies, we integrated ASPAM modules into its architecture (Fig. 2), with the number of modules varying based on the optimization of model depth (2–4 layers). Hyperparameters were optimized using a metaheuristic optimization algorithm (OA), facilitating the process of architecture depth and other hyperparameter fine-tuning. Fixed hyperparameters across all models included transpose convolution kernel size (2,2), 2×2 max pooling, and batch size of 128, the latter optimized for accuracy, GPU memory, and performance. The number of filters in the convolution layers doubled in the encoder and halved in the decoder, reaching one feature at the final layer. Two dropouts and a batch normalization layer with Tanh activation were incorporated into each encoder and decoder segment to reduce the model's sensitivity to initialization parameters and provide regularization. (Wang et al., 2021). The AG and SE modules were incorporated into U-Net based on the architectures developed by John and Zhang (2022) and Rundo et al.

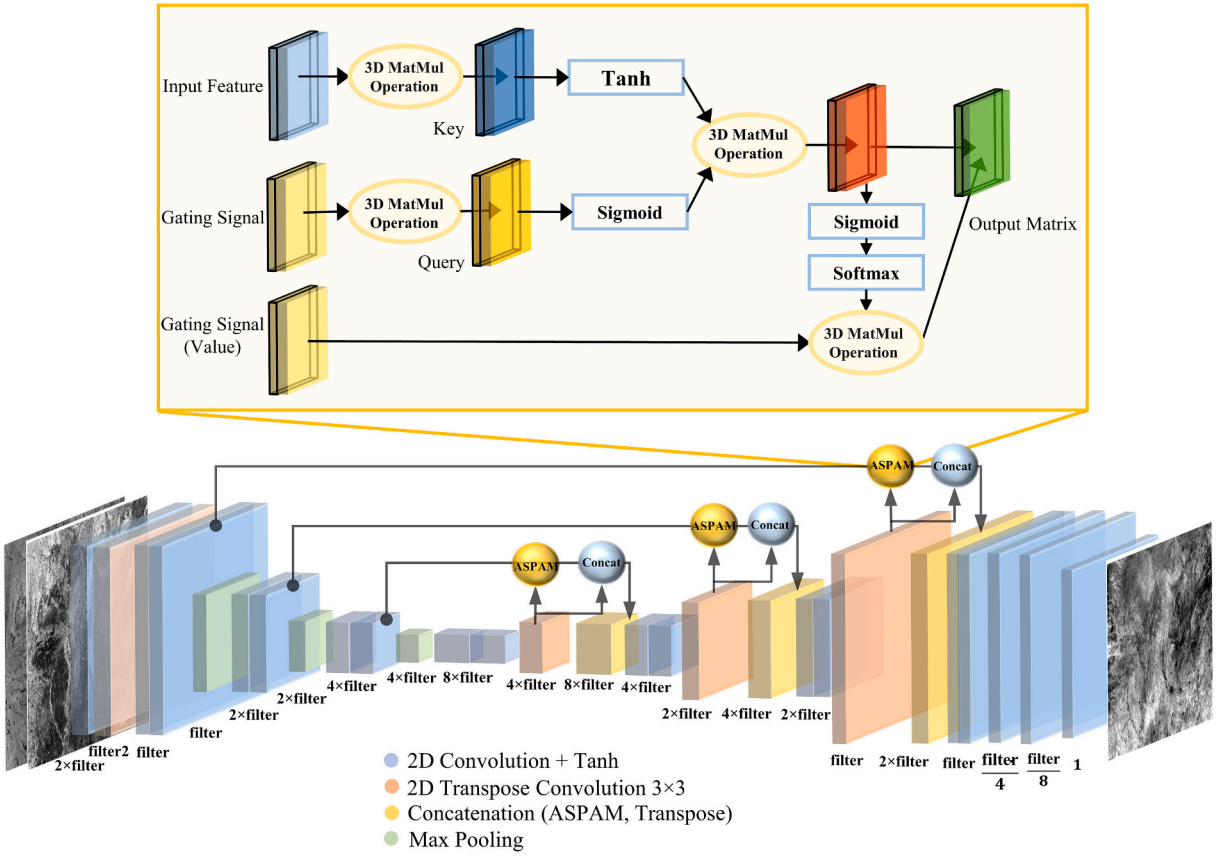


Fig. 2. Schematic representation of the spatial attention-based U-Net and the detailed structure of the novel activated spatial attention mechanism (ASPAM). ASPAM takes the up-sampling feature as the gating signal from the encoding part and the input feature as the key from the decoding section. The architecture proceeds by concatenating the output matrix with the input feature.

(2019), respectively, with the SE modules of the encoder from the latter article excluded.

2.3.3. Harris Hawks optimization (HHO)

Hyperparameter tuning is a critical step in developing DL algorithms, as optimized parameters can significantly enhance the model's performance. Nonetheless, despite the increasing use of DL models in hydrology, the prevailing practice remains experimental trial-and-error and reliance on recommended values (Li et al., 2021b). These approaches, however, can be time-consuming and do not account for the combined effects of hyperparameters (Sameen et al., 2020), which could be problematic when comparing the performance of several models (Alizadeh et al., 2021).

To address this, a reliable and fast optimization algorithm is essential for developing an AutoDL-based model that encompasses all phases, from image processing to the automatic determination of the best architecture and optimal hyperparameters. Accordingly, we employed the Harris Hawks Optimization (HHO), a novel swarm algorithm introduced by Heidari et al. (2019). Despite its demonstrated superiority over other optimization algorithms (Moayedi and Mosavi, 2021a, 2021b), the application of HHO in the domain of atmospheric and hydrologic predictions has yet to be extensively investigated.

HHO mathematically simulates Harris's hawk hunting tactics, involving exploration, surprise, hunting, and modification of chase patterns based on the dynamics of the situation and the prey's escape behaviors. HHO, similar to other optimization methods, encompasses stages of exploration, and exploitation as well as a transitional phase. Detailed mathematical information about this algorithm can be found in the article written by Heidari et al. (2019).

2.3.4. Framework of U-Net-based models

The main objective of this study is to propose a new method for large-scale estimation of daily ET_a through AutoDL models. This approach aims to automate the ET_a extraction process while ensuring accuracy comparable to energy-based models such as SSEBop. Both methods use MOD11A2 LST as input data. However, the AutoDL models differ from the SSEBop model in that they do not require the utilization of meteorological input data from multiple sources. Instead, they use the MODIS products as described in Section 2.1.1 facilitating data collection procedure. By learning to simulate the structure of the SSEBop model, the AutoDL models can eliminate the need for auxiliary inputs and automate the ET_a estimation process.

The flowchart of AutoDL models is illustrated in Fig. 3. After performing the required data preprocessing for the three main types of inputs, as detailed in Section 2.1.2, the initial 12 years of data was selected for training, followed by 2 years for validation, with the final two years (09-30-2017 to 09-22-2019) allocated for testing the DL model's performance for each month and study area. The HHO was run for five iterations to determine the optimal set of hyperparameters and architecture. The initial population for the algorithm was set to 5 based on experimental trial. The hyperparameters of each model were optimized for the first month of the year and the first study area, and the optimized sets were applied to the other months and study areas. Table 3 presents the optimized hyperparameters and their corresponding search

