GIScience & Remote Sensing

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

Advancement of a diagnostic prediction model for spatiotemporal calibration of earth observation data: a case study on projecting forest net primary production in the mid-latitude region

Eunbeen Park, Hyun-Woo Jo, Gregory Scott Biging, Jong Ahn Chun, Seong Woo Jeon, Yowhan Son, Florian Kraxner & Woo-Kyun Lee

To cite this article: Eunbeen Park, Hyun-Woo Jo, Gregory Scott Biging, Jong Ahn Chun, Seong Woo Jeon, Yowhan Son, Florian Kraxner & Woo-Kyun Lee (2024) Advancement of a diagnostic prediction model for spatiotemporal calibration of earth observation data: a case study on projecting forest net primary production in the mid-latitude region, GIScience & Remote Sensing, 61:1, 2401247, DOI: 10.1080/15481603.2024.2401247

To link to this article: https://doi.org/10.1080/15481603.2024.2401247

Published online: 12 Sep 2024.

GISCIENCE & REMOTE SENSING
2024, VOL. 61, NO. 1, 2401247
https://doi.org/10.1080/15481603.2024.2401247

Advancement of a diagnostic prediction model for spatiotemporal calibration of earth observation data: a case study on projecting forest net primary production in the mid-latitude region

Eunbeen Park[a,b], Hyun-Woo Jo[a,b], Gregory Scott Biging[c], Jong Ahn Chun[d], Seong Woo Jeon[e], Yowhan Son[e], Florian Kraxner[a] and Woo-Kyun Lee[e]

[a]Agriculture Forestry and Ecosystem Services (AFE) Group, Biodiversity and Natural Resource (BNR) Program, International Institute for Applied System Analysis (IIASA), Laxenburg, Austria; [b]OJEong Resilience Institute (OJERI), Korea University, Seoul, Republic of Korea; [c]Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA; [d]Climate Analytics Department, APEC Climate Center, Busan, Republic of Korea; [e]Department of Environmental Science & Ecological Engineering, Korea University, Seoul, Republic of Korea

ABSTRACT

Developing a precise and interpretable spatiotemporal model is need for establishing evidence-based adaptation strategies on climate change-driven disasters. This study introduced a diagnostic prediction concept as a generalized modeling framework for enhancing modeling precision and interpretability and demonstrate a case study of estimating forest net primary production (NPP) in a mid-latitude region (MLR) by developing a diagnostic NPP diagnostic prediction model (DNPM). The diagnostic prediction concept starts with modeling meteorology and static environmental data, referred as a prognostic prediction part. Then, its outcome is refined with spatiotemporal residual calibration in the diagnostic prediction part, of which result undergo spatial, temporal, and spatiotemporally explicit validation methods. For the case of DNPM, a prognostic NPP prediction model (PNPM) was set, using a multilinear regression on SPEI 3, temperature, and static environmental features extracted from topography and soil by a random forest. Subsequently, during the diagnostic process of DNPM, we calibrated the primary outcome based on the temporal pattern captured at the time-series residual of PNPM. The results highlighted the superiority of the DNPM over the PNPM. Spatiotemporal validation showed that the DNPM achieved higher accuracy, with Pearson correlation coefficients (r) ranging from 0.975 to 0.992 and root mean squared error (RMSE) between 38.99 and 70.23 gC/m^2/year across all climate zones. Similarly, temporal validation indicated that DNPM outperformed the PNPM, with r values of 0.233 to 0.494 and RMSE of 46.01 to 70.75 gC/m^2/year, compared to the PNPM's r values of 0.192 to 0.406 and RMSE of 55.23 to 89.31 gC/m^2/year. This study showed enhanced diagnostic prediction concept can be applied to diverse environmental modeling approaches, offering valuable insights for climate adaptation and forest policy formulation. By accurately predicting various environmental targets, including drought and forest NPP, this approach aids in making informed policy decisions across different scales.

ARTICLE HISTORY
Received 20 March 2024
Accepted 2 September 2024

KEYWORDS
Diagnostic prediction model; Earth observation data; Spatiotemporal modeling; Net primary production; Mid latitude region

1. Introduction

The current global scenario is witnessing substantial changes in water cycle patterns, leading to an increase in hydrometeorological disasters and their associated impacts. The Intergovernmental Panel on Climate Change (IPCC) has highlighted that these changes