remote sensing

MDPI

Article

A Transfer Learning Approach for Estimating All-Weather Daily Net Radiation over the Tibetan Plateau: Site-Scale Evaluation and Spatial Extension

Lingjie Liu [1], Yan Li [1,*], Lin Zhao [1], Jinliang Hou [2], Lingxiao Wang [1] and Guojie Hu [3]

[1] School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China; 202312110009@nuist.edu.cn (L.L.); lzhao@nuist.edu.cn (L.Z.); lx.wang@nuist.edu.cn (L.W.)
[2] Heihe Remote Sensing Experimental Research Station, State Key Laboratory of Cryospheric Science and Frozen Soil Engineering, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; jlhours@lzb.ac.cn
[3] Cryosphere Research Station on the Qinghai-Tibet Plateau, State Key Laboratory of Cryospheric Science and Frozen Soil Engineering, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; huguojie123@lzb.ac.cn
* Correspondence: yan.li@nuist.edu.cn

Highlights

What are the main findings?

- The DenseMLP-based transfer learning framework shows promising performance in generating 1 km daily net radiation over the Tibetan Plateau.
- The method maintains low site-specific bias and captures seasonal variability well.

What are the implications of the main findings?

- The 1 km daily net radiation product provides a high-resolution data source for land surface energy budget analysis over the Tibetan Plateau.
- The framework offers a pathway for long-term, high-resolution daily net radiation product development in data-sparse regions.

Abstract

Accurate estimation of daily net radiation (R_{n_daily}) at high spatial resolution (1 km) over the Tibetan Plateau (TP) is crucial for understanding land surface energy budgets and climate dynamics. This study proposes a densely connected multilayer perceptron (DenseMLP)-based transfer learning framework, with a two-stage strategy (coarse pre-training on GLASS R_{n_daily}, followed by fine-tuning on limited TP ground observations) using MODIS land surface parameters and auxiliary data to generate 1 km R_{n_daily}. When evaluated on the training set, the proposed model achieves an overall R^2 of 0.87, MAE of 16.06 W m^{-2}, RMSE of 21.94 W m^{-2}, and a near-zero bias of 0.07 W m^{-2}. On an independent test set, the model maintains robust performance with R^2 = 0.83, MAE = 17.43 W m^{-2}, RMSE = 22.55 W m^{-2}, and bias = −1.12 W m^{-2}. The method exhibits consistently low bias across individual sites (mostly within ±3.7 W m^{-2}) and accurately captures seasonal variability. When applied to the entire TP for 2018, the 1 km R_{n_daily} product reveals clear aspect-related terrain effects and a distinct annual cycle. This framework effectively mitigates site-dependent errors, providing a useful reference for long-term R_n product development over the TP.

Keywords: daily net radiation; transfer learning; Tibetan Plateau

Check for updates

Academic Editor: Dimitris Kaskaoutis

Received: 11 June 2026
Revised: 22 June 2026
Accepted: 26 June 2026
Published: 29 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Surface net radiation (R_n) serves as the core energy driver underpinning all biophysical transfer processes that structure and sustain terrestrial ecosystems, including latent and sensible heat convection; heat conduction through soil, water, vegetation, and atmospheric air masses; and endergonic photoreactions mediated by chlorophyll-containing plants [1–4]. This fundamental energy control is particularly pronounced on the Tibetan Plateau (TP), a region characterized by its high altitude and unique land–atmosphere interactions [5]. The plateau's elevated heating effect, driven by the transformation of surface R_n into sensible heat flux, serves as a critical heat source that regulates the formation and evolution of the Asian monsoon system [6]. At the local scale, the partitioning of R_n into latent, sensible, and ground heat fluxes directly governs the water use efficiency of alpine ecosystems [7] and the thermal stability of underlying permafrost [8]. Accordingly, accurate quantification and reliable spatial estimation of surface R_n are essential for unraveling the complex land–atmosphere–cryosphere interactions.

Ground-based observations of surface radiation, typically obtained from instruments such as four-component radiometers, provide the most direct and accurate measurements of R_n by quantifying individual radiation fluxes (i.e., downwelling and upwelling components of both shortwave and longwave radiation) that comprise the surface radiation balance [9]. However, the installation and maintenance of radiometers are time-consuming and expensive, resulting in extremely sparse observational networks in most regions. This scarcity is further exacerbated on the TP, where harsh environmental conditions and logistical difficulties make sustained radiation measurements particularly challenging to obtain [10]. These observational gaps thus call for alternative approaches that leverage remote sensing and meteorological data to estimate R_n at regional scales over the TP and beyond [11–15].

Reanalysis products provide continuous gridded radiation data by assimilating multi-source observations, yet they are characterized by coarse spatial resolution (>5 km) and exhibit substantial biases in complex topographic regions, such as the TP [16]. To obtain accurate R_n estimates at higher spatial resolution, various estimation methods based on remote sensing data have been developed, which can be broadly categorized into physical and data-driven approaches. Although physically based radiative transfer models (RTMs) can capture the spatial heterogeneity of R_n with relatively complete coverage, they are predominantly applied at satellite overpass times, and their extension to daily scales remains challenging [17,18]. In contrast, data-driven models, especially those using machine learning (ML), have become increasingly popular, as they can effectively characterize the complex nonlinear relationships between R_n and various satellite or meteorological variables [11,12,19–21]. Recent studies further show that ML methods enable direct estimation of regional daily R_n (R_{n_daily}) from satellite observations, avoiding the need for intermediate radiative transfer modeling [12,19,22]. Nevertheless, the majority of existing machine learning models still face substantial challenges in practical applications over regions with sparse ground observations and complex terrain, such as the TP.

In the specific application scenario of regional R_{n_daily} estimation, transfer learning (TL) serves as a natural extension of ML modeling under the conditions of limited local samples and weak labels. This technique allows the model to integrate prior knowledge extracted from external datasets, thereby constructing robust predictive models even when local ground observations are extremely limited [23]. For example, Miao et al. [24] developed a multi-scale TL framework in the Heihe River Basin to estimate R_{n_daily} at a 1 km spatial resolution using limited ground-based observations. In another relevant study, Li et al. [21] constructed a ResNet-based TL model, using data augmentation in the contiguous United States, and transferred it to data-scarce regions in mainland China and tropical areas for

regional R_{n_daily} estimation at a 9 km spatial resolution. Their results showed that the TL framework achieved significantly higher estimation accuracy in the target regions compared to a directly trained model.

The aforementioned studies fully demonstrate that transfer learning can effectively alleviate the constraints imposed by sparse ground observations on model training. However, existing frameworks predominantly employ convolutional networks, such as ResNet, in a patch-to-value regression manner to capture spatial features from remote sensing imagery. In practice, Li et al. [21] adopted this patch-to-value regression design and produced coarse-resolution R_n outputs at 9 km, falling short of the requirements for fine-scale applications. Miao et al. [24] also achieved 1 km daily R_n estimation using a convolutional DenseNet with two-stage transfer learning, yet this convolution-based architecture comes with inherent spatial inductive bias and may show fluctuating fine-tuning performance under scarce in situ data. To address this, this paper proposes a TL framework based on a densely connected multilayer perceptron (DenseMLP), adopting a two-stage strategy of "coarse-resolution pre-training and fine-resolution fine-tuning". Our fully connected DenseMLP serves as an alternative architecture that circumvents the spatial inductive bias inherent to convolutional structures, delivering improved cross-resolution transferability and greater fine-tuning robustness with limited ground observations, while maintaining competitive regression performance for multivariate daily R_n estimation. Specifically, a foundational model is first pre-trained using R_{n_daily} obtained from global land surface satellite (GLASS) products with a relatively coarse spatial resolution (0.05°), and subsequently fine-tuned with limited ground-based observations from the TP. The model inputs include land surface parameters (e.g., reflectance and LST) from a Moderate Resolution Imaging Spectroradiometer (MODIS) at 1 km resolution, along with other auxiliary radiative, meteorological, and topographic data. This enables the estimation of R_{n_daily} at a 1 km resolution, providing methodological support for high-resolution net radiation modeling in data-scarce regions.

2. Materials and Methods

2.1. Materials

2.1.1. Study Area

The Tibetan Plateau (TP) spans approximately 2.5 million square kilometers across Southwestern China with average elevation exceeding 4000 m above sea level (a.s.l.) (Figure 1). This extreme high-elevation topography creates a unique atmospheric environment and directly drives the formation of distinct cryospheric features. The plateau's high-altitude continental climate is marked by pronounced spatial variability, with widespread cold, aridity, and strong winds. Mean annual air temperatures, for instance, range from below -5 °C in the central high-elevation core to near 0 °C in the southern and eastern peripheries, while large diurnal temperature fluctuations and short, cool summers prevail across the region [25]. Combined with the foregoing climatic conditions, permafrost is the dominant subsurface cryospheric feature on the TP. The plateau hosts extensive permafrost, covering approximately 1.06×10^6 km^2, which accounts for about 40% of its total land area [26]. This extensive distribution renders the TP as the largest and highest-altitude permafrost region in the mid-to-low latitudes of the globe. Given the sparse distribution of ground observation stations and the complex terrain, the plateau represents a typical data-scarce region, making it an ideal testbed for transfer learning-based high-resolution environmental parameter estimation.

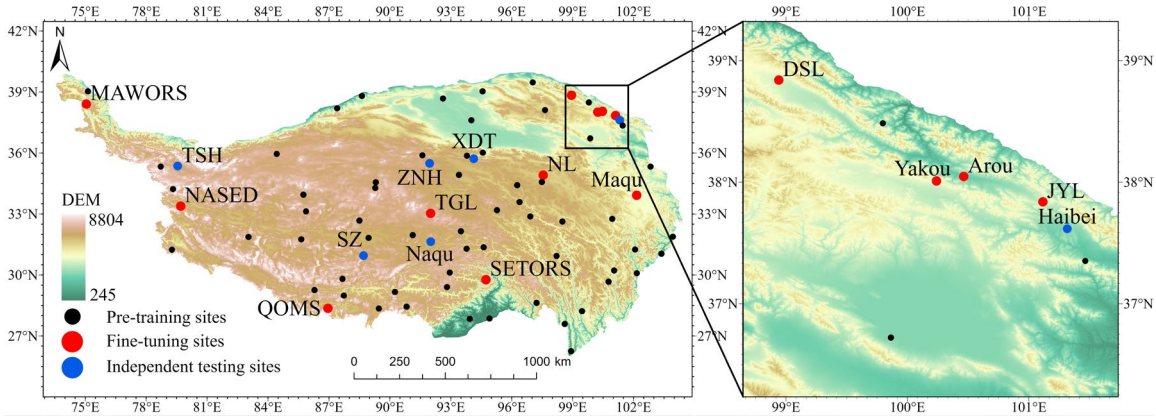


Figure 1. Map of the study area across the Tibetan Plateau, showing sites for model training and independent testing, with topography based on a digital elevation model (DEM).