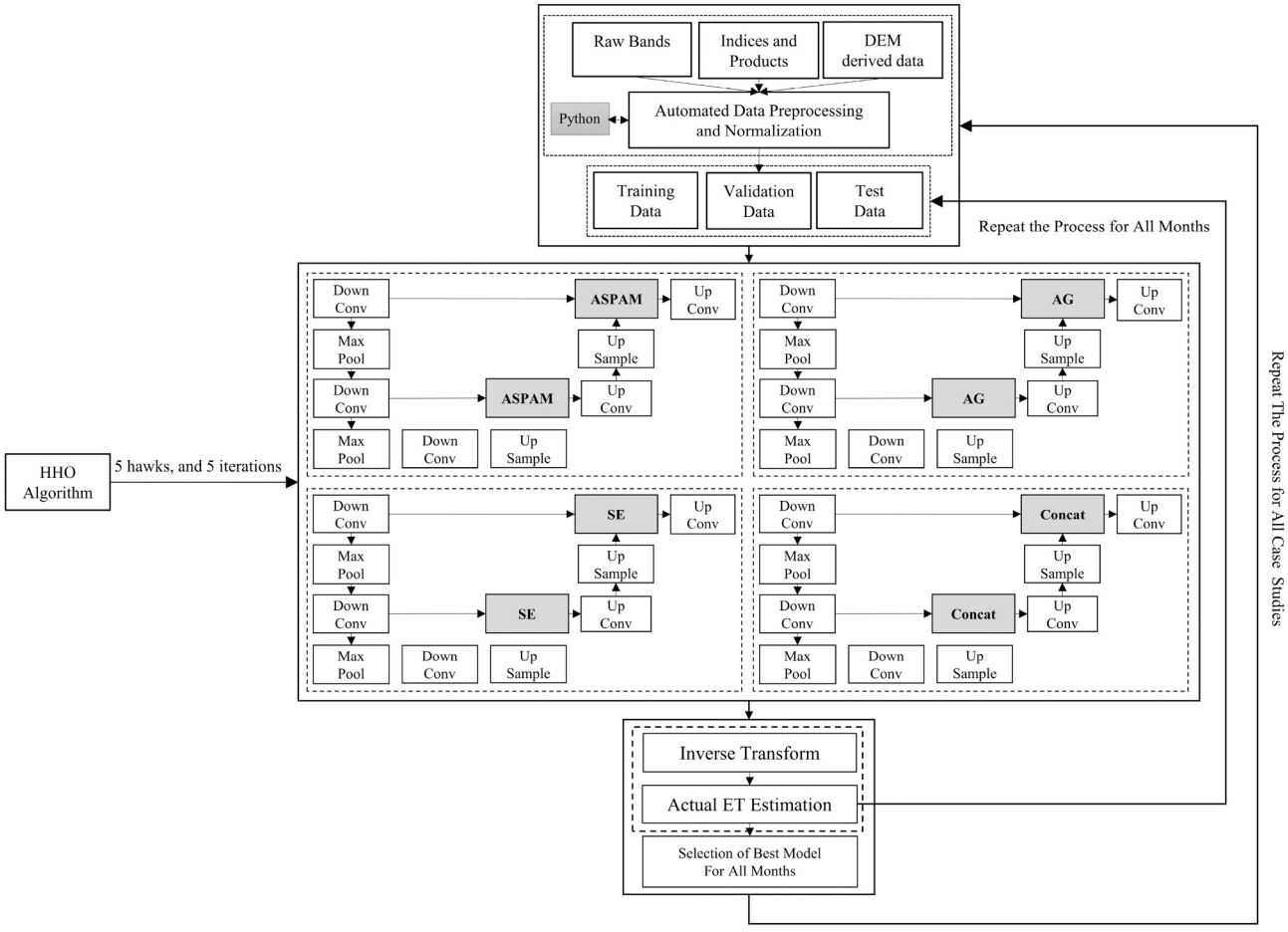


Fig. 3. Illustration of the AutoDL model framework for actual evapotranspiration (ET_a) estimation, employing the novel ASPAM-U-Net architecture in conjunction with the Harris Hawks optimization (HHO) algorithm for hyperparameter fine-tuning. Data was preprocessed and segmented for monthly DL training across study areas. Using the initial hyperparameter optimizations, the most effective U-Net-based architecture was selected per month, based on the evaluation metrics.

Table 3
Details on the hyperparameters of each DL model and their corresponding range.

Hyperparameters	Range	ASPAM-U-NET	AG-U-NET	SE-U-NET	Simple-U-NET
Dropout	[0.2, 0.5]	✓	✓	✓	✓
Kernel Size	(3,3), (5,5)	✓	✓	✓	✓
Filters (of the First Layer)	16,32,48,64	✓	✓	✓	✓
Filters	16,32,48,64	✓	✓	✓	✓
Learning Rate	[0.00001, 0.005]	✓	✓	✓	✓
Architecture Depth	[1,3]	✓	✓	✓	✓
Intermediate Channels 1	16,32,48,64	–	✓	–	–
Intermediate Channels 2	16,32,48,64	–	✓	–	–

spaces. The best model for each month was then selected from four U-Net-based architectures based on evaluation metrics. The process was repeated for the remaining months and study areas.

2.3.5. Evaluation Metrics

The performance of four AutoDL models was evaluated across different climatology types and temporal windows throughout the year to reflect the impact of weather fluctuations. Daily ET_a predictions were combined to provide a monthly assessment. Several metrics were used to evaluate the accuracy of the models. The primary accuracy measures for deterministic estimation included Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and coefficient of determination (R^2).

MAE and MAPE are more meaningful when the overall impact of errors is proportional to the increases in error, with MAPE providing the error in percentage form, making it easier to interpret across different cases. MAE and MAPE can be written as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left[\left|ET_i^p - ET_i^m\right|\right] \tag{2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left[\left|ET_i^p - ET_i^m\right|\right]}{ET_i^m} \tag{3}$$

where ET_i^p indicates the i^{th} pixel ET prediction, ET_i^m represents the modeled ET, and n is the total number of pixels in a test raster.

Another important metric, R^2, indicates the extent to which the model explains the variance in the dependent variable. An R^2 value of one signifies a highly accurate forecast, while an R^2 value of zero implies that the model cannot explain any variation in the outcome. In cases where the chosen model performs worse than a simple horizontal line, R^2 may be negative. The R^2 can be calculated using Eq. (4):

$$R^2 = 1 - \frac{RSS = \sum_{1}^{n}\left(ET_i^p - ET_i^m\right)^2}{TSS = \sum_{1}^{n}\left(ET_i^m - \overline{ET}^m\right)^2} \tag{4}$$

where RSS stands for the residual sum of squares, TSS denotes the total sum of squares, and $\overline{ET}^m = \frac{1}{n}\sum_{i=1}^{n} ET_i^m$ is the mean of modeled ET.

RMSE (Root Mean Square Error) penalizes larger errors and is more favorable for evaluating the performance of models for extreme values. The RMSE can be formulated as Eq. (5):

$$RMSE = \sqrt{\frac{\sum_{1}^{n}\left(ET_i^p - ET_i^m\right)^2}{n}} \tag{5}$$

3. Results and discussion

A total of four AutoDL models, including three with spatial and channel attention mechanisms, were employed using the HHO algorithm and validated across three distinct study areas within the CONUS throughout the year to assess their performance under varying climatological conditions for daily ET_a estimation. As discussed in Section 2.1.1, the models were primarily constructed using MODIS satellite optical and thermal data, eliminating the reliance on ancillary meteorological information. The models were trained using outputs from the SSEBop, an energy-based model for ET_a estimation. The performance of the models was evaluated on a monthly and six-monthly basis by averaging daily ET_a estimates.

The development of an AutoDL model depends on the implementation of a robust optimization algorithm. Fig. 4 depicts the results obtained from the application of HHO in determining the best set of DL hyperparameters. As illustrated in Fig. 4, the initial populations yielded higher validation and training errors as they were randomly selected. As the algorithm proceeded through subsequent iterations, the errors progressively decreased, ultimately reaching the optimum set. As a result, in comparison to the best values in the first iteration, the HHO reduced the validation error (normalized MAE) by 18.03%, 17.34%, 5.79%, and 15.28% for the ASPAM, AG, SE, and Simple-U-Net models, respectively. Therefore, as an effective metaheuristic optimization algorithm, HHO could facilitate and expedite the process of hyperparameter tuning.

3.1. AutoDL model's goodness of fit evaluation

Fig. 5 presents a comparative analysis of the R^2 values for four DL models across three study areas over the course of 12 months. All models evaluated in the three study areas exhibited notably high R^2 values during the first half of the year, maintaining errors within an acceptable range (Chen, 2016; Singh and Senay, 2015). This is particularly remarkable considering that meteorological data were not utilized and that there were differences in processing methods between input and output variables, as discussed in Section 2.1.1.

Study area 1 is characterized by a hot, humid subtropical climate and is mainly covered by evergreen broadleaf forests (41.8%), shrubland (30.6%), and needle-leaf forests (21.26%), based on the annual LAI classification. In the study area, although all models exhibited similar R^2 patterns during the first six months, ASPAM-U-Net outperformed other models with a 17.5% accuracy improvement in the third month and 5.5% overall. Its R^2 values ranged from 0.65 to 0.82, surpassing the minimum and maximum values of other models. Subsequent best performances were observed by the AG-, Simple-, and SE-U-Net models, in that order.

Study area 2 features a humid continental mild/hot summer climate and consists of 57.93% shrublands, 18.42% savannas, and 8.82% broadleaf croplands. Similar to study area 1, it has a relatively flat topography but cooler temperatures in general. Similar to previous regions, the ASPAM-U-Net outperformed others by increasing accuracy by 4.3% during the first half of the year, with R^2 values ranging from 0.53 to 0.79, slightly lower than the prior area. The highest maximum R^2 values were achieved by the ASPAM-U-Net, followed by the AG-, Simple-, and SE-U-Net models in a similar order as the previous case.

Study area 3 encompasses a combination of cold semi-arid and desert climates, summer Mediterranean climates, and humid continental summer climates and is comprised of 62.44% shrublands, 12.94% broadleaf croplands, and 8.03% broadleaf forests. Unlike the previous two areas, this region exhibits greater topographical and climatological variety. Except for the second month, all models demonstrated comparable performance, where, in contrast to study areas 1 and 2, the AG-U-Net surpassed others, followed by the Simple-, ASPAM-, and SE-U-Net models. The AG-U-Net improved the R^2 by 6.87% during the first half of the year, achieving the highest minimum and maximum values (0.59 and 0.91).

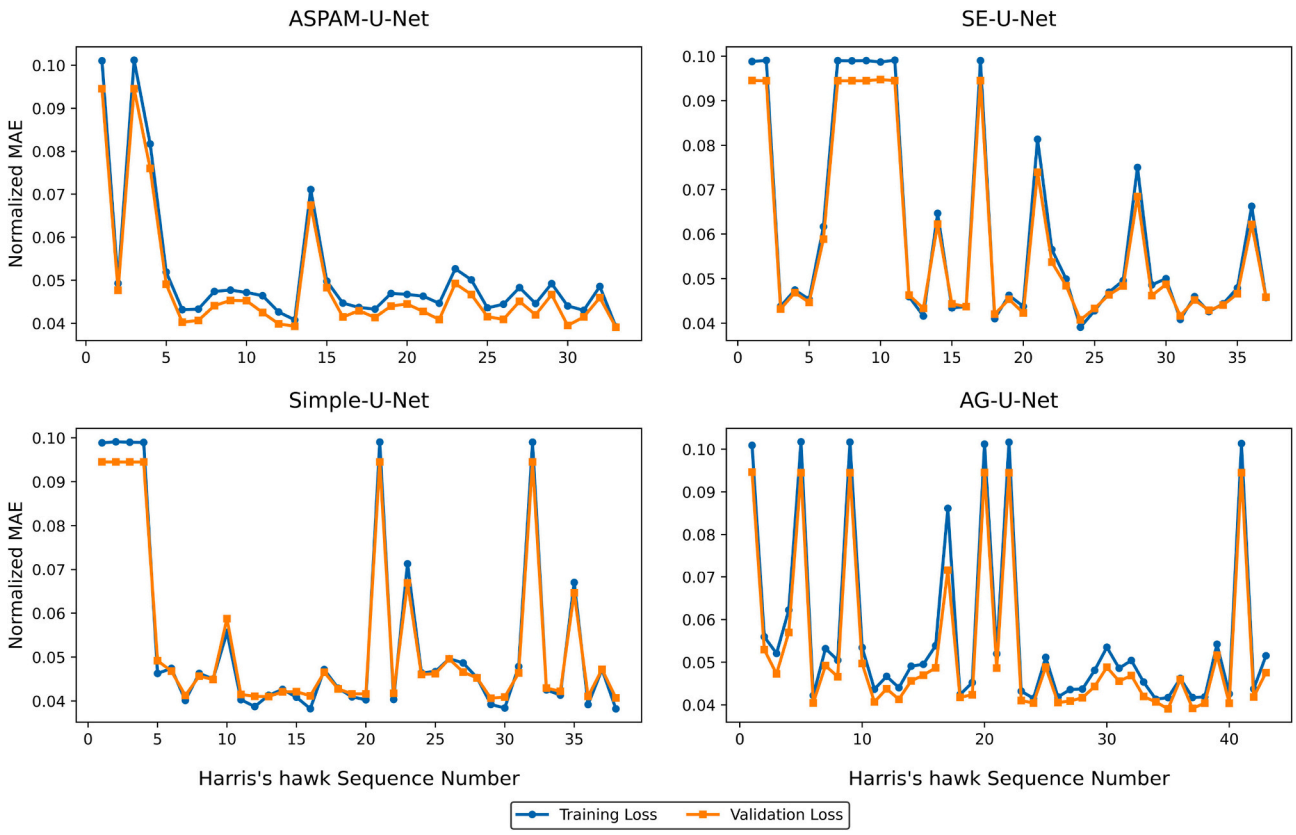


Fig. 4. Performance analysis of Harris Hawks optimization algorithm utilized for the fine-tuning of U-NET-based models based on MAE values of training and validation data. Initially, random population selections resulted in high errors. As the algorithm iterated, validation errors notably reduced, proving its efficacy in hyperparameter tuning.

In contrast to the elevated R^2 values observed during the first six warm months, the values of R^2 were lower and near zero during the second colder half of the year in all study areas, particularly from the 8th to the 11th months. This phenomenon aligns with findings from Xu et al. (2019). Evaluating the performance of SSEBop, Xu et al. (2019) found that this model exhibits larger uncertainties in the colder months of the year. These uncertainties are attributed to inaccuracies in satellite-retrieved LST, resulting in substantial errors in SSEBop-ET_a calculations (Xu et al., 2019). Accordingly, lower reliability and R^2 values during the latter half of the year can ratify the hypothesis that colder months lead to increased uncertainty in the target variable (SSEBop-ET_a), thereby complicating the establishment of meaningful relationships between input variables and the target variable for DL models. This can elucidate the higher R^2 values observed in study area 3 compared to study area 1 and in study area 1 compared to study area 2, as study area 3 features warmer and drier climates, followed by study areas 1 and 2. This pattern positively influenced the certainty of the target SSEBop-ET_a variable.

Considering the potential impact of cold weather on MODIS LST data, some studies have excluded observations from the cold months (Ji et al., 2019). Similarly, the present study confines the analysis to the first six warmer months of the year, when LST and SSEBop-ET_a are presumed to exhibit high certainty.

3.2. Monthly performance evaluation

Fig. 6 compares the distributions of absolute error (AE) and absolute percentage error (APE) for four DL models across three study areas during the first six months of the year. In the first study area, the ASPAM-U-Net outperformed others, showing lower AE and APE values in the 25th, 50th, and 75th percentiles, where other models exhibited a similar performance. This model reduced AE and APE errors by 8.35% and 11%, respectively, with lower extreme error values compared to the other models in most cases. Consistent with the result of R^2, the highest accuracy was observed during the 4th and 5th months, corresponding to the hottest months in the Northern Hemisphere. The ASPAM-U-Net maintained the range of average AE and APE (MAE and MAPE) between 6.38 and 16.36%, and 0.38–0.53 mm per day (mm/d), respectively, yielding the lowest maximum and minimum values for the specified metrics.

In the second study area, all models produced comparable results, with the ASPAM-U-Net reducing MAE and MAPE by 5.7% and 9.3%, respectively. Analogous to the first study area, the ASPAM-U-Net achieved the lowest 25th, 50th, and 75th percentiles for both AE and APE in most months, followed by the AG-, Simple-, and SE-U-Net models, while error results were substantially lower in the 4th and 5th months. The ASPAM-U-Net exhibited the narrowest variation range for the MAPE metric (7.02–37.56%). For MAE, this range spanned from 0.40 mm/d (by ASPAM-U-Net) to 0.73 mm/d (by AG-U-Net). In comparison to previous study areas, these ranges were higher, likely attributable to the colder temperatures in study area 2, leading to decreased certainty in SSEBop-ET_a values, as hitherto discussed.

In contrast to the prior study areas, in study area 3, the AG-U-Net generated superior performance, followed by the Simple-, ASPAM-, and SE-U-Net models. The AG-U-Net decreased MAE and MAPE by 12.43% and 14.76%, respectively, achieving the lowest 25th, 50th, and 75th percentiles for both AE and APE in the majority of cases. Following a similar trend to previous regions, the 6th month exhibited the smallest variation range for AE and APE, followed by the 5th and 4th months. The AG-U-Net achieved the lowest minimum and maximum values for MAPE and MAE, ranging between 15 and 29.17%, and 0.22–0.38 mm/d, respectively. AE values and ranges in this region were smaller compared

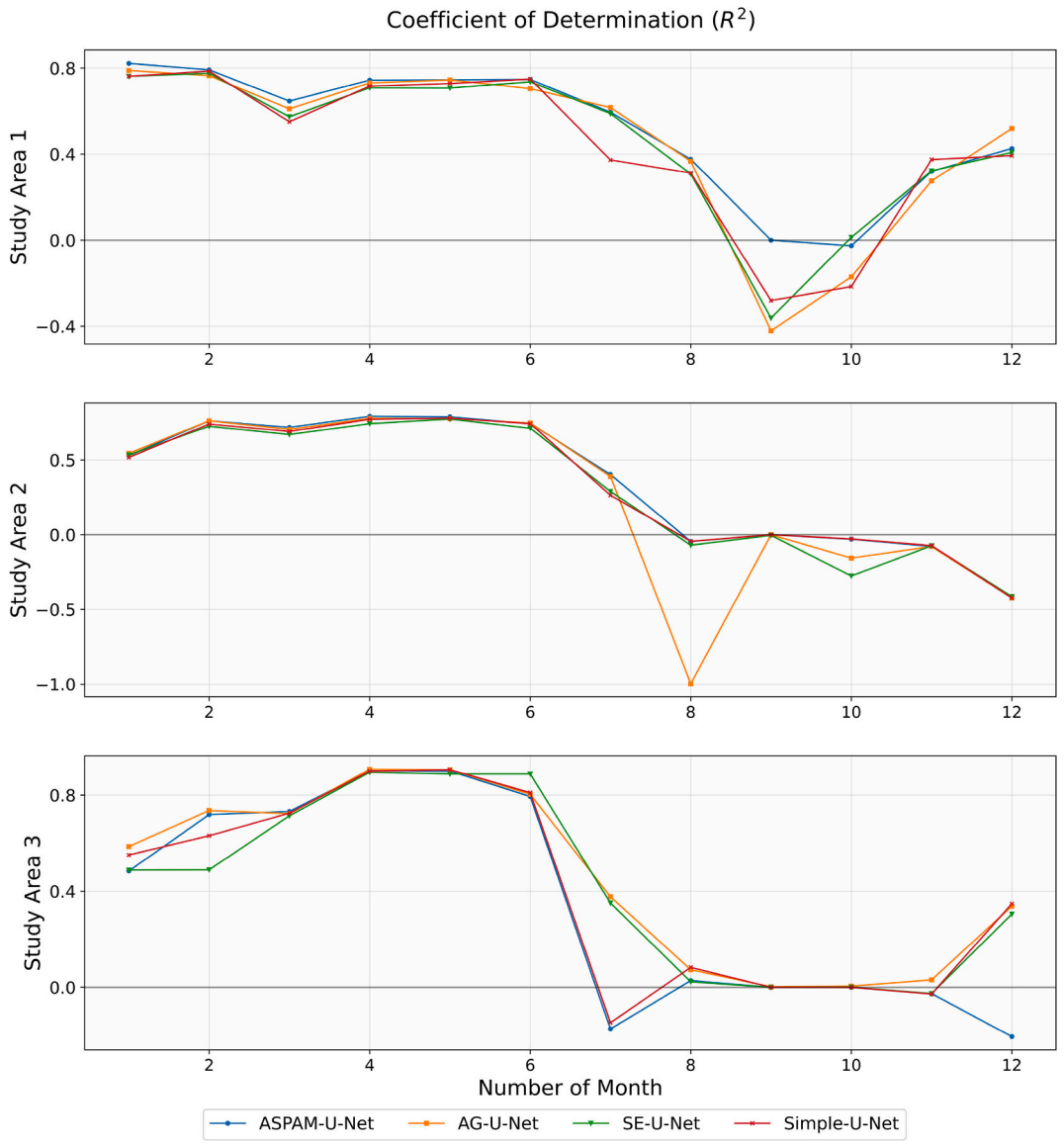


Fig. 5. Comparative analysis of monthly coefficient of determination (R^2) values among four DL models across three distinct study areas over a 12-month timeframe. In the first half of the year, all models exhibited high R^2 values across all study areas, with ASPAM-U-Net notably outperforming in areas 1 and 2. Inaccuracy in satellite-derived LST and subsequent SSEBop-ET_a calculation in colder months introduced uncertainty and lowered R^2 values in this time of the year, complicating meaningful DL model input-target relationships.

to previous regions, possibly due to the lower vegetation coverage, resulting in reduced ET_a values. On the other hand, since this region features warmer climates, the models were expected to produce low values for APE. Xu et al. (2019) noted that SSEBop exhibits higher relative uncertainties in the western CONUS as compared to the eastern CONUS, which may be attributed to the complex terrain in the western mountainous regions. Moreover, models tend to generate larger uncertainties in ET_a estimation over arid, semi-arid, and bare land areas (Chen et al., 2016).