2.1.2. In Situ Observations

This study utilizes in situ observations from seventeen field stations across the TP (Figure 1, Table 1). Eleven stations are used for model training, and the remaining six independent stations are reserved exclusively for out-of-site independent testing. Training stations (Arou, DSL, JYL, Maqu, MAWORS, NASED, NL, QOMS, SETORS, TGL, Yakou) span elevations of 3033 m to 5100 m a.s.l., dominated by alpine meadow, alpine steppe, and alpine desert, with data compiled from the long-term quality-assured land–atmosphere interaction datasets of the Tibetan Plateau [27]. We note that the Maqu dataset covers only 2019. The six independent testing stations (Haibei, Naqu, SZ, TSH, XDT, ZNH) range from 3199 m to 4838 m a.s.l. across alpine meadow, alpine grassland, and alpine desert, with all observations limited to 2018. Specifically, data at Naqu are obtained from the China terrestrial ecosystem flux research network (ChinaFLUX; http://www.chinaflux.org/) at a half-hourly resolution. Haibei and SZ are derived from meteorological data of the surface environment and observation network in China's cold region (2018), with daily-scale data [28]. TSH, XDT, and ZNH are derived from the integrated monitoring dataset of permafrost on the Tibetan Plateau (IMDP) and only daily-scale data are available for these sites [29]. Daily mean net radiation was calculated from four-component radiation flux observations at the corresponding stations. Overall, the training stations yielded 5717 valid daily records (2019–2020), and the six testing stations yielded 2157 valid daily records (2018).

Table 1. Basic information of the observation sites across the Tibetan Plateau.

Site	Longitude	Latitude	Altitude	Surface Type	Temporal Resolution	Observation Period
Arou	100.4°E	38.04°N	3033 m	Alpine meadow	1 h	2019–2020
DSL	98.94°E	38.84°N	3739 m	Alpine meadow	1 h	2019–2020
JYL	101.11°E	37.83°N	3750 m	Alpine meadow	1 h	2019–2020
Maqu	102.15°E	33.92°N	3434 m	Alpine meadow	30 min	2019
MAWORS	75.0°E	38.41°N	3668 m	Alpine desert	1 h	2019–2020
NASED	79.70°E	33.39°N	4270 m	Alpine desert	1 h	2019–2020
NL	97.55°E	34.91°N	4280 m	Alpine meadow	30 min	2019–2020
QOMS	86.95°E	28.36°N	4298 m	Alpine desert	1 h	2019–2020
SETORS	94.73°E	29.77°N	3327 m	Alpine meadow	1 h	2019–2020
TGL	92.00°E	33.03°N	5100 m	Alpine meadow	30 min	2019–2020
Yakou	100.24°E	38.01°N	4148 m	Alpine meadow	1 h	2019–2020
Haibei	101.31°E	37.61°N	3199 m	Alpine meadow	1 d	2018

Table 1. *Cont.*

Site	Longitude	Latitude	Altitude	Surface Type	Temporal Resolution	Observation Period
Naqu	92.01°E	31.64°N	4585 m	Alpine meadow	30 min	2018
SZ	88.70°E	35.95°N	4675 m	Alpine grassland	1 d	2018
TSH	79.55°E	35.36°N	4838 m	Alpine desert	1 d	2018
XDT	94.13°E	35.72°N	4456 m	Alpine meadow	1 d	2018
ZNH	91.96°E	35.49°N	4775 m	Alpine desert	1 d	2018

2.1.3. Regional Data

For pre-training the DenseMLP_TL model, R_{n_daily} retrieved from 0.05° GLASS products [30] was adopted as the pre-training target due to its reliable spatiotemporal patterns over the TP [11,22], as well as its superior accuracy against other mainstream R_n products at the global scale [22]. To estimate R_{n_daily}, a multi-source input feature space (Table 2) is constructed for the DenseMLP_TL model, which includes all-weather reflectance in optical bands (covering visible, near-infrared, and shortwave infrared wavelengths) along with downwelling shortwave radiation (Srad) and topographic factors (elevation, slope, and aspect) to characterize shortwave processes, as well as all-weather land surface temperature (LST), air temperature (Ta), and downwelling longwave radiation (Lrad) to capture longwave radiation components. Specifically, the optical band reflectance (b01–b07) is derived from the MOD09GA product (MODIS/Terra Surface Reflectance Daily L2G Global 500 m SIN Grid), which provides daily surface reflectance at a spatial resolution of 500 m. The digital elevation data are acquired from the Multi-Error-Removed Improved-Terrain DEM (MERIT DEM) at a spatial resolution of 3″, from which slope and aspect were subsequently derived. Srad, Lrad, and Ta are obtained from the TPMFD dataset (a high-resolution, near-surface meteorological forcing dataset for the Third Pole region, 1979–2023) [25], which features hourly, daily, and monthly temporal resolution at 1/30° spatial resolution. LST is sourced from the daily 1 km all-weather land surface temperature dataset for Western China (TRIMS LST-TP V2), with a spatial resolution of 1 km and a temporal coverage of 2000–2024 [31]. All data were unified to a daily temporal resolution and spatially aligned to a consistent 1 km grid under the geographic coordinate system (latitude/longitude), with resampling strategies tailored to each dataset's native resolution. Specifically, coarse GLASS R_n and TPMFD radiation/meteorological data were bilinearly interpolated to 1 km, while finer MOD09GA reflectance and MERIT DEM were aggregated via area averaging.

Table 2. Multi-source data used for daily net radiation estimation.

Dataset	Physical Variables	Temporal Resolution	Spatial Resolution	Period
GLASS	Daily net radiation	Daily	0.05°	2018–2020
MOD09GA	Optical band reflectance	Daily	500 m	2018–2020
TPMFD	Downwelling shortwave radiation, Downwelling longwave radiation, Air temperature	Daily	1/30°	2018–2020
TRIMS	Land surface temperature	Daily	1 km	2018–2020
MERIT DEM	Elevation, slope, aspect	N/A	3″	2018–2020

To align with the site-scale evaluation and spatial extension objectives of this study, we first performed pre-training using 60 spatially representative static sites (Figure 1) rather than daily random samples, which were carefully selected via a multi-step terrain-stratified sampling strategy. Initially, 300 candidate points were randomly generated

within the TP boundary. We then extracted core topographic variables (DEM, slope, and aspect) for all candidates and established integrated terrain stratification, where DEM was divided into four quantile classes, slope was divided into three quantile classes, and aspect was categorized into four directional bins to form comprehensive "DEM–slope–aspect " combination strata. Based on the maximum–minimum distance method, we finally screened 60 sampling points by guaranteeing at least one site for each non-empty stratum and proportionally allocating the remaining samples across all strata. Furthermore, we validated the land cover representativeness of the selected points using the TP land cover dataset. The final 60 points cover the major TP land cover types, among which alpine meadow and alpine steppe account for 66.7% of all samples, highly consistent with the regional areal proportion of 61.2%. For model training, daily time series of 1 km input variables from 2018 to 2020 were extracted at each site.

The model was then fine-tuned using all pixel-matched observations collected from the eleven in situ stations during 2019–2020 as the target variable, with six independent unseen stations (all from 2018) retained for later independent testing. This design enforces strict spatiotemporal separation of training and testing data; the model trains on a unique set of stations for 2019–2020 and is assessed on separate stations in an entirely unused year (2018), supporting rigorous testing of spatial and temporal generalization. Finally, the trained DenseMLP_TL model was applied to the 2018 regional data to estimate the 1 km all-weather R_{n_daily} across the TP for the year 2018.

2.2. Methods

2.2.1. The DenseMLP Network

The DenseMLP network constructed in this study adopts the concept of cross-layer dense connections from densely connected networks (DenseNets) [32], overcoming the limitation of traditional MLPs that rely solely on layer-by-layer unidirectional propagation. Specifically, a cross-layer concatenation strategy is employed for fully connected layers within each dense block. Within a dense block, the input of any layer is formed by concatenating the output features of all preceding layers in that block along the channel dimension, thereby enabling feature reuse and efficient forward propagation. During back-propagation, gradients can propagate across layers, alleviating the vanishing or dissipating gradient problem in deep networks and enhancing the model's ability to represent strongly nonlinear variables.

The model input layer consists of multi-dimensional features that have undergone spatial matching and normalization preprocessing. After all input features are concatenated along the dimension, they are fed into the backbone network, which is composed of alternating dense blocks and transition layers. Each dense block contains three fully connected layers, with inter-layer connections following the dense connectivity rule, which can be mathematically expressed as,

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \quad (1)$$

where $[x_0, x_1, \ldots, x_{l-1}]$ denotes the concatenation of output features from all preceding layers; x_l represents the output of the l-th layer; and H_l denotes a nonlinear transformation comprising a fully connected layer, batch normalization, and Rectified Linear Unit (ReLU) activation function. Within a dense block, the cumulative feature dimension increases linearly with the preset growth rate w, thereby enriching feature representation. A transition layer is placed between adjacent dense blocks, which reduces the dimensionality through a linear fully connected layer, with the compression rate θ controlling the degree of reduction, thereby balancing feature reuse and model complexity. In this framework, the hyperparameters w and the number of blocks (d) represent the width and depth of

model, dominating the feature learning capacity and model complexity. Larger w strengthens feature extraction, whereas deeper d promotes multi-level feature fusion, collectively governing model accuracy, parameter volume, and training efficiency.

Following the dense connectivity rule in Equation (1), the backbone network comprising alternating dense blocks and transition layers extracts the terminal feature X. At the end of the backbone network, a fully connected layer with 32 neurons integrates the features, followed by a dropout layer to suppress overfitting, and finally a single-neuron linear output layer produces the estimation result, corresponding to the regression equation,

$$R_{n_daily} = \sigma(W \cdot X + b) \tag{2}$$

where W is the weight matrix, b is the bias term, and σ is the linear activation function.

2.2.2. Transfer Learning Strategy

Transfer learning is a machine learning paradigm that leverages existing knowledge to address similar new problems. Its core idea lies in capturing the shared structural information between the source and target domains, thereby reducing the reliance of the target task on large-scale samples and improving both learning efficiency and model generalization [33]. In the TP, where radiation observation stations are sparse and available samples are limited, training a model from scratch using only target-domain samples often leads to poor stability and weak generalization. Therefore, it is necessary to leverage prior knowledge from large-sample source domains to provide parameter initialization and knowledge transfer support for R_{n_daily} estimation in the target region.

The domain and task are core concepts in transfer learning. A domain D consists of a feature space and the marginal probability of the features, defined as,

$$\mathrm{D} = \{\mathcal{X}, P(X)\} \tag{3}$$

where $\mathcal{X}$ denotes the multi-source input feature space in this study; and $P(X)$ represents the marginal distribution of features under specific data sources, spatial resolutions, and underlying surface conditions, reflecting the overall distribution characteristics of the data. A task T is defined by the label space and the prediction function, expressed as,

$$\mathrm{T} = \{\mathcal{Y}, f(\cdot)\} \tag{4}$$

In this study, $\mathcal{Y} \subseteq R$ is the value space of R_{n_daily}, and $f(\cdot)$ is the DenseMLP regression mapping function to be trained. The core objective of the model is to achieve an accurate fit between the predicted value $\hat{y}_i = f(x_i)$ and the true observed value y_i, ensuring estimation accuracy.

Based on the above definitions, the source domain and target domain can be expressed as $D_s = \{\mathcal{X}_s, P_s(X)\}$ and $D_T = \{\mathcal{X}_T, P_T(X)\}$, respectively. In this case, D_s is a coarse-resolution sample set labeled with R_{n_daily} derived from GLASS products at spatial resolution of 0.05°, while D_T is a high-resolution region with target spatial resolution of 1 km, constrained by ground observations on the TP. After preprocessing, including spatial resampling to align the feature spaces, the two domains share homogeneous features (i.e., $\mathcal{X}_s \approx \mathcal{X}_T$), constituting a cross-domain regression problem. However, their label spaces remain heterogeneous ($\mathcal{Y}_s \neq \mathcal{Y}_T$) due to differences in data sources and spatial scales.

Differences in data sources, spatial resolutions, and underlying surfaces leads to $P_s(X) \neq P_T(X)$, indicating a clear domain shift. Such distributional discrepancies directly affect model generalization, leading to low estimation accuracy across the TP when models

are trained directly. Therefore, it is necessary to mitigate inter-domain discrepancies through transfer learning to optimize model adaptability.

This study adopts a DenseMLP-based transfer learning strategy, denoted as DenseMLP_TL, which consists of pre-training followed by supervised fine-tuning (Figure 2). First, the Dense-MLP model is pre-trained using source-domain samples $\left(x_s^i, y_s^i\right)$ by minimizing the mean squared error loss (L_s) in the source domain,

$$L_s = \frac{1}{N_s}\sum_{i=1}^{N_s}\left(\hat{y}_s^i - y_s^i\right)^2 \quad (5)$$

where N_s is the number of samples in the source domain, and $\hat{y}_s^i$ is the predicted value from pre-training. Subsequently, the model parameters are fine-tuned using a small number of site samples from the target domain, with only the final dense block and the output head being updated to adapt to the observational characteristics of the plateau, ultimately achieving full-area estimation of R_{n_daily} at a 1 km resolution. This approach fully leverages prior knowledge from the source domain to address the challenge of modeling with limited samples.

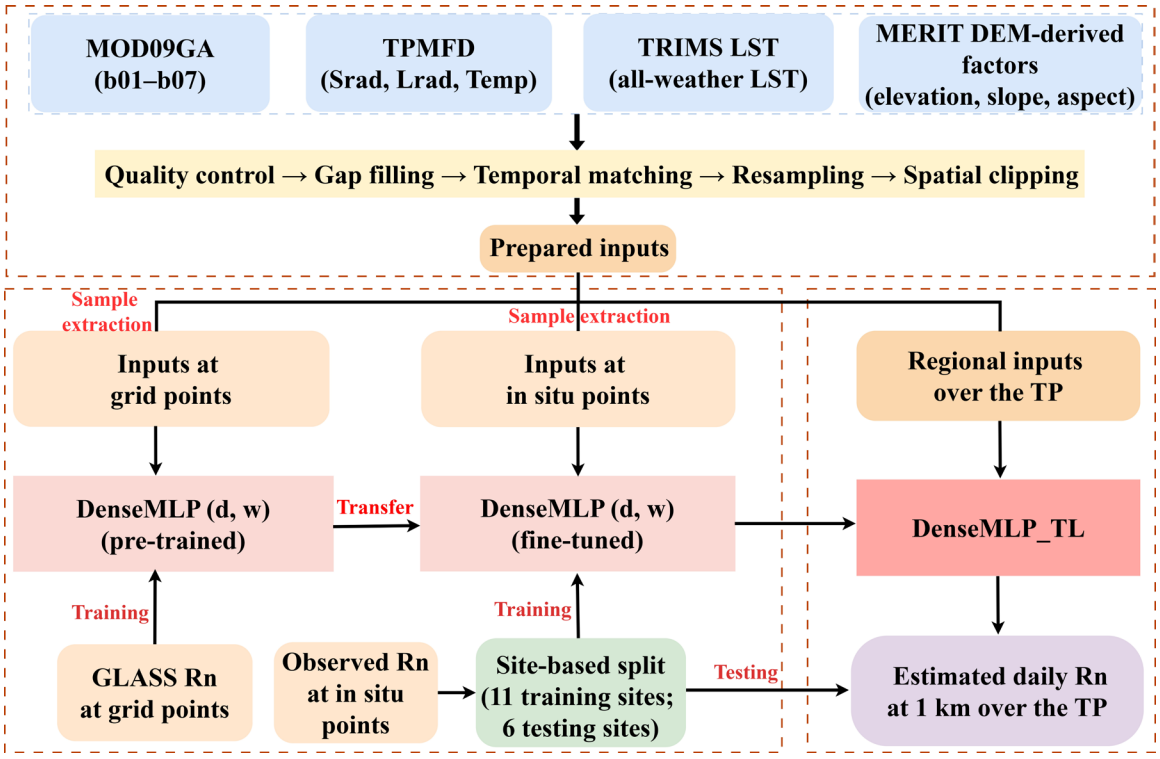


Figure 2. The workflow of DenseMLP-based transfer learning framework for the daily net radiation estimating.

In this work, the 0.05° GLASS product is adopted as the source-domain data for model pre-training. As reported in Ma et al. [34], this coarse-resolution dataset exhibits site-specific biases across the TP, but such biases do not undermine our transfer learning framework. The pre-training process focuses on large-scale spatial characteristics and universal physical relationships between R_{n_daily} and its driving factors rather than reproducing site-level biased values, and the following fine-tuning with field observations can effectively remove these biases. Meanwhile, GLASS R_{n_daily} has proven to be more accurate than many commonly used R_n products [22], including Clouds and the Earth's Radiant Energy System Edition 4A (CERES4A), the fifth generation of European Centre for Medium-Range Weather Forecasts Reanalysis (ERA5), and FLUXCOM_RS, and presents consistent spatial patterns

with other datasets over the TP [11,22]. For these reasons, GLASS remains a suitable choice for model pre-training.

3. Results

3.1. Site-Scale Evaluation

To evaluate the model performance, we conducted a direct evaluation by comparing the estimated R_{n_daily} with observations from seventeen in situ stations (eleven stations for model training and six stations for independent testing). In addition, a cross-validation was performed using the GLASS R_{n_daily} as a benchmark, where the discrepancies between observations and GLASS were compared against those between observations and our estimates. The scatter plots with error statistics for the training set and the independent test set are shown in Figures 3 and 4, respectively. From the scatter plots, the R_{n_daily} estimates from DenseMLP_TL are in close agreement with the observations across the seventeen stations for both the training and test sets. The points are tightly clustered around the 1:1 line, with no significant systematic bias. In contrast, the GLASS product exhibits greater scatter and notable discrepancies. Specifically, GLASS shows clear underestimations at MAWORS, NL, SETORS, TGL, Naqu, and ZNH, while overestimations are evident at JYL, QOMS, and Yakou. Such site-specific biases are not observed in our estimates.

In terms of error metrics, the performance of DenseMLP_TL and GLASS on the training set is as follows (also see the corresponding scatter plot). DenseMLP_TL achieves an overall R^2 of 0.87, with MAE of 16.06 W m^{-2}, RMSE of 21.94 W m^{-2}, and a near-zero bias of 0.07 W m^{-2}. In contrast, the GLASS product performs far worse, with an overall R^2 of just 0.70, MAE of 26.03 W m^{-2}, RMSE of 34.26 W m^{-2}, and a large negative bias of −9.30 W m^{-2}. At individual sites, the performance gap is even more striking. For example, at SETORS station, DenseMLP_TL achieves a training R^2 of 0.68 and a bias of 0.50 W m^{-2}, whereas GLASS, despite yielding an R^2 of 0.68, shows an extreme bias of −36.71 W m^{-2}. At JYL station, the model delivers an R^2 of 0.89, MAE of 22.43 W m^{-2}, and RMSE of 31.87 W m^{-2}, compared to GLASS's R^2 of 0.79, MAE of 43.02 W m^{-2}, and RMSE of 53.89 W m^{-2}. Similarly, at NL station, DenseMLP_TL maintains a low bias of 0.39 W m^{-2}, while GLASS shows a large negative bias of −27.45 W m^{-2}, with both MAE and RMSE values substantially higher than those of the proposed model.

For the independent test set, the model maintains robust and stable performance, with an overall R^2 of 0.83, MAE of 17.43 W m^{-2}, RMSE of 22.55 W m^{-2}, and a near-zero bias of −1.12 W m^{-2}. This represents only a slight drop from the training set, suggesting a reduced risk of overfitting to the training data. By direct comparison, GLASS yields substantially poorer performance in our evaluation, with an overall R^2 of only 0.81, MAE of 20.62 W m^{-2}, RMSE of 26.27 W m^{-2}, and a notable negative bias of −11.65 W m^{-2}. Across individual sites, GLASS exhibits drastically fluctuating bias—consistently negative but varying widely in magnitude—from near-zero at HB (−1.27 W m^{-2}) to severe underestimation at Naqu (−25.94 W m^{-2}), as well as with other sites, such as XDT (−11.35 W m^{-2}), ZNH (−10.29 W m^{-2}), TSH (−10.17 W m^{-2}), and SZ (−10.15 W m^{-2}). In contrast, DenseMLP_TL keeps low bias across all stations, with bias values within approximately ±3.7 W m^{-2} (e.g., 2.68 W m^{-2} at Haibei, −1.83 W m^{-2} at Naqu, −2.69 W m^{-2} at SZ, −1.63 W m^{-2} at TSH, −3.61 W m^{-2} at XDT, 0.27 W m^{-2} at ZNH) and no extreme site-specific deviations. In terms of RMSE, at Haibei and XDT, the two models yield comparable values, with both around 21 W m^{-2} at Haibei and around 26 W m^{-2} at XDT. For the remaining sites, DenseMLP_TL outperforms GLASS, often by a considerable margin; for instance, at Naqu (23.69 vs. 33.53 W m^{-2}) and TSH (16.06 vs. 20.88 W m^{-2}). These results confirm that the DenseMLP_TL model effectively mitigates the site-dependent systematic errors present in the GLASS product, achieving higher overall accuracy and better generalization.