3.3. Six-month spatial performance analysis

Fig. 7 illustrates the average ET_a maps generated by AutoDL models for the three study areas over a six-month period. A comparison of the estimated ET_a maps with the corresponding SSEBop-ET_a revealed similar spatial patterns for the investigated regions, indicating that the DL models effectively captured the spatial variability of ET_a. The ET_a variabilities (mean $\pm$ standard deviation) for the study areas 1 through 3 were 5.24 $\pm$ 0.66, 3.79 $\pm$ 0.55, and 1.69 $\pm$ 1.21 mm/d, respectively. In study area 3, the diverse climate and topography coupled with sparse vegetation resulted in a broad ET_a variability, whereas study area 1, with its dense vegetation and warmer climate, exhibited the highest average ET_a. Overall, all AutoDL models maintained the average and standard deviation values closely aligned with the target benchmarks, while the ASPAM-U-Net, followed by the AG-U-Net achieved superior accuracy.

By comparing Fig. 7 with Fig. 1 (b), it can be inferred that the land cover type had a strong correlation with the ET_a magnitude. In all study areas, urban areas and grasslands experienced lower ET_a values, followed by savannas, whereas water bodies and evergreen needleleaf forests had the highest ET_a. Furthermore, in study area 2, croplands demonstrated moderate ET_a, while in study area 3, shrublands and non-vegetated lands displayed small ET_a values. As a result, in addition to temperature, estimated ET_a values revealed a correlation with vegetation density, as represented in the DL models by indices such as the NDVI.

Fig. 8 illustrates the spatial variation in MAPE, RMSE, and MAE errors for the four DL models across all study areas. By comparing the

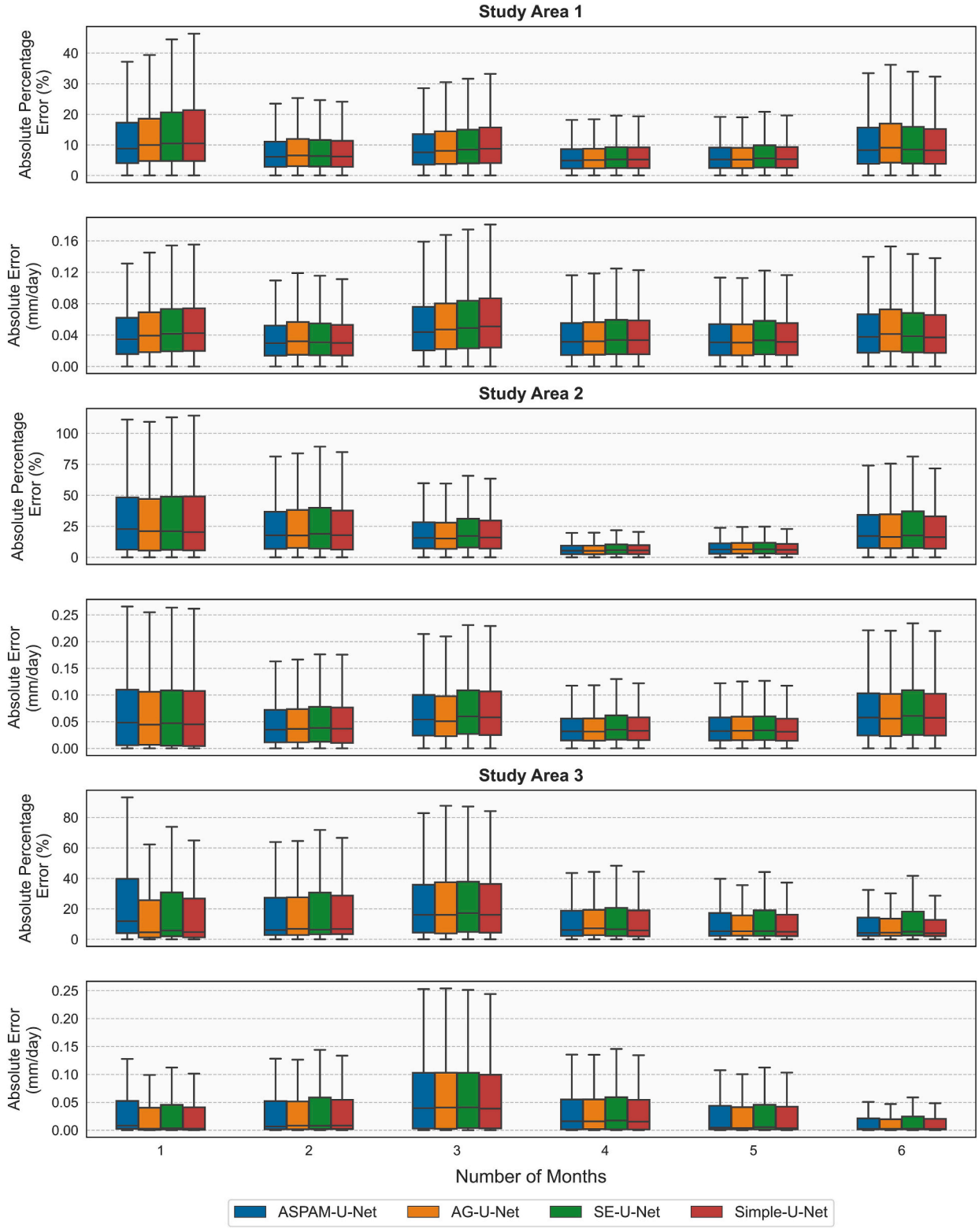


Fig. 6. Boxplot representation of Absolute Error and Absolute Percentage Error metrics among four DL models across three distinct study areas during the first six-months of the year. In study areas 1 and 2, ASPAM-U-Net surpassed other models in AE and APE. Accuracy peaked during warmer months in each area. Study area 2 had higher MAE due to its colder temperature, while study area 3 had higher APE values due to a more complex terrain.

results in this figure with Fig. 1 (b) for study areas 1 and 2, it can be inferred that water bodies, urban areas, grasslands, and savannas exhibited the highest values of errors, particularly in terms of MAE and RMSE, while areas covered by deciduous and evergreen broadleaf forests, evergreen needleleaf forests, and crop fields demonstrated the highest accuracy. This implies that in these two similar regions (in terms of stable climatic and topographical conditions), the models demonstrated enhanced performance in ET_a estimation over areas with high canopy density. Still, in both study areas, the ASPAM-U-Net not only reduced the extreme values of errors but also achieved lower average

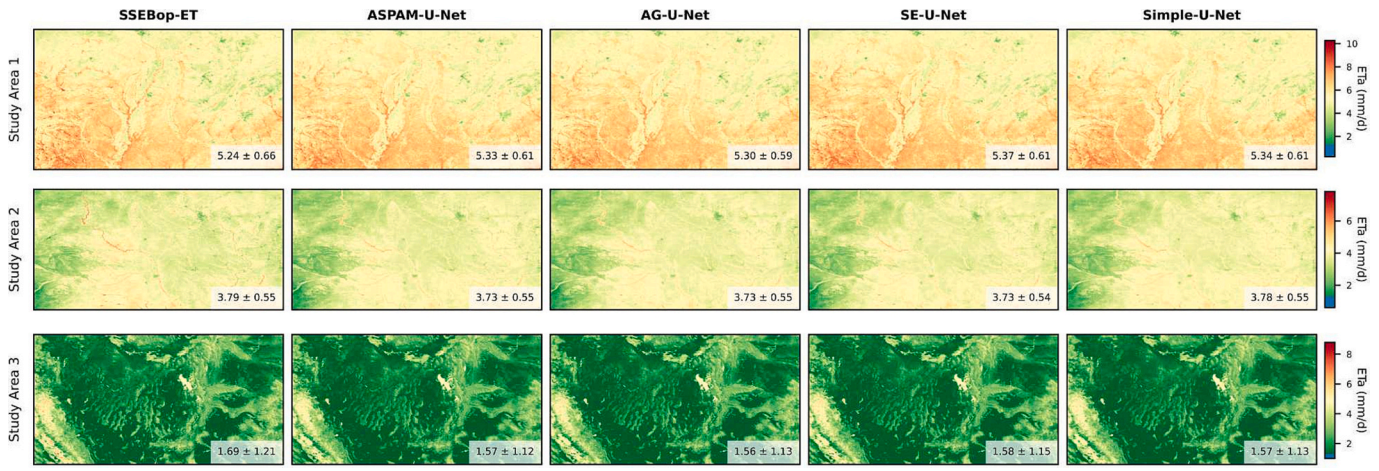


Fig. 7. Comparison of spatial variability in six-month mean values of modeled ET_a and SSEBop-ET_a. Each subplot's lower-right corner displays the corresponding mean value ± standard deviation of ET_a, expressed in millimeters per day (mm/d). Estimated ET_a and SSEBop-ET_a showed similar spatial patterns, demonstrating DL models' effectiveness in capturing ET_a variability. Study areas showed varied ET_a due to climate and vegetation, yet AutoDL models consistently aligned with benchmarks.

errors. Specifically, as depicted in Fig. 9, the ASPAM-U-Net had average MAPE, RMSE, and.