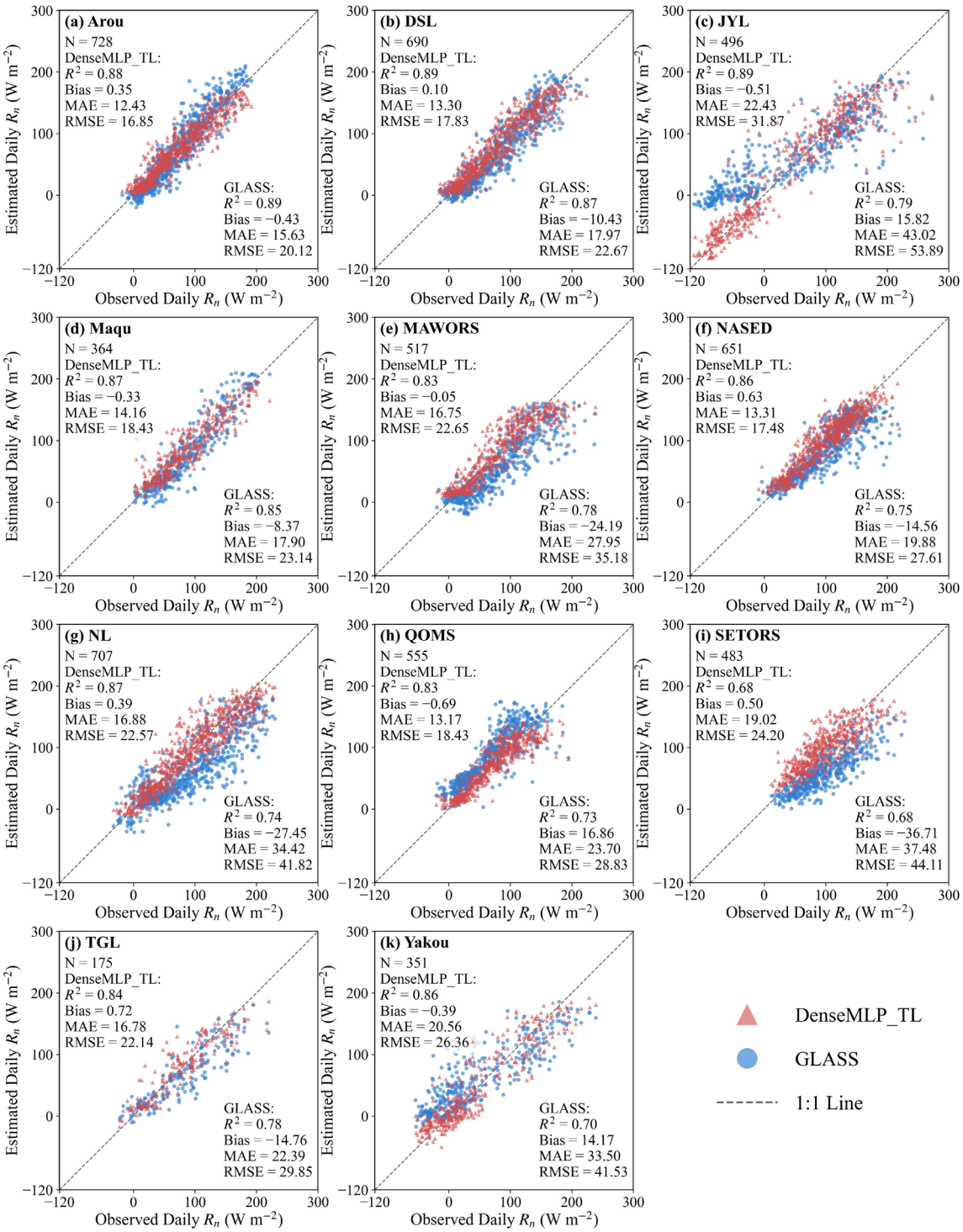


Figure 3. Scatter plots of observed daily R_n from DenseMLP_TL and GLASS versus observed daily R_n of eleven stations for model training.

To further evaluate the temporal performance of the proposed model, we compared the time series of R_{n_daily} from DenseMLP_TL estimates, the GLASS product, and in situ observations at the six test stations over the year 2018 (Figure 5). The figure shows that both DenseMLP_TL estimates and the GLASS product effectively capture the overall temporal characteristics of R_{n_daily}. The lowest R_{n_daily} values occur in winter (November, December, and January), followed by a gradual increase in spring, with the highest values appearing in July and August, and a subsequent decrease in autumn. In addition, both DenseMLP_TL and GLASS are capable of capturing the fluctuations associated with cloudy or rainy condi-

tions. However, compared with the GLASS product, the DenseMLP_TL estimates show better agreement with in situ observations, especially on clear-sky days. GLASS exhibits a pronounced underestimation at Naqu, XDT, and ZNH, whereas DenseMLP_TL estimates are closer to the in situ observations. These results demonstrate that the DenseMLP_TL model achieves superior performance in capturing both the seasonal dynamics and the day-to-day variability of daily R_n, with significantly improved accuracy and stability compared to the GLASS benchmark.

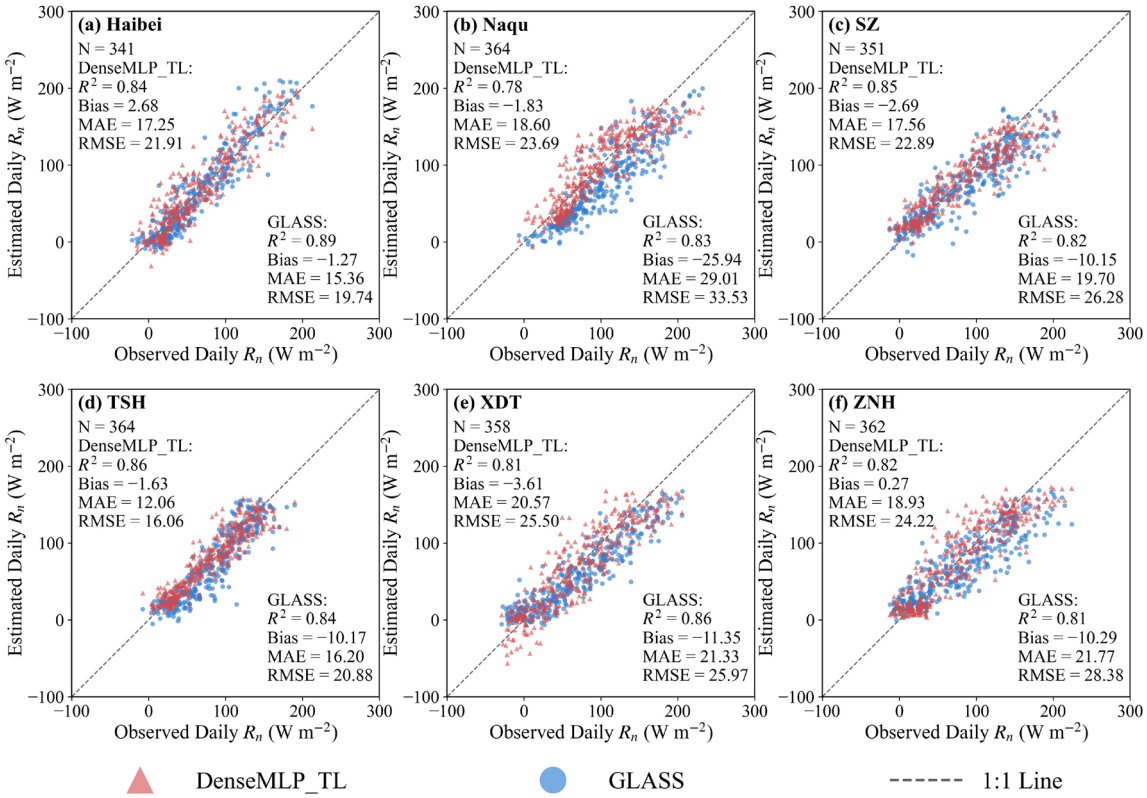


Figure 4. Scatter plots of estimated daily R_n from DenseMLP_TL and GLASS versus observed daily R_n of six stations for independent testing.

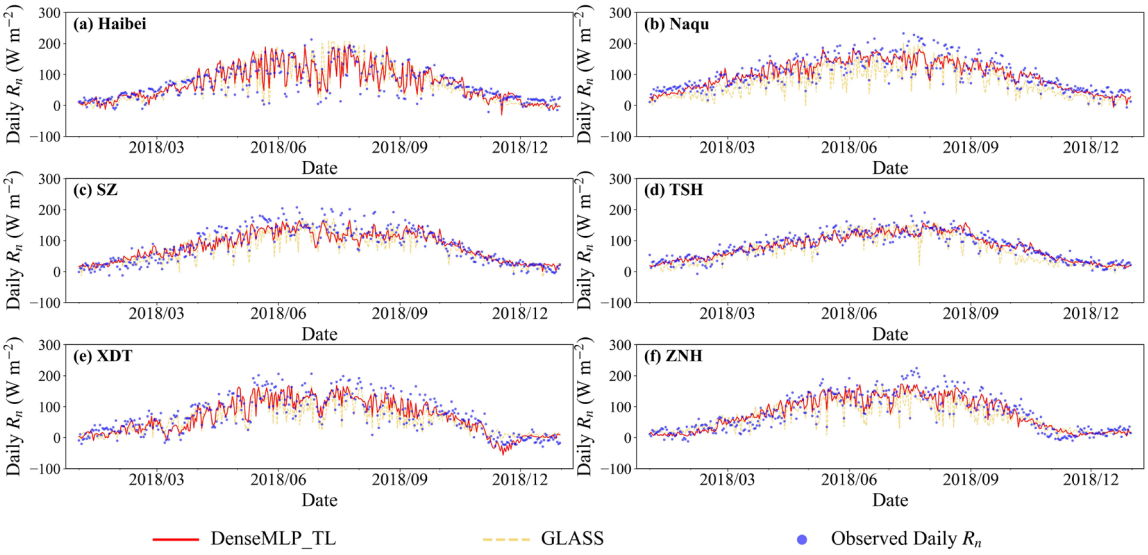


Figure 5. Time series of daily R_n from DenseMLP_TL estimates, GLASS product, and in situ observations at the six test stations over the year 2018.

3.2. Spatial Extension

Here, this section extends the DenseMLP_TL model to estimate 1 km all-weather daily R_n across the TP for year 2018. Figure 6 presents the spatial distribution of the estimated monthly mean R_n from January to December 2018. The spatial distribution of monthly mean R_n shows a consistent pattern throughout the year. Higher values occur in low-elevation regions, such as the eastern plateau and the southern margin, while lower values are found over the high-altitude zones across the central and northeastern plateau, reflecting a clear topographical control. In addition, low-elevation regions covered by high-albedo desert surfaces, such as the Qaidam Basin, show distinctly lower R_n compared with other lowland areas. This spatial distribution agrees well with previous studies [12,22,34]. The pronounced elevation-dependent variation further confirms that topography plays a dominant role in shaping the spatial heterogeneity of all-weather R_n, which supports the conclusion that topographic correction is essential for accurate radiation estimation over complex terrain [35]. The monthly R_n estimates also capture a distinct seasonal cycle across the TP. Low values (<50 W m^{-2}) dominate in winter (November–January). From March to May, R_n increases steadily, reaching a peak in June (>150 W m^{-2} over most areas). High values persist through July and August, then gradually decrease after September. This seasonal pattern is consistent with previous observational studies on the TP [27,36], which have consistently reported minimum net radiation in winter and maximum in summer.