MAE values of 10.17%, 0.573 mm/d, and 0.43 mm/d in study area 1, and 24.58%, 0.82 mm/d, and 0.575 mm/d in study area 2, respectively, which was lower than the other models.

In study area 3, the highest values of MAPE, RMSE, and MAE were primarily associated with water bodies, urban areas, and savannas, while grasslands, evergreen needleleaf forests, and crop fields experienced moderate values of errors. However, in contrast to previous regions, areas with minimal to no canopy coverage (i.e., shrublands and non-vegetated areas) exhibited the highest accuracy. This observation indicates that, in addition to canopy density, the overall regional climatology also influences models' performance. In this region, the AG-U-Net demonstrated superior performance with average RMSE, MAE, and MAPE values of 0.69 mm/d, 0.37 mm/d, and 21.48% respectively, followed by the Simple-, ASPAM-, and SE-U-Net.

By comparing Fig. 1 (b), Fig. 7, and Fig. 8, it can be inferred that despite the absence of a concrete relationship between the magnitude of error and ET_a values, both exhibit a strong correlation with land cover types. Furthermore, according to the monthly and average six-month analysis of the results, incorporating spatial attention mechanisms in the U-Net architecture proved to be efficient improving the performance in all study areas, validating the hypothesis that incorporating spatial dependencies among input variables into the DL models is effective. More importantly, the developed ASPAM mechanism successfully outperformed other DL models in two out of three study areas and exhibited comparable performance in study area 3.

Fig. 9 presents the scatter plots comparing the six-month average values of predicted ET_a with SSEBop-ET_a for all DL models across study areas. The average R^2 values across all regions were closely aligned, ranging from 0.79 (by SE-U-Net) to 0.85 (by the ASPAM-U-Net). This result indicates that all models generally demonstrated comparable performance. In the first two study areas, the ASPAM-U-Net exhibited the highest R^2, followed by the AG-, Simple-, and SE-U-Net models, while in study area 3, the AG-U-Net outperformed the others, followed by the Simple-, and ASPAM-U-Net models. This pattern is consistent with the monthly analysis. Although the data are diverse across all study areas, the highest density of values is near the diagonal line, indicating that the majority of points had high accuracy.

Based on the results, the main objective of the present study, proposing an AutoDL model for large-scale daily ET_a estimation as an alternative to energy-based models has been successfully achieved. However, further validation of the model using measured data such as Eddy covariance ET_a, was beyond the scope of this study (Babaeian et al., 2022; Yan et al., 2023), and is recommended for future studies.

3.4. Sensitivity analysis

As described in the image processing section, the data were categorized into three groups: raw bands, derived indices and products, and DEM-derived products. To assess the influence of each group on model accuracy and identify the most influential group, a Gaussian noise with a standard deviation (σ) equal to 5% of σ of each feature was added to four distinct combinations of these data groups: i) a combination of the first and third groups; ii) a combination of the second and third groups; iii) the first group of data alone; and iv) the second group of data alone. These combinations were applied to the first and fourth months in all three study areas as inputs for the ASPAM-U-Net model. Fig. 10 presents the test data results for four distinct metrics. In all scenarios, the model's performance was reduced compared to the utilization of all three data sets without perturbation, indicating that utilizing all three data groups with the lowest uncertainty is necessary to achieve the highest accuracy.

Based on Fig. 10, the third data group (DEM-derived products) displayed slight sensitivity to the introduced noise. This suggests that these features might hold less significance in the modeling framework. On the other hand, the second dataset was noticeably more vulnerable to added noise, consistently showing higher sensitivity than the first data group across all study areas, especially in terms of the R^2 metric. However, the reduction in accuracy was not consistent, varying both by study location and month of the year. The higher sensitivity of this group of data might be attributed to the physical correlation between vegetation and water indices with ET_a, especially the LST. In other words, understanding the vegetation and moisture characteristics of a surface is expected to contribute to the ET_a of that surface.

4. Summary and conclusion

In this paper, a novel data-driven framework for mapping ET_a is proposed. It is shown that, contrary to classic energy-based models, the proposed methodology overcomes the challenges associated with acquiring multi-source meteorological data by utilizing only optical and thermal products of MODIS and DEM-derived data while providing comparable accuracy. The main contributions of this research are summarized as follows:

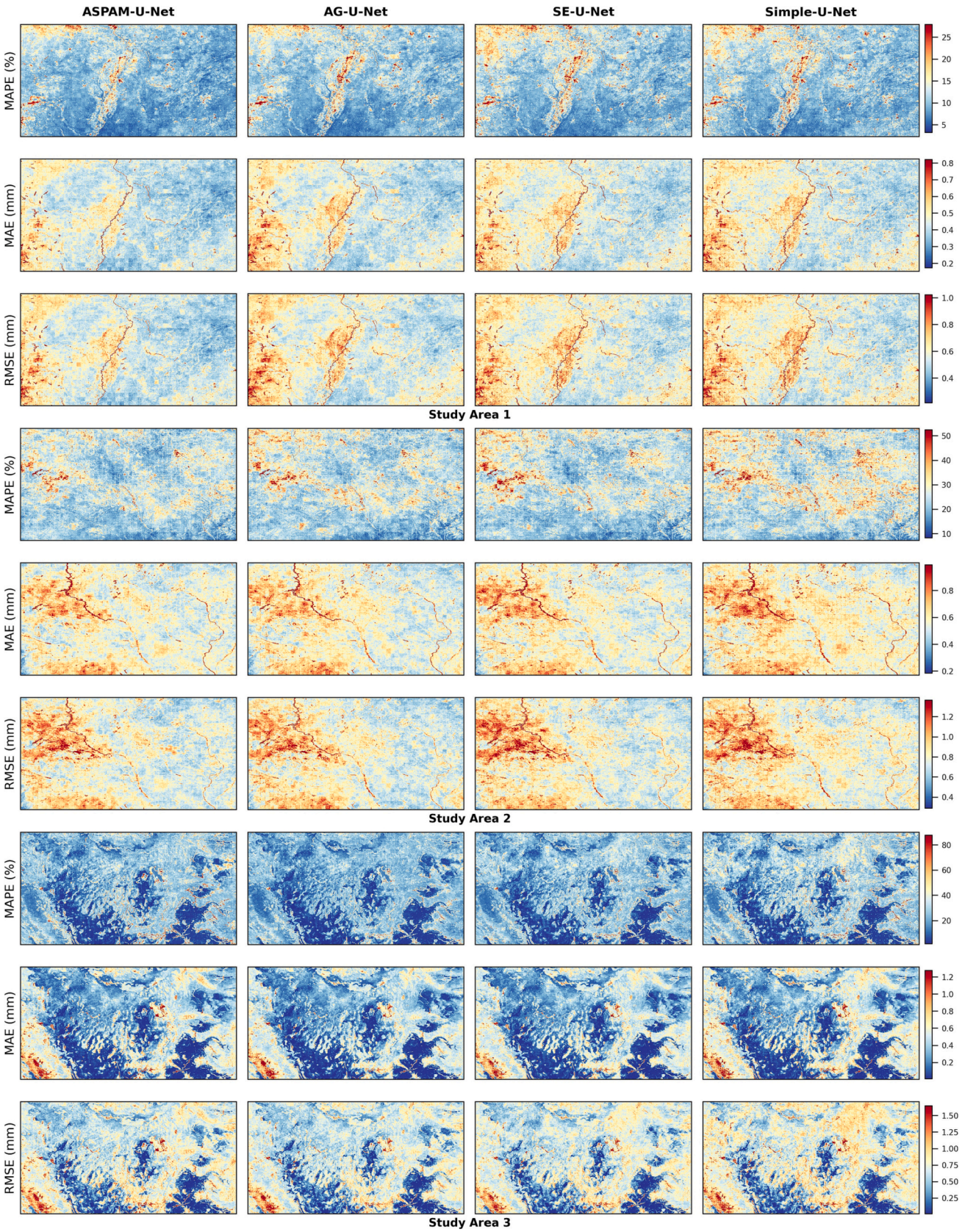


Fig. 8. Maps illustrating the spatial distribution of six-month mean MAPE, MAE, and RMSE errors compared to SSEBop model output from the four DL models across all study areas. Each pair of consecutive rows displays the spatial variability of MAPE, MAE, and RMSE for a particular study area. In study areas 1 and 2, ASPAM-U-Net minimized extreme errors and had lower average metrics, while in study area 3, AG-U-Net was superior.