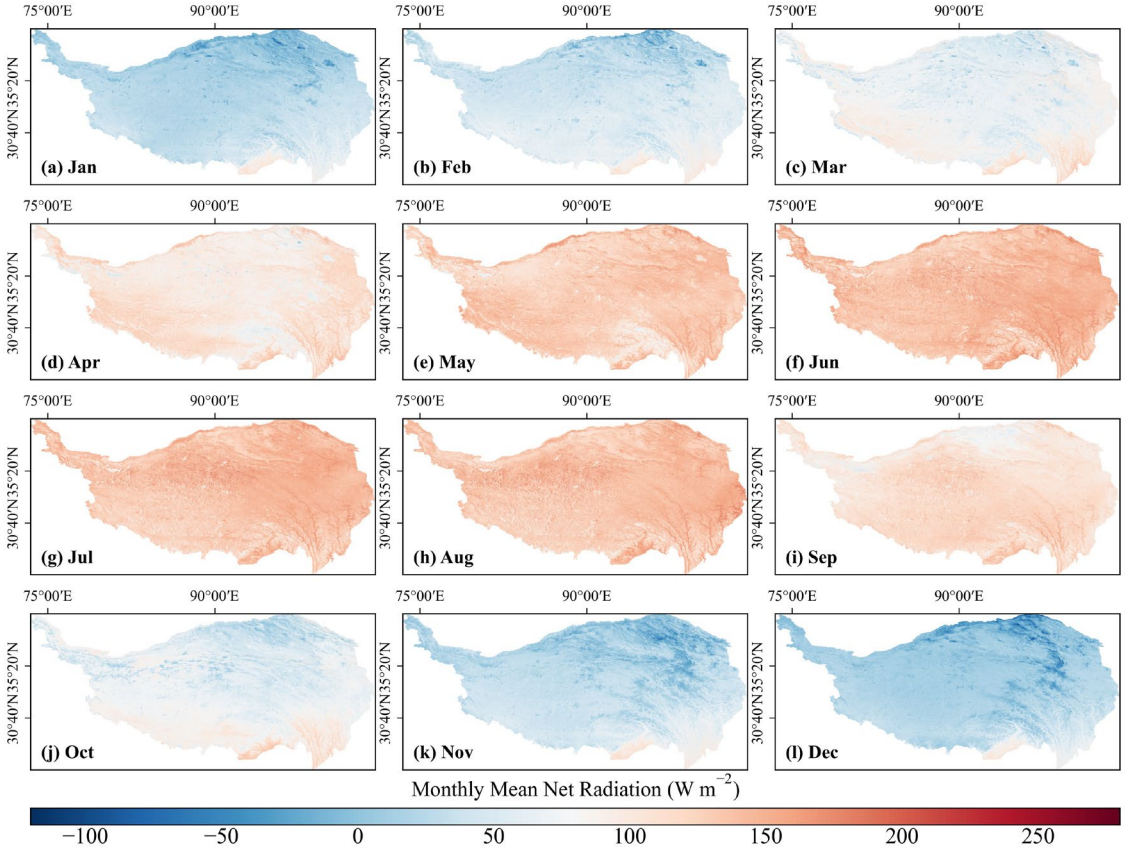


Figure 6. Monthly mean all-weather R_n maps across the Tibetan Plateau for 2018 (January–December).

4. Discussion

4.1. Model Analysis and Ablation Experiments

4.1.1. Model Evaluation with Different Hyperparameter Combinations

As elaborated in Section 2.2.1, the growth rate w (representing model width) and the number of dense blocks d (representing model depth) are two core hyperparameters

governing the feature learning capability, structural complexity, computational cost, and generalization performance of the proposed DenseMLP_TL model. Reasonable configuration of these two hyperparameters is critical to balancing feature representation richness, model overfitting risk, and training efficiency, especially for regional R_{n_daily} estimation tasks that require both high prediction accuracy and stable generalization. To identify the optimal hyperparameter combination for the target estimation task, this study systematically designed a controlled hyperparameter tuning experiment, constructing a total of 25 independent models with gradient variations in d ranging from 1 to 5 and *w* ranging from 8 to 128. The performance of all candidate models, including R^2, MAE, RMSE, and bias, was compared and visualized in Figure 7.

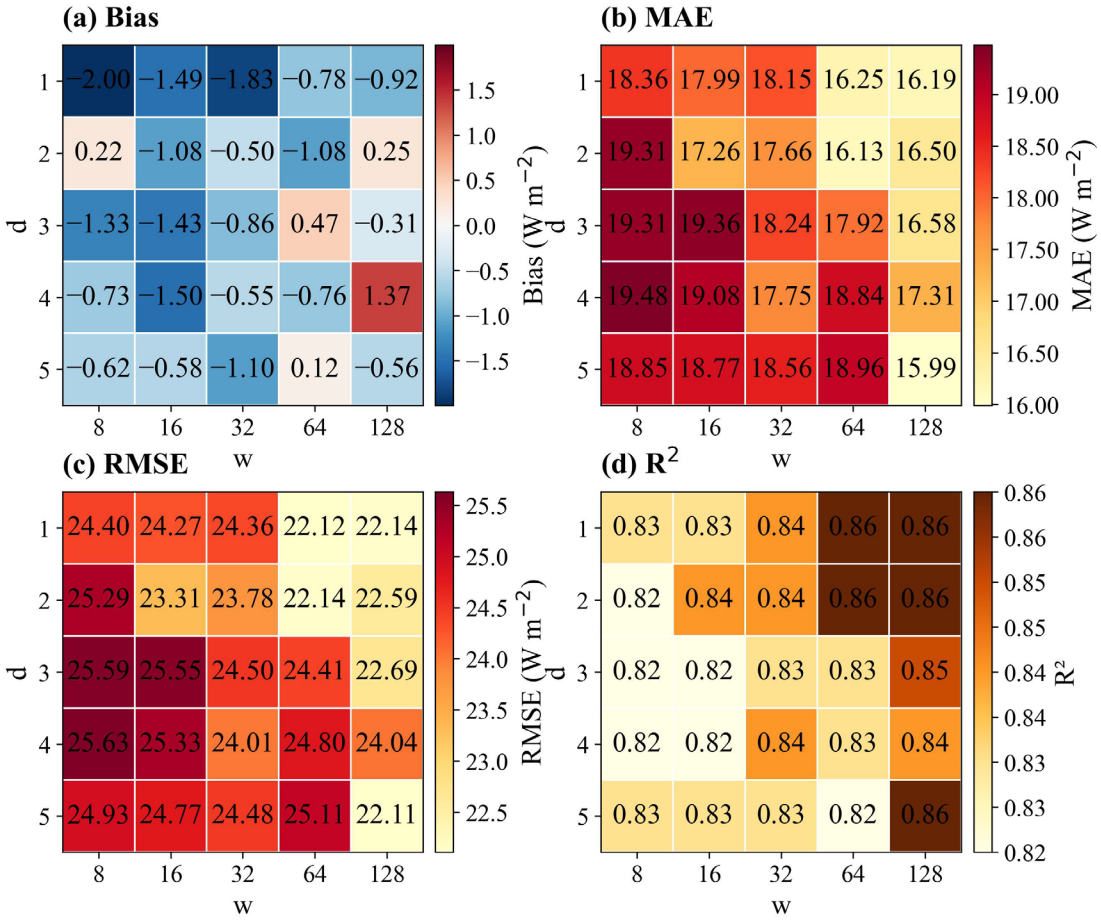


Figure 7. Performance of DenseMLP models with hyperparameter combinations.

In terms of model width *w*, within a fixed model depth (where d is 1, 2, or 3), a gradual increase in w from 8 to 128 generally enhances the model's feature extraction capability at the initial stage. The validation R^2 shows an overall upward trend, while both MAE and RMSE decrease continuously. This is because a larger growth rate *w* enables each dense block to capture more refined spatial and temporal feature information related to R_{n_daily} [37,38]. It enriches the feature representation space and improves the fitting ability of the model to the nonlinear variation pattern of R_{n_daily} [37,39]. Meanwhile, the systematic bias of the model remains relatively low within the appropriate *w* range, indicating no obvious systematic deviation in the prediction results. When the model width exceeds a certain threshold under a fixed depth, excessively large *w* leads to a sharp increase in the number of trainable parameters [38]. This results in model overfitting to the pre-training data and weakened generalization ability on the validation dataset [40]. Such

deterioration is manifested as a significant drop in R^2 and a marked rise in MAE and RMSE, particularly for deeper model configurations where the performance gain from increased width diminishes after w reaches 32 (e.g., d = 4, d = 5).

For model depth d, the impact on estimation performance presents a generally positive but highly nonlinear relationship jointly constrained by model width [38,41,42]. When the depth increases from 1 to 5, the deepened network structure facilitates cross-block feature reuse and multi-level feature transformation. This allows the model to mine high-order abstract features related to R_{n_daily}. In the presence of sufficient model width (e.g., w = 64 or 128), increasing depth does not always continuously lift accuracy, as shallow configurations d = 1 and d = 2 already achieve competitive performance, with the best overall performance achieved at d = 5, w = 128. However, when the width is limited (e.g., w = 8 or 16), deeper networks do not always yield better results, indicating that depth and width interact nonlinearly. Nonetheless, for the optimal width of 128, the deepest configuration (d = 5) yields the highest R^2 and lowest error metrics, suggesting that within the explored range, deeper is beneficial when combined with adequate width.

Based on the comprehensive comparison of all 25 hyperparameter combinations, the optimal model configuration for this study is determined as d = 5 and w = 128, which achieves the best overall validation performance with a R^2 of 0.86, MAE of 15.99 W m^{-2}, RMSE of 22.11 W m^{-2}, and a low bias of −0.56 W m^{-2}. This combination strikes an ideal balance between model width and depth: the sufficient depth (d = 5) allows effective multi-level feature transformation and reuse, while the large width (w = 128) provides sufficient feature representation capability, together maximizing the estimation accuracy of R_{n_daily} while maintaining robust generalization performance. In contrast, overly shallow or narrow models lack sufficient feature learning capacity and fail to meet the accuracy requirements of the target task [39,43], while excessively wide models with insufficient depth suffer from limited performance improvement rather than severe overfitting and redundant parameters [38,44]. This hyperparameter tuning result also confirms that targeted optimization of core structural hyperparameters is indispensable for constructing a high-precision and stable R_{n_daily} estimation model, and blind pursuit of deeper or wider network structures is not conducive to improving the practical application effect of the model in regional R_{n_daily} simulation [45,46].

4.1.2. Ablation Study on Input Variables and Model Components

To verify the physical rationality of the input variables and clarify the independent contributions of the DenseMLP architecture and transfer learning (TL) strategy, we conducted a series of ablation experiments. Key variable groups, including Srad and Lrad, LST, terrain factors, and MODIS optical reflectance bands, were sequentially removed from the full DenseMLP_TL model (E0). Two baseline models without transfer learning were also established: a vanilla DenseMLP adopting the same input set (E5), and another DenseMLP integrated with all inputs plus coarse-resolution GLASS R_n as a physical constraint (E6). Furthermore, to isolate the inherent contribution of the DenseMLP architecture, we added a structure-controlled baseline (E7): a standard MLP trained with the same input variables and identical transfer learning pipeline. All models were evaluated on the training-site validation set (eleven in situ stations with a 90/10 train–validation split) and a fully independent test set consisting of six stations. The results are summarized in Table 3.

Removing both Srad and Lrad (E1) led to noticeable performance degradation across the two datasets. On the independent testing set, R^2 decreased from 0.83 (E0) to 0.75, while MAE rose from 17.43 to 20.95 W m^{-2} and RMSE increased from 22.55 to 27.36 W m^{-2}. This result is reasonable, since R_n is physically derived from downwelling radiation fluxes minus upwelling components, making Srad and Lrad inherently strong predic-

tors for R_n estimation [9,18]. Many previous studies have also proven that Srad and Lrad can effectively retrieve R_n via either radiative transfer parameterization or data-driven approaches [1,11,14,24,34]. Even so, E1 still outperformed the transfer learning-free DenseMLP (E5, R^2 = 0.68, RMSE = 33.77 W m^{-2} on the testing set) and yielded a far smaller bias than the original GLASS R_n product (−1.43 vs. −11.65 W m^{-2}). This demonstrates that Srad and Lrad are valuable though not indispensable inputs for our model.

Table 3. Performance of the full model and ablation variants on the training-site validation set and the independent test set.

Name	Experiment Setting	Evaluation Dataset	R^2	Bias (W m^{-2})	MAE (W m^{-2})	RMSE (W m^{-2})
E0	DenseMLP_TL with all variables	Validation set	0.88	−0.49	15.66	21.27
		Test set	0.83	−1.12	17.43	22.55
E1	DenseMLP_TL without Srad and Lrad	Validation set	0.80	−1.73	20.51	27.15
		Test set	0.75	−1.43	20.95	27.36
E2	DenseMLP_TL without LST	Validation set	0.79	0.64	21.13	27.68
		Test set	0.80	0.13	18.93	24.38
E3	DenseMLP_TL without terrain	Validation set	0.87	−0.19	16.01	21.57
		Test set	0.81	0.73	18.04	23.52
E4	DenseMLP_TL without optical band reflectance	Validation set	0.83	0.79	19.31	25.47
		Test set	0.80	0.83	18.68	23.99
E5	DenseMLP with all variables	Validation set	0.86	1.49	17.41	23.15
		Test set	0.68	−10.07	25.70	33.77
E6	DenseMLP with all variables plus GLASS R_n	Validation set	0.88	−1.68	15.04	21.39
		Test set	0.80	−13.5	23.57	29.53
E7	MLP_TL with all variable	Validation set	0.78	0.10	22.01	28.47
		Test set	0.79	−0.81	19.86	25.08

Eliminating LST (E2), terrain factors (E3), or optical reflectance bands (E4) only caused mild accuracy loss compared with the full model (E0). On the independent test set, E2 achieved R^2 = 0.80 and RMSE = 24.38 W m^{-2}; E3 achieved R^2 = 0.81 and RMSE = 23.52 W m^{-2}; and E4 achieved R^2 = 0.80 and RMSE = 23.99 W m^{-2}. All these variants performed considerably better than the two non-TL baselines (E5 and E6). Each input group provides complementary information for R_n retrieval. Topography regulates the spatial redistribution of surface radiation [35], LST characterizes near-surface thermal status, and optical reflectance reflects surface spectral properties.

Comparison between the two non-TL models (E5 and E6) reveals distinct differences. Incorporating GLASS R_n as an additional physical constraint (E6) effectively improved performance on the training-site validation set relative to E5, with RMSE declining from 23.15 to 21.39 W m^{-2} and R^2 rising from 0.86 to 0.88. This agrees with existing findings that coarse-resolution radiation products can act as effective physical constraints to boost model accuracy [22]. Nevertheless, E6 suffered severe generalization degradation, with its RMSE increasing sharply from 21.39 W m^{-2} on the validation set to 29.53 W m^{-2} on the test set, accompanied by a large negative bias of −13.50 W m^{-2}. In contrast, transfer learning fundamentally alleviates this issue. By pre-training on massive GLASS R_n datasets, the DenseMLP_TL framework learns generalized feature patterns and achieves stable prediction across independent stations. This superiority is reflected in the consistent accuracy of the full model (E0) between validation and test sets (RMSE: 21.27 vs. 22.55 W m^{-2}). Moreover, all DenseMLP_TL variants after removing individual input variables (E1–E4)

still outperformed non-TL models on the independent test set, fully verifying the strong robustness of the proposed framework.

Most critically, the E7 baseline (standard MLP_TL with identical inputs and TL configuration) enables a strict ablation of network architecture effects. On the independent test set, E7 achieved R^2 = 0.79 and RMSE = 25.08 W m^{-2}, which are clearly inferior to the full DenseMLP_TL model (E0: R^2 = 0.83, RMSE = 22.55 W m^{-2}). This controlled comparison explicitly verifies the independent, non-negligible contribution of the dense connection architecture. Under a completely identical transfer learning framework and input regime, the vanilla MLP fails to extract and transfer cross-scale geospatial and radiative features as effectively as DenseMLP, leading to weaker generalization across diverse sites. Therefore, while transfer learning addresses the scarcity of in situ observations and improves overall generalization, the DenseMLP structure further enhances retrieval accuracy and stability under varied geographical conditions.

4.2. Comparison with Previous Studies

To evaluate the reliability, advancement, and regional applicability of the proposed DenseMLP_TL model for R_{n_daily} estimation over the TP, we conduct two comparative analyses: (1) a comparison with representative R_{n_daily} estimation methods to highlight advantages in algorithm design and regional adaptability, and (2) a comparison with mainstream reanalysis and remote sensing datasets to assess spatial distribution and estimation accuracy.

4.2.1. Comparison with Existing Estimation Methods

Although existing studies [11,12,19,21,22,24] have made remarkable progress in R_{n_daily} estimation, they still exhibit clear limitations in regional adaptability, environmental constraints, and applicability to complex terrain, which hinder their direct application to our study area. Specifically, most previous approaches suffer from one or more common issues, such as ignoring topographic effects, being restricted to clear-sky conditions, lacking evaluation in complex high-altitude terrain, or relying on training samples that are difficult to obtain in regions like the TP.

Against this background, the proposed DenseMLP_TL model addresses these deficiencies through targeted optimization. In terms of algorithm design and spatial resolution, Li et al. [21] employed a transfer learning framework based on ResNet, but its output remains at a coarse resolution of 9 km, limiting fine-scale applications. Jiang et al. [11] produced high-resolution (30 m) products, yet their method is restricted to clear-sky conditions, which constrains its utility in cloud-prone regions. Jing et al. [12] generated a high-resolution daily net radiation dataset for China (CHiRAD) at a 0.05° resolution, which still smooths out terrain-induced variations. Chen et al. [19] designed a model tailored for high-latitude regions by exploiting the very high temporal repeat frequency of MODIS observations (~6 to 20 times per day), yet such frequent overpasses are not available over the TP. Li et al. [22] developed a random forest model constrained by ERA5 to address the limited overpass frequency of MODIS at mid–low latitudes, achieving strong global evaluation accuracy. However, their approach relies heavily on ERA5 product, which may carry biases over the TP. Miao et al. [24] also adopted a two-stage transfer learning strategy and a convolutional DenseNet to estimate 1 km daily R_n over the Heihe River Basin. Differing from their convolutional framework, our study develops a fully connected DenseMLP, which eliminates artificial spatial inductive bias and better conforms to the physical characteristics of R_n estimation.

By comparison, our model operates under all-weather conditions while maintaining 1 km spatial detail, and offers two key advantages: (i) all-weather estimation capability for

complete temporal coverage, and (ii) a transfer learning strategy that ensures robust generalization with limited local observations. These features make our approach particularly suitable for R_{n_daily} over the TP.

4.2.2. Comparison with Mainstream R_n Datasets

On the basis of comparing with existing estimation methods, this study further conducts a systematic comparison between the DenseMLP_TL estimated results and four R_{n_daily} products, including Global Land Data Assimilation System (GLDAS) with 0.25° spatial resolution [47], ERA5-Land with 0.1° spatial resolution [48], GLASS with 0.05° spatial resolution [30], CHiRAD with 0.05° spatial resolution [12]. Since the trained DenseMLP_TL model was applied to the regional data for 2018 to estimate the 1 km all-weather R_{n_daily} across the TP, the comparative analysis in this section is based on data from 2018 to ensure the consistency and rationality of the comparison. The comparison focuses on two core aspects: spatial distribution consistency (using original products) and site-scale estimation accuracy (after unified linear calibration to ensure fairness).

As illustrated in Figure 8, the spatial patterns of the annual mean R_n estimated by the DenseMLP_TL are generally consistent with those of GLDAS, ERA5-Land, GLASS, and CHiRAD, all reflecting the overall characteristics of R_n distribution in the TP (e.g., high values in low-altitude valleys and low values in high-altitude cold regions) [35]. Quantitatively, DenseMLP_TL and GLDAS exhibit the most consistent estimates, with both showing that over 50% of the study area has annual mean R_n above 75 W m^{-2}, whereas ERA5-Land, GLASS, and CHiRAD yield generally lower values. However, obvious differences exist in spatial detail. With its higher spatial resolution (1 km) (Figure 8e), DenseMLP_TL shows substantially more spatial detail than the coarse-resolution products (Figure 8a–d). In particular, it clearly distinguishes R_n differences between adjacent mountain slopes and valleys, whereas the other four datasets smooth out such fine-scale variations.

To provide quantitative evidence of the model's ability to capture slope-aspect effects, we further analyzed how R_n varies with slope aspect, a topographic factor long recognized as a primary control on surface radiation balance [49–52]. As shown in Table 4, the DenseMLP_TL product exhibits distinct aspect-related R_n contrasts: the east–west difference reaches 16.59 W m^{-2} and the south–north difference 2.29 W m^{-2}, consistent with the expected pattern that sunlit slopes (east and south) receive higher R_n than shaded slopes (west and north) [52]. In contrast, the four coarse-resolution or reanalysis products (GLDAS, ERA5-Land, GLASS, and CHiRAD) show negligible aspect-induced variations, with all inter-aspect differences below 1 W m^{-2}. This comparison demonstrates that coarse-resolution products fail to resolve fine-scale R_n contrasts between differently oriented slopes, whereas the 1 km DenseMLP_TL reproduces aspect-related terrain effects.

Table 4. Annual mean net radiation (W m^{-2}) for different aspects and aspect-induced contrasts (east–west and south–north differences) across five products.

Product	North	East	South	West	East–West	South–North
GLDAS	87.41	87.35	87.75	87.60	−0.25	0.34
ERA5-Land	65.21	65.43	65.33	65.43	0.00	0.12
GLASS	66.20	66.43	66.37	66.57	−0.14	0.17
CHiRAD	65.49	65.66	64.95	65.87	−0.21	−0.54
DenseMLP_TL	83.48	92.22	85.77	75.63	16.59	2.29

It should be noted that for this spatial comparison, all four mainstream products are presented in their original form, because linear correction coefficients derived from ground stations (elevation range 3000–5100 m) cannot be safely extrapolated to the entire

TP without causing physically unrealistic overcorrection in areas outside the observed elevation range.