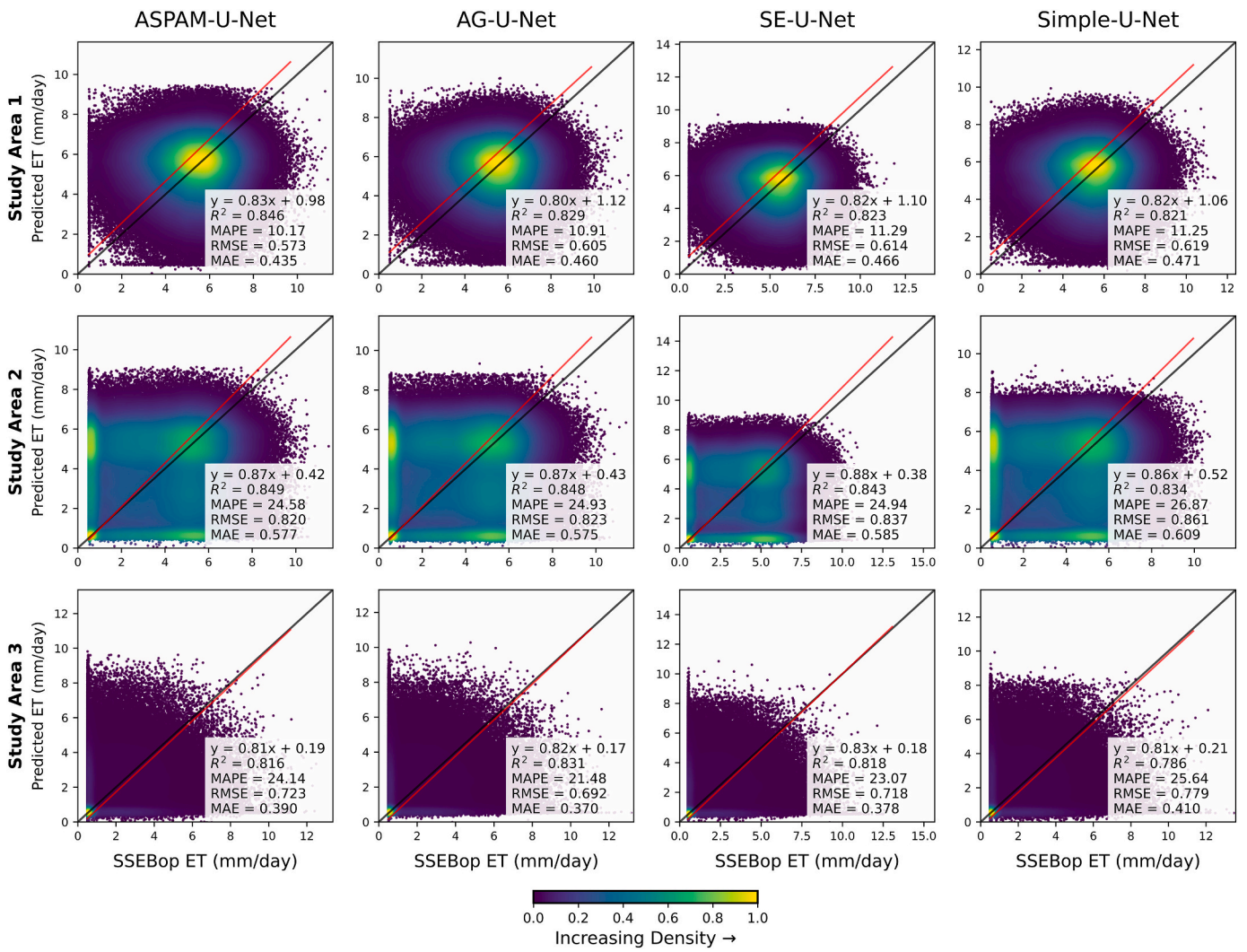


Fig. 9. Scatter plots, depicting the correlation between six-month mean values of estimated ET_a and SSEBop-ET_a obtained from four DL models across three study areas. Each row and column correspond to a specific study area and DL model, respectively. The black and red lines represent the diagonal line, and fitted linear line, respectively. The lower-right corner of each subplot presents the fitted line equation, R^2 values, MAPE (in percentage), RMSE (in mm/d) and MAE (in mm/d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

- Development of an AutoDL framework for large-scale ET_a estimation utilizing HHO and Python packages that automates the entire process from importing raw images to saving trained FCNN models: this is especially helpful because, despite the acceptable performance of energy-based and DL models, their development (e.g., architectural design and hyperparameter tuning in DL) requires expert knowledge. Efforts like the present study can facilitate the application of innovative DL methodologies in atmospheric and hydrology studies and contribute to the generalizability of the developed models.
- Presenting the ASPAM-U-Net as a novel U-Net model that incorporates an activated spatial attention mechanism based on matrix multiplication into the U-Net model as well as comparing its performance to other attention mechanisms (AG and SE) and the simple U-Net model: the ASPAM-U-Net model generally demonstrated improved accuracy in ET_a estimation through a more effective incorporation of spatial dependencies.
- Evaluation of the aforementioned models across three hydro-climatologically distinct regions throughout the year revealed their suitability for ET_a estimation, particularly during spring and summer when ET_a products exhibit lower levels of uncertainty.
- Analysis of the effect of various inputs on model performance through sensitivity analysis: this analysis showed that the derived vegetation and water indices and raw bands of the MODIS instrument played a more significant role compared to DEM data for ET_a mapping.

The main limitation of the current study, similar to all data-driven approaches, is the models' susceptibility to the uncertainty of the target ET_a parameter rather than the capabilities of developed AutoDL models. Consequently, training the models with more accurate target values derived from other more reliable ET_a mapping models is recommended. Although excluding meteorological inputs simplified the ET_a estimation, incorporation of such data in subsequent studies could potentially enhance the models' accuracy. Based on the promising results of the present study, higher-resolution ET_a mapping can be achieved by using high resolution satellites with thermal infrared data (e.g. Landsat) as inputs and incorporating new fine-resolution ET_a products, such as ECOSTRESS satellite data, for future studies. Similarly, the current 8-day temporal resolution of this approach can be further enhanced using data with higher temporal resolution.

Author contributions

AGB reviewed the existing literature, developed AutoDL models,

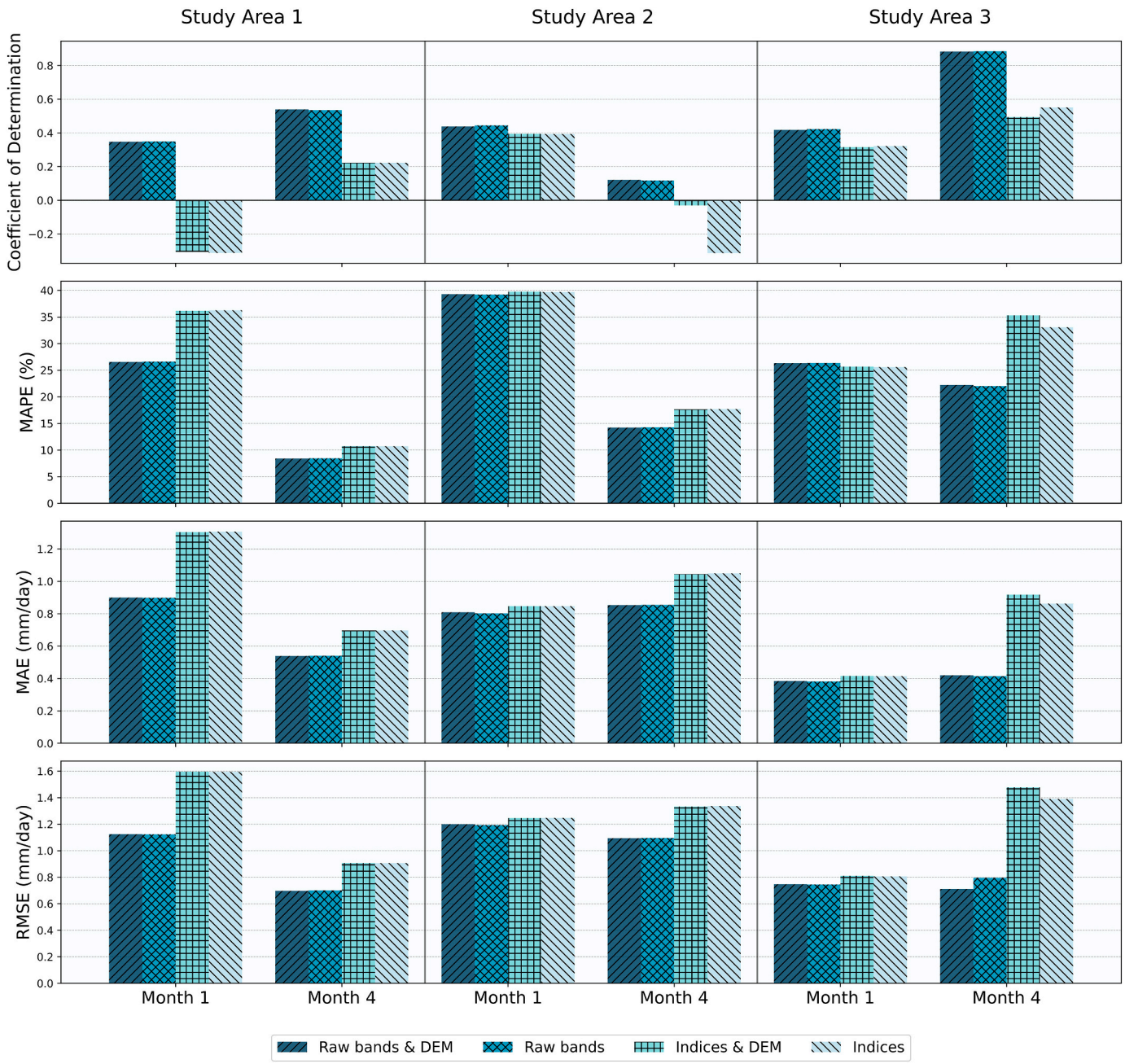


Fig. 10. Results of sensitivity analysis for four unique combinations of data groups across all three study areas during the first and fourth months of the year, obtained by ASPAM-U-Net. Adding noise to all groups of data resulted in reduction of accuracy in all evlaution metrics while second group of data exhibited higher sensitivity.

including the ASPAM attention mechanism, obtained and processed data, conceived and designed the methodology, generated the results, and compiled the first version of the article. A. Ahmadi and HK provided technical comments and assisted in writing the summary and abstract sections and formatting the article. A. Abbasi, SJ, and HH contributed to the supervision of AGB, assisted with the writing, improved subsequent revisions, and approved the final version of the manuscript. Additionally, SJ contributed to funding acquisitions.