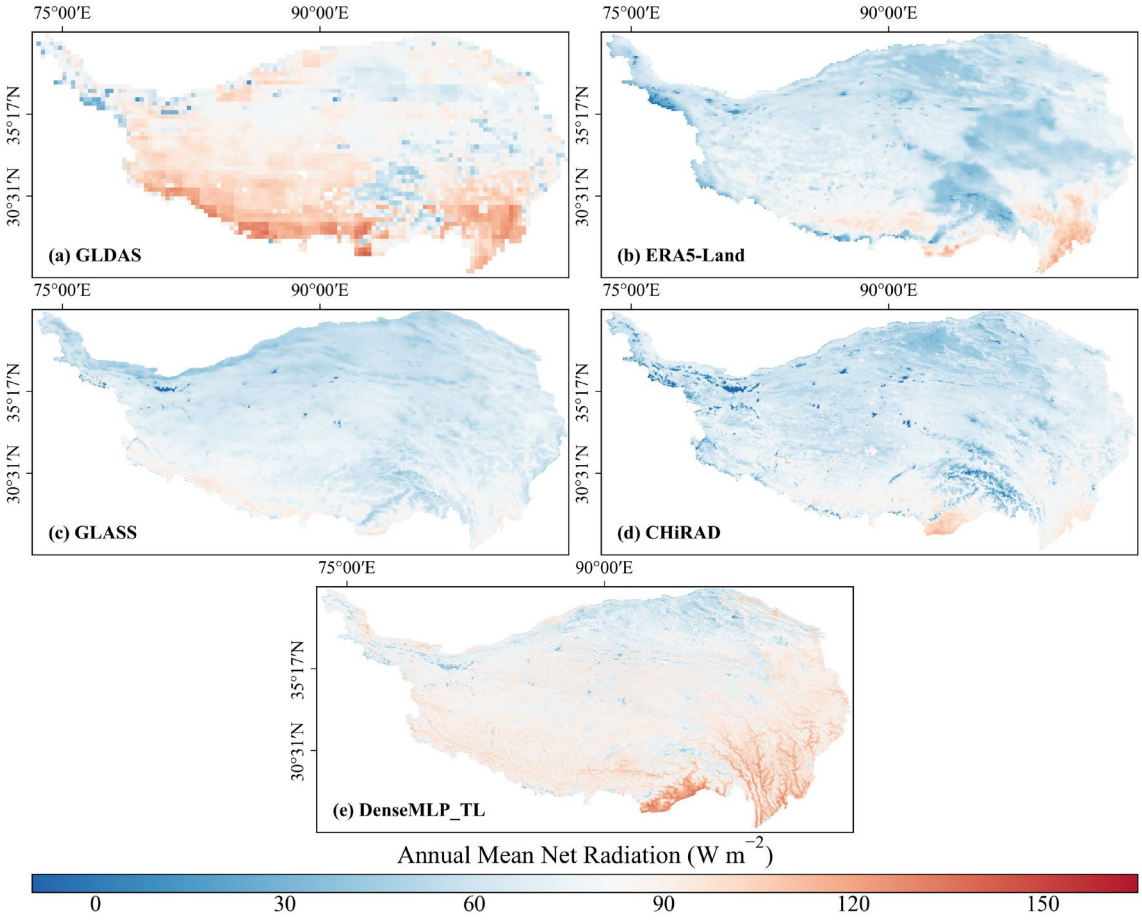


Figure 8. Spatial distribution of annual mean R_n over the Tibetan Plateau in 2018 derived from (**a**) GLDAS, (**b**) ERA5-land, (**c**) GLASS, (**d**) CHiRAD, and (**e**) DenseMLP_TL.

To quantitatively evaluate the accuracy of the five R_n products at station locations while ensuring a fair comparison with the fine-tuned DenseMLP_TL model, we applied a unified first-order linear regression calibration ($R^{cal}_{n_daily} = a \times R^{orig}_{n_daily} + b$) to GLDAS, ERA5-Land, GLASS, and CHiRAD following the linear scaling method commonly used in remote sensing product bias correction [53]. The calibration coefficients (a, b) were derived from the observations of the eleven training sites, which are the same sites used for fine-tuning DenseMLP_TL. This calibration removes systematic biases at the training sites and allows an equitable comparison at the six fully independent test stations (which were not used for fine-tuning in DenseMLP_TL model). Table 5 summarizes the performance of the original and calibrated products together with DenseMLP_TL at the six test stations.

Among all datasets, DenseMLP_TL achieves the best overall performance, with the highest R^2 (0.83), the smallest bias (-1.12 W m^{-2}), and the lowest MAE (17.43 W m^{-2}) and RMSE (22.55 W m^{-2}). After calibration, the biases of the four mainstream products are substantially reduced (e.g., GLASS from -11.65 to -4.08 W m^{-2}, ERA5-Land from -18.03 to -2.23 W m^{-2}). However, their R^2, MAE, and RMSE remain notably worse than those of DenseMLP_TL. GLASS shows the best among the calibrated products (R^2 = 0.81, MAE = 18.60 W m^{-2}, RMSE = 23.89 W m^{-2}), but still underperforms DenseMLP_TL. The superior performance of DenseMLP_TL can be attributed to its higher spatial resolution (1 km) and its ability to capture terrain-induced spatial heterogeneity, as discussed earlier, which may lead to more accurate point-to-pixel matching at station locations. These

comparisons further demonstrate that the proposed model effectively reduces both random and systematic errors, providing a more reliable estimate of R_n over the complex terrain of the TP.

Table 5. Site-scale performance of original and calibrated mainstream products versus DenseMLP_TL at the six independent test stations.

Source	a	b	R^2	Bias (W m^{-2})	MAE (W m^{-2})	RMSE (W m^{-2})
GLDAS (original)	N/A	N/A	0.75	9.74	22.75	28.96
GLDAS (calibrated)	0.96	−10.79	0.75	−3.99	21.25	27.63
ERA5_Land (original)	N/A	N/A	0.72	−18.03	27.2	34.17
ERA5_Land(calibrated)	0.9	21.63	0.72	−2.23	22.22	28.53
GLASS (original)	N/A	N/A	0.81	−11.65	20.62	26.27
GLASS(calibrated)	1.02	5.97	0.81	−4.08	18.6	23.89
CHiRAD (original)	N/A	N/A	0.75	−17.00	25.93	32.01
CHiRAD(calibrated)	0.97	10.86	0.75	−7.71	22.40	28.07
DenseMLP_TL	N/A	N/A	0.83	−1.12	17.43	22.55

Notably, there are fundamental structural differences between DenseMLP_TL and the four benchmark R_n datasets. DenseMLP_TL is a nonlinear multi-source model fine-tuned using local plateau observations, whereas GLDAS, ERA5-Land, GLASS, and CHiRAD are publicly available precomputed products without local calibration, which can only be adjusted via post hoc linear scaling. Despite these structural disparities, we applied consistent linear calibration to all datasets to remove systematic bias and support reliable cross-product evaluation across the TP. Our results reveal the evident advantages of DenseMLP_TL relative to these benchmark products within this study region, though we acknowledge the comparison is not perfectly equivalent when accounting for differences in model complexity and fitting flexibility.

5. Conclusions

To obtain accurate R_{n_daily} at 1 km spatial resolution over the TP, this study proposed a DenseMLP-based transfer learning framework with a two-stage strategy (coarse pre-training on GLASS R_{n_daily}, followed by fine-tuning on limited TP ground observations) using land surface parameters from MODIS products and other auxiliary radiative, meteorological, and topographic data. The proposed model was evaluated using six independent testing stations distributed across the TP. From the scatter plots, the DenseMLP_TL estimates are tightly clustered around the 1:1 line for both training and test sets, indicating close agreement with ground observations and no significant systematic bias. From a quantitative perspective, DenseMLP_TL achieves an overall R^2 of 0.87, MAE of 16.06 W m^{-2}, RMSE of 21.94 W m^{-2}, and a near-zero bias of 0.07 W m^{-2} on the training set. On the independent test set, it maintains robust performance, with an overall R^2 of 0.83, MAE of 17.43 W m^{-2}, RMSE of 22.55 W m^{-2}, and a bias of −1.12 W m^{-2}. Additionally, the proposed method exhibits consistent and low bias across individual sites, with bias values mostly within ±3.7 W m^{-2}. Time-series comparisons further demonstrate that the model accurately captures the seasonal dynamics of R_{n_daily} (lowest in winter, highest in summer) and the day-to-day fluctuations associated with cloudy or rainy conditions, consistently tracking the in situ observations across all six test stations. Compared with the GLASS benchmark, DenseMLP_TL estimates show closer alignment with observed values and fewer site-specific discrepancies, confirming that the proposed framework effectively mitigates site-dependent systematic errors and achieves high accuracy and good spatial generalization for R_{n_daily} estimation over the TP.

Subsequently, the proposed framework was used to generate 1 km resolution R_{n_daily} over the entire TP for 2018. The spatial distribution of R_{n_daily} is clearly controlled by topography, with higher values in low-elevation regions and lower values over high-altitude zones. Notably lower R_{n_daily} is observed in high-albedo desert areas, such as the Qaidam Basin. The monthly mean R_n follows a distinct annual cycle, with low values in winter, a peak in June, and a gradual decrease after September. Compared with mainstream products (GLDAS, ERA5-Land, GLASS, and CHiRAD), the spatial pattern of annual mean R_n from DenseMLP_TL is generally consistent. Importantly, benefiting from its 1 km resolution, DenseMLP_TL clearly resolves aspect-dependent R_n contrasts (up to 16.59 W m^{-2}) that are entirely absent in coarse-resolution datasets, underscoring the value of fine spatial resolution for aspect-sensitive R_n estimation. Quantitatively, DenseMLP_TL consistently outperforms all compared products across evaluation metrics. These spatial and seasonal patterns, along with inter-product comparisons, further illustrate the generalization capability of the proposed framework. In future work, we plan to develop a long-term, 1 km daily R_n product over the TP based on the DenseMLP_TL model.

Author Contributions: Conceptualization, Y.L. and L.Z.; methodology, L.L., Y.L. and J.H.; software, L.L.; validation, L.L. and Y.L.; formal analysis, L.L. and Y.L.; investigation, L.L., L.W. and G.H.; resources, L.Z.; data curation, L.L.; writing—original draft preparation, L.L.; writing—review and editing, Y.L., L.Z., J.H., L.W. and G.H.; visualization, L.L. and Y.L.; supervision, L.Z.; project administration, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant numbers 42301157 and 42322608, and the Natural Science Foundation of Jiangsu Province, grant number BK20230429.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jiang, B.; Zhang, Y.; Liang, S.; Wohlfahrt, G.; Arain, A.; Cescatti, A.; Georgiadis, T.; Jia, K.; Kiely, G.; Lund, M.; et al. Empirical estimation of daytime net radiation from shortwave radiation and ancillary information. *Agric. For. Meteorol.* **2015**, *211–212*, 23–36. [CrossRef]
2. Kaminsky, K.Z.; Dubayah, R. Estimation of surface net radiation in the boreal forest and northern prairie from shortwave flux measurements. *J. Geophys. Res. Atmos.* **1997**, *102*, 29707–29716. [CrossRef]
3. Sellers, P.J.; Dickinson, R.E.; Randall, D.A.; Betts, A.K.; Hall, F.G.; Berry, J.A.; Collatz, G.J.; Denning, A.S.; Mooney, H.A.; Nobre, C.A.; et al. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* **1997**, *275*, 502–509. [CrossRef] [PubMed]
4. Zeng, Z.-J.; Tang, C.-S.; Cheng, Q.; Yang, Z.-M.; Hao, S.-F.; Shi, B. Effect of net radiation on soil temperature field considering slope angle. *J. Rock Mech. Geotech. Eng.* **2025**, *18*, 3038–3055. [CrossRef]
5. Hsu, H.H.; Liu, X. Relationship between the Tibetan Plateau heating and East Asian summer monsoon rainfall. *Geophys. Res. Lett.* **2003**, *30*, 517–537. [CrossRef]
6. Wu, G.; Liu, Y.; He, B.; Bao, Q.; Duan, A.; Jin, F.-F. Thermal Controls on the Asian Summer Monsoon. *Sci. Rep.* **2012**, *2*, 404. [CrossRef] [PubMed]
7. Xu, M.; An, T.; Zheng, Z.; Zhang, T.; Zhang, Y.; Yu, G.; Hao, Y. Variability in evapotranspiration shifts from meteorological to biological control under wet versus drought conditions in an alpine meadow. *J. Plant Ecol.* **2022**, *15*, 921–932. [CrossRef]
8. Zhu, M.; Wang, J.; Ivanov, V.; Sheshukov, A.; Zhou, W.; Zhang, L.; Mazepa, V.; Sokolov, A.; Valdayskikh, V. An Analytical Model of Active Layer Depth Under Changing Ground Heat Flux. *J. Geophys. Res. Atmos.* **2024**, *129*, e2023JD039453. [CrossRef]
9. Blonquist, J.M.; Tanner, B.D.; Bugbee, B. Evaluation of measurement accuracy and comparison of two new and three traditional net radiometers. *Agric. For. Meteorol.* **2009**, *149*, 1709–1721. [CrossRef]
10. Li, P.; Zhong, L.; Ma, Y.; Fu, Y.; Cheng, M.; Wang, X.; Qi, Y.; Wang, Z. Estimation of 1 km downwelling shortwave radiation over the Tibetan Plateau under all-sky conditions. *Atmos. Chem. Phys.* **2023**, *23*, 9265–9285. [CrossRef]

11. Jiang, B.; Han, J.; Liang, H.; Liang, S.; Yin, X.; Peng, J.; He, T.; Ma, Y. The Hi-GLASS all-wave daily net radiation product: Algorithm and product validation. *Sci. Remote Sens.* **2023**, *7*, 100080. [CrossRef]
12. Jing, Y.; Bai, P.; Yang, Z. CHiRAD: A high-resolution daily net radiation dataset for China generated using meteorological and albedo data. *J. Hydrol.* **2025**, *654*, 132854. [CrossRef]
13. Liang, S.; Wang, D.; He, T.; Yu, Y. Remote sensing of earth's energy budget: Synthesis and review. *Int. J. Digit. Earth* **2019**, *12*, 737–780. [CrossRef]
14. Rains, D.; Trigo, I.; Dutra, E.; Ermida, S.; Ghent, D.; Hulsman, P.; Gómez-Dans, J.; Miralles, D.G. High-resolution (1 km) all-sky net radiation over Europe enabled by the merging of land surface temperature retrievals from geostationary and polar-orbiting satellites. *Earth Syst. Sci. Data* **2024**, *16*, 567–593. [CrossRef]
15. Xu, J.; Liang, S.; Ma, H.; Chen, Y.; Li, W.; Ma, Y.; Zhao, X.; Jiang, B.; Zhang, X.; Guan, S. Joint estimation of global daily 1 km surface radiation budget components from MODIS observations (2000−2023) using conservation-constrained deep neural networks. *Remote Sens. Environ.* **2026**, *333*, 115135. [CrossRef]
16. Wang, G.; Wang, T.; Xue, H. Validation and comparison of surface shortwave and longwave radiation products over the three poles. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102538. [CrossRef]
17. Bisht, G.; Bras, R.L. Estimation of net radiation from the MODIS data under all sky conditions: Southern Great Plains case study. *Remote Sens. Environ.* **2010**, *114*, 1522–1534. [CrossRef]
18. Bisht, G.; Venturini, V.; Islam, S.; Jiang, L. Estimation of the net radiation using MODIS (Moderate Resolution Imaging Spectroradiometer) data for clear sky days. *Remote Sens. Environ.* **2005**, *97*, 52–67. [CrossRef]
19. Chen, J.; He, T.; Jiang, B.; Liang, S. Estimation of all-sky all-wave daily net radiation at high latitudes from MODIS data. *Remote Sens. Environ.* **2020**, *245*, 111842. [CrossRef]
20. Foyo-Moreno, I.; Lozano, I.L.; Alados, I.; Guerrero-Rascado, J.L. A new model to estimate daytime net surface radiation under all sky conditions. *Atmos. Res.* **2025**, *315*, 107886. [CrossRef]
21. Li, S.; Jiang, B.; Liang, S.; Xiao, X.; Peng, J.; Liang, H.; Han, J.; Yin, X. Estimation of surface all-wave net radiation from MODIS data using deep residual neural network based on limited samples. *ISPRS J. Photogramm. Remote Sens.* **2025**, *225*, 131–143. [CrossRef]
22. Li, S.; Jiang, B.; Peng, J.; Liang, H.; Han, J.; Yao, Y.; Zhang, X.; Cheng, J.; Zhao, X.; Liu, Q.; et al. Estimation of the All-Wave All-Sky Land Surface Daily Net Radiation at Mid-Low Latitudes from MODIS Data Based on ERA5 Constraints. *Remote Sens.* **2022**, *14*, 33.
23. Hemmati, E.; Sahebi, M.R. Surface soil moisture retrieval based on transfer learning using SAR data on a local scale. *Int. J. Remote Sens.* **2024**, *45*, 2374–2406. [CrossRef]
24. Miao, S.; He, Q.; Zhu, L.; Yu, M.; Gu, Y.; Zhou, M. A Cross-Resolution Surface Net Radiative Inversion Based on Transfer Learning Methods. *Remote Sens.* **2024**, *16*, 2450. [CrossRef]
25. Jiang, Y.; Tang, W.; Yang, K.; He, J.; Shao, C.; Zhou, X.; Lu, H.; Chen, Y.; Li, X.; Shi, J. Development of a high-resolution near-surface meteorological forcing dataset for the Third Pole region. *Sci. China Earth Sci.* **2025**, *68*, 1274–1290. [CrossRef]
26. Zou, D.; Zhao, L.; Sheng, Y.; Chen, J.; Hu, G.; Wu, T.; Wu, J.; Xie, C.; Wu, X.; Pang, Q.; et al. A new map of permafrost distribution on the Tibetan Plateau. *Cryosphere* **2017**, *11*, 2527–2542. [CrossRef]
27. Ma, Y.; Xie, Z.; Chen, Y.; Liu, S.; Che, T.; Xu, Z.; Shang, L.; He, X.; Meng, X.; Ma, W.; et al. Dataset of spatially extensive long-term quality-assured land–atmosphere interactions over the Tibetan Plateau. *Earth Syst. Sci. Data* **2024**, *16*, 3017–3043. [CrossRef]
28. Zhu, L. *Meteorological Data of Surface Environment and Observation Network in China's Cold Region (2018)*; National Tibetan Plateau Data Center: Beijing, China, 2020.
29. Zhao, L.; Zou, D.; Hu, G.; Wu, T.; Du, E.; Liu, G.; Xiao, Y.; Li, R.; Pang, Q.; Qiao, Y.; et al. A synthesis dataset of permafrost thermal state for the Qinghai–Tibet (Xizang) Plateau, China. *Earth Syst. Sci. Data* **2021**, *13*, 4207–4218. [CrossRef]
30. Zhao, X.; Liang, S.; Liu, S.; Yuan, W.; Xiao, Z.; Liu, Q.; Cheng, J.; Zhang, X.; Tang, H.; Zhang, X.; et al. The Global Land Surface Satellite (GLASS) Remote Sensing Data Processing System and Products. *Remote Sens.* **2013**, *5*, 2436–2450. [CrossRef]
31. Tang, W.; Zhou, J.; Ma, J.; Wang, Z.; Ding, L.; Zhang, X.; Zhang, X. TRIMS LST: A daily 1 km all-weather land surface temperature dataset for China's landmass and surrounding areas (2000–2022). *Earth Syst. Sci. Data* **2024**, *16*, 387–419. [CrossRef]
32. Huang, G.; Liu, Z.; Pleiss, G.; Maaten, L.V.D.; Weinberger, K.Q. Convolutional Networks with Dense Connectivity. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 8704–8716. [CrossRef] [PubMed]
33. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359.
34. Ma, B.; Ma, Y.; Ma, W. Estimation of All-Weather Daily Surface Net Radiation over the Tibetan Plateau Using an Optimized CNN Model. *Remote Sens.* **2025**, *17*, 3894. [CrossRef]
35. Han, C.; Ma, Y.; Chen, X.; Su, Z. Estimates of land surface heat fluxes of the Mt. Everest region over the Tibetan Plateau utilizing ASTER data. *Atmos. Res.* **2016**, *168*, 180–190. [CrossRef]
36. Yao, J.; Zhao, L.; Gu, L.; Qiao, Y.; Jiao, K. The surface energy budget in the permafrost region of the Tibetan Plateau. *Atmos. Res.* **2011**, *102*, 394–407. [CrossRef]
37. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017. [CrossRef]

38. Tan, M.; Le, Q.V. EfficientNet Rethinking Model Scaling for Convolutional Neural Networks. *Int. Conf. Mach. Learn.* **2019**, *97*, 6105–6114. [CrossRef]
39. Lu, Z.; Pu, H.; Wang, F.; Hu, Z.; Wang, L. The Expressive Power of Neural Networks: A View from the Width. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
40. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
41. He, K.; Zhang, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
42. Wu, Z.; Shen, C.; van den Hengel, A. Wider or Deeper: Revisiting the ResNet Model for Visual Recognition. *Pattern Recognit.* **2019**, *90*, 119–133. [CrossRef]
43. Eldan, R.; Shamir, O. The Power of Depth for Feedforward Neural Networks. In *29th Annual Conference on Learning Theory*; JMLR: Cambridge, MA, USA, 2016; Volume 49, pp. 1–34.
44. Raghu, M.; Poole, B.; Kleinberg, J.; Ganguli, S.; Dickstein, J.S. On the Expressive Power of Deep Neural Networks. In Proceedings of the ICML'17: 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; Volume 70, pp. 2847–2854.
45. Zhu, L.; Dai, J.; Liu, Y.; Yuan, S.; Qin, T.; Walker, J.P. A cross-resolution transfer learning approach for soil moisture retrieval from Sentinel-1 using limited training samples. *Remote Sens. Environ.* **2024**, *301*, 113944. [CrossRef]
46. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
47. Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation System. *Bull. Am. Meteorol. Soc.* **2004**, *85*, 381–394. [CrossRef]
48. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **2021**, *13*, 4349–4383. [CrossRef]
49. Fan, X.; Lin, Z.; Niu, F.; Lan, A.; Yao, M.; Li, W. Near-surface heat transfer at two gentle slope sites with differing aspects, Qinghai-Tibet Plateau. *Front. Environ. Sci.* **2022**, *10*, 1037331. [CrossRef]
50. Weng, D.; Sun, Z.; Shi, B. The calculation and analysis of slope net radiation in China. *J. Meteorol. Res.* **1991**, *5*, 242–251.
51. Pan, Y.; Ren, L.; Xiang, X.; Huo, J.; Meng, D.; Wang, Y.; Yu, C.; Huang, Y. Effect of Slope Aspect on Plant Above- and Belowground Functional Traits of Alpine Meadow on the Qinghai-Tibet Plateau, China. *J. Geophys. Res. Biogeosci.* **2023**, *128*, e2022JG007268. [CrossRef]
52. Li, W.; Lin, Z.; Fan, X.; Yao, M.; Wu, X. Experimental study on the influence of embankment slope direction on near-surface thermal conditions in permafrost region, Qinghai-Tibet Plateau. *Case Stud. Therm. Eng.* **2023**, *49*, 103311. [CrossRef]
53. Ngah Onana, D.A.; Owona Atangana, P.B.; Mbuko Tcheutchoua, M.; Nyamsi, W.W. Assessing Satellite-Based Data Products Estimating Daily Means of Solar Irradiance at Surface over South Cameroon Plateau and Potential Improvements. *Remote Sens.* **2026**, *18*, 1434. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.