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.

Acknowledgment

AGB was funded by the Erasmus student mobility program for a 3-month traineeship at Lund University, Sweden.

References

Abtew, W., Melesse, A., 2013. Evaporation and Evapotranspiration: Measurements and Estimations, pp. 1–206. https://doi.org/10.1007/978-94-007-4737-1.

Ahmadi, A., Daccache, A., Snyder, R.L., Suvočarev, K., 2022. Meteorological driving forces of reference evapotranspiration and their trends in California. Sci. Total Environ. 849, 157823 https://doi.org/10.1016/j.scitotenv.2022.157823.
Alizadeh, B., Ghaderi Bafti, A., Kamangir, H., Zhang, Y., Wright, D.B., Franz, K.J., 2021. A novel attention-based LSTM cell post-processor coupled with bayesian optimization for streamflow prediction. J. Hydrol. 601, 126526 https://doi.org/10.1016/j.jhydrol.2021.126526.
Amatya, D.M., Irmak, S., Gowda, P., Sun, G., Nettles, J.E., Douglas-Mankin, K.R., 2016. Ecosystem evapotranspiration: challenges in measurements, estimates, and modeling. Trans. ASABE 59 (2), 555–560. https://doi.org/10.13031/TRANS.59.11808.
Asgari Taghanaki, S., Abhishek, K., Cohen, J.P., Cohen-Adad, J., Hamarneh, G., 2021. Deep semantic segmentation of natural and medical images: a review. Artif. Intell. Rev. 54 (1), 137–178. https://doi.org/10.1007/s10462-020-09854-1.
Assouline, S., Mahrer, Y., 1993. Evaporation from Lake Kinneret: 1. Eddy correlation system measurements and energy budget estimates. Water Resour. Res. 29 (4), 901–910. https://doi.org/10.1029/92WR02432.
Babaeian, E., Paheding, S., Siddique, N., Devabhaktuni, V.K., Tuller, M., 2022. Short- and mid-term forecasts of actual evapotranspiration with deep learning. J. Hydrol. 612, 128078 https://doi.org/10.1016/j.jhydrol.2022.128078.
Bastiaanssen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998a. A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formul. J. Hydrol. 212–213, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4.
Bastiaanssen, W.G.M., et al., 1998b. A remote sensing surface energy balance algorithm for land (SEBAL): part 2: Validation. J. Hydrol. 212–213, 213–229. https://doi.org/10.1016/S0022-1694(98)00254-6.
Chen, M., et al., 2016. Uncertainty analysis of the operational simplified surface energy balance (SSEBop) model at multiple flux tower sites. J. Hydrol. 536, 384–399. https://doi.org/10.1016/j.jhydrol.2016.02.026.
Chen, I.T., et al., 2018. Exploring the spatio-temporal interrelation between groundwater and surface water by using the self-organizing maps. J. Hydrol. 556, 131–142. https://doi.org/10.1016/j.jhydrol.2017.10.015.
Chen, X., et al., 2020. The importance of short lag-time in the runoff forecasting model based on long short-term memory. J. Hydrol. 589, 125359 https://doi.org/10.1016/J.JHYDROL.2020.125359.
Cui, H., et al., 2021. Multiscale attention guided U-Net architecture for cardiac segmentation in short-axis MRI images. Comput. Methods Prog. Biomed. 206, 106142 https://doi.org/10.1016/j.cmpb.2021.106142.
Dayal, D., et al., 2021. Streamflow estimation using satellite-retrieved water fluxes and machine learning technique over monsoon-dominated catchments of India. Hydrol. Sci. J. 66 (4), 656–671. https://doi.org/10.1080/02626667.2021.1889557.
Ding, Y., et al., 2019. Spatio-Temporal attention LSTM Model for Flood Forecasting. iThings/GreenCom/CPSCom/SmartData 2019, 458-465. https://doi.org/10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00095.
Doulabian, S., Ghasemi Tousi, E., Aghlmand, R., Alizadeh, B., Ghaderi Bafti, A., Abbasi, A., 2021. Evaluation of integrating SWAT Model into a multi-criteria decision analysis towards reliable rainwater harvesting systems. Water 13, 1935. https://doi.org/10.3390/w13141935.
Fisher, J.B., et al., 2017. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resour. Res. 53 (4) https://doi.org/10.1002/2016wr020175.
Gebremedhin, M., et al., 2020. Evaluation of the Operational Simplified Surface Energy Balance Model for Pastureland Evapotranspiration Mapping and Drought monitoring in North Central Kentucky. Adv. Meteorol. 2020 https://doi.org/10.1155/2020/1386468.
Gowda, P.H., et al., 2008. ET mapping for agricultural water management: present status and challenges. Irrig. Sci. 26 (3), 223–237. https://doi.org/10.1007/S00271-007-0088-6.
Heidari, A.A., et al., 2019. Harris hawks optimization: Algorithm and applications. Futur. Gener. Comput. Syst. 97, 849–872. https://doi.org/10.1016/J.FUTURE.2019.02.028.
Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-Excitation Networks. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 7132–7141. https://doi.org/10.1109/CVPR.2018.00745.
Ji, L., et al., 2019. Evaluating the temperature difference parameter in the SSEBop model with satellite-observed land surface temperature data. Remote Sens. 11 (16) https://doi.org/10.3390/rs11161947.
John, D., Zhang, C., 2022. An attention-based U-Net for detecting deforestation within satellite sensor imagery. Int. J. Appl. Earth Obs. Geoinf. 107, 102685 https://doi.org/10.1016/j.jag.2022.102685.
Lenters, J.D., et al., 2005. Effects of climate variability on Lake evaporation: results from a long-term energy budget study of Sparkling Lake, northern Wisconsin (USA). J. Hydrol. 308, 168–195. https://doi.org/10.1016/J.JHYDROL.2004.10.028.
Li, Y., Huang, C., Hou, J., Gu, J., Zhu, G., Li, X., 2017. Mapping daily evapotranspiration based on spatiotemporal fusion of ASTER and MODIS images over irrigated agricultural areas in the Heihe River Basin, Northwest China. Agric. For. Meteorol. 244–245, 82–97. https://doi.org/10.1016/J.AGRFORMET.2017.05.023.
Li, X., Wei, Y., Wang, L., Fu, S., Wang, C., 2021a. MSGSE-Net: Multi-scale guided squeeze-and-excitation network for subcortical brain structure segmentation. Neurocomputing 461, 228–243. https://doi.org/10.1016/j.neucom.2021.07.018.
Li, H., et al., 2021b. Mapping salt marsh along coastal South Carolina using U-Net. ISPRS J. Photogramm. Remote Sens. 179, 121–132. https://doi.org/10.1016/j.isprsjprs.2021.07.011.
Li, X., Cheng, K., Huang, T., Qiu, Z., Tan, S., 2022a. Research on short term prediction method of thermal hydraulic transient operation parameters based on automated deep learning. Ann. Nucl. Energy 165, 108777. https://doi.org/10.1016/J.ANUCENE.2021.108777.
Li, Y., Wang, W., Wang, G., Tan, Q., 2022b. Actual evapotranspiration estimation over the Tuojiang River Basin based on a hybrid CNN-RF model. J. Hydrol. 610 (December 2021), 127788 https://doi.org/10.1016/j.jhydrol.2022.127788.
Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965.
Ma, J., Ding, Y., Cheng, J.C.P., Jiang, F., Gan, V.J.L., Xu, Z., 2020. A Lag-FLSTM deep learning network based on Bayesian Optimization for multi-sequential-variant PM2.5 prediction. Sustain. Cities Soc. 60 (May), 102237 https://doi.org/10.1016/j.scs.2020.102237.
Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P., 2017. Convolutional neural networks for large-scale remote-sensing image classification. IEEE Trans. Geosci. Remote Sens. 55 (2), 645–657. https://doi.org/10.1109/TGRS.2016.2612821.
McCulloch, J.S.G., 1976. The world water balance mean annual global, continental and maritime precipitation, evaporation and run-off: A. Baumgartner and E. Reichel. Oldenbourg Verlag GmbH, München, 1975. Agric. Water Manag. 1 (1), 100–101. https://doi.org/10.1016/0378-3774(76)90012-3.
Mementi, M., Choudhury, B.J., 2007. Parameterization of land surface evaporation by means of location dependent potential evaporation and surface temperature range. In: Exchange Processes at the Land Surface for a Range of Space and Time Scales, 212. IAHS Publ, pp. 561–568.
Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., Terzopoulos, D., 2021. Image segmentation using deep learning: a survey. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2021.3059968.
Moayedi, H., Mosavi, A., 2021a. Electrical power prediction through a combination of multilayer perceptron with water cycle ant lion and satin bowerbird searching optimizers. Sustainability 13 (4), 2336. https://doi.org/10.3390/SU13042336.
Moayedi, H., Mosavi, A., 2021b. Synthesizing multi-layer perceptron network with ant lion biogeography-based dragonfly algorithm evolutionary strategy invasive weed and league champion optimization hybrid algorithms in predicting heating load in residential buildings. Sustainability 13 (6), 3198. https://doi.org/10.3390/SU13063198.
Mu, Q., Zhao, M., Running, S.W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sens. Environ. https://doi.org/10.1016/j.rse.2011.02.019.
Polonskaia, I.S., Aliev, I.R., Nikitin, N.O., 2021. Automated evolutionary design of CNN classifiers for object recognition on satellite images. Procedia Comput. Sci. 193, 210–219. https://doi.org/10.1016/j.procs.2021.10.021.
Prasad, D.V.V., Venkataramana, L.Y., Kumar, P.S., Prasannamedha, G., Harshana, S., Srividya, S.J., Harrinei, K., Indraganti, S., 2022. Analysis and prediction of water quality using deep learning and auto deep learning techniques. Sci. Total Environ. 821, 153311 https://doi.org/10.1016/J.SCITOTENV.2022.153311.
Rabiei, S., Jalilvand, E., Tajrishy, M., 2021. A method to estimate surface soil moisture and map the irrigated cropland area using Sentinel-1 and Sentinel-2 data. Sustainability 13 (20), 11355. https://doi.org/10.3390/su132011355.
Rahimpour, M., Rahimzadegan, M., 2021. Assessment of surface energy balance algorithm for land and operational simplified surface energy balance algorithm over freshwater and saline water bodies in Urmia Lake Basin. Theor. Appl. Climatol. 143 (3–4), 1457–1472. https://doi.org/10.1007/s00704-020-03472-1.
Rundo, L., Han, C., Nagano, Y., Zhang, J., Hataya, R., Militello, C., Tangherloni, A., Nobile, M.S., Ferretti, C., Besozzi, D., Gilardi, M.C., Vitabile, S., Mauri, G., Nakayama, H., Cazzaniga, P., 2019. USE-Net: incorporating squeeze-and-excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets. Neurocomputing 365, 31–43. https://doi.org/10.1016/j.neucom.2019.07.006.
Sahraei, R., Ghorbanian, A., Kanani-Sadat, Y., Jamali, S., Homayouni, S., 2023. Mangrove plantation suitability mapping by integrating multi criteria decision making geospatial approach and remote sensing data. Geo-Spat. Inf. Sci. https://doi.org/10.1080/10095020.2023.2167615.
Sameen, M.I., Pradhan, B., Lee, S., 2020. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. CATENA 186, 104249. https://doi.org/10.1016/J.CATENA.2019.104249.
Schlemper, J., Oktay, O., Schaap, M., Heinrich, M., Kainz, B., Glocker, B., Rueckert, D., 2019. Attention gated networks: Learning to leverage salient regions in medical images. Med. Image Anal. 53, 197–207. https://doi.org/10.1016/j.media.2019.01.012.
Senay, G.B., 2018. Satellite psychrometric formulation of the operational simplified surface energy balance (Ssebop) model for quantifying and mapping evapotranspiration. Appl. Eng. Agric. 34 (3), 555–566. https://doi.org/10.13031/aea.12614.
Senay, G.B., Bohms, S., Singh, R.K., Gowda, P.H., Velpuri, N.M., Alemu, H., Verdin, J.P., 2013. Operational evapotranspiration mapping using remote sensing and weather datasets: a new parameterization for the SSEB approach. JAWRA J. Am. Water Resour. Assoc. 49 (3), 577–591. https://doi.org/10.1111/JAWR.12057.
Setiawan, I.N., Kurniawan, R., Yuniarto, B., 2021. Science DIRECT PARAMETER OPTIMIZATION OF SUPPORT VECTOR REGRESSION using Harris Hawks Optimization. Procedia Comput. Sci. 179 (2020), 17–24. https://doi.org/10.1016/j.procs.2020.12.003.
Singh, R.K., Senay, G.B., 2015. Comparison of four different energy balance models for estimating evapotranspiration in the Midwestern United States. Water 8 (1), 9. https://doi.org/10.3390/w8010009.
Stoian, A., Poulain, V., Inglada, J., Poughon, V., Derksen, D., 2019. Land cover maps production with high resolution satellite image Time series and convolutional neural

networks: adaptations and limits for operational systems. Remote Sens. 11 (17), 1986. https://doi.org/10.3390/RS11171986.
Su, Z., 2002. The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. Hydrol. Earth Syst. Sci. 6 (1), 85–99. https://doi.org/10.5194/HESS-6-85-2002.
Su, T., Xie, D., Feng, T., Huang, B., Qian, Z., Feng, G., Wu, Y., 2022. Quantifying the contribution of terrestrial water storage to actual evapotranspiration trends by the extended Budyko model in Northwest China. Atmos. Res. 273 (February), 106147 https://doi.org/10.1016/j.atmosres.2022.106147.
Taccari, M.L., Nuttall, J., Chen, X., Wang, H., Minnema, B., Jimack, P.K., 2022. Attention U-Net as a surrogate model for groundwater prediction. Adv. Water Resour. 104169 https://doi.org/10.1016/J.ADVWATRES.2022.104169.
Tang, R., Li, Z.L., Tang, B., 2010. An application of the Ts–VI triangle method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: Implementation and validation. Remote Sens. Environ. 114 (3), 540–551. https://doi.org/10.1016/J.RSE.2009.10.012.
Tanny, J., Cohen, S., Assouline, S., Lange, F., Grava, A., Berger, D., Teltch, B., Parlange, M.B., 2008. Evaporation from a small water reservoir: direct measurements and estimates. J. Hydrol. 351 (1–2), 218–229. https://doi.org/10.1016/J.JHYDROL.2007.12.012.
Tran, P.V., 2016. A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis MRI. https://doi.org/10.48550/arxiv.1604.00494.
Wang, W., Xiao, W., Cao, C., Gao, Z., Hu, Z., Liu, S., Shen, S., Wang, L., Xiao, Q., Xu, J., Yang, D., Lee, X., 2014. Temporal and spatial variations in radiation and energy balance across a large freshwater lake in China. J. Hydrol. 511, 811–824. https://doi.org/10.1016/J.JHYDROL.2014.02.012.
Wang, J., Lv, P., Wang, H., Shi, C., 2021. SAR-U-Net: Squeeze-and-excitation block and atrous spatial pyramid pooling based residual U-Net for automatic liver segmentation in Computed Tomography. Comput. Methods Prog. Biomed. 208, 106268 https://doi.org/10.1016/j.cmpb.2021.106268.
Wang, X., Zhong, L., Ma, Y., Fu, Y., Han, C., Li, P., Wang, Z., Qi, Y., 2023. Estimation of hourly actual evapotranspiration over the Tibetan Plateau from multi-source data. Atmos. Res. 281, 106475 https://doi.org/10.1016/j.atmosres.2022.106475.
Weerasinghe, I., Bastiaanssen, W., Mul, M., Jia, L., Van Griensven, A., 2020. Can we trust remote sensing evapotranspiration products over Africa. Hydrol. Earth Syst. Sci. 24 (3), 1565–1586. https://doi.org/10.5194/HESS-24-1565-2020.
Winter, T.C., Buso, D.C., Rosenberry, D.O., Likens, G.E., Sturrock, A.M., Mau, D.P., 2003. Evaporation determined by the energy-budget method for Mirror Lake, New Hampshire. Limnol. Oceanogr. 48 (3), 995–1009. https://doi.org/10.4319/LO.2003.48.3.0995.
Xu, T., Guo, Z., Xia, Y., Ferreira, V.G., Liu, S., Wang, K., Yao, Y., Zhang, X., Zhao, C., 2019. Evaluation of twelve evapotranspiration products from machine learning, remote sensing and land surface models over conterminous United States. J. Hydrol. 578, 124105 https://doi.org/10.1016/j.jhydrol.2019.124105.
Xu, Z., Wang, S., Stanislawski, L.V., Jiang, Z., Jaroenchai, N., Sainju, A.M., Shavers, E., Usery, E.L., Chen, L., Li, Z., Su, B., 2021. An attention U-Net model for detection of fine-scale hydrologic streamlines. Environ. Model. Softw. 140, 104992 https://doi.org/10.1016/j.envsoft.2021.104992.
Yan, X., Yang, N., Ao, R., Mohammadian, A., Liu, J., Cao, H., Yin, P., 2023. Deep learning for daily potential evapotranspiration using a HS-LSTM approach. Atmos. Res. 292 (June), 106856 https://doi.org/10.1016/j.atmosres.2023.106856.
Yang, R., Ahmed, Z.U., Schulthess, U.C., Kamal, M., Rai, R., 2020. Detecting functional field units from satellite images in smallholder farming systems using a deep learning based computer vision approach: a case study from Bangladesh. Remote Sens. Appl. Soc. Environ. 20, 100413 https://doi.org/10.1016/j.rsase.2020.100413.
Zhuang, Q., Shi, Y., Shao, H., Zhao, G., Chen, D., 2021. Evaluating the ssebop and rspmpt models for irrigated fields daily evapotranspiration mapping with modis and cmads data. Agriculture 11 (5). https://doi.org/10.3390/agriculture11050424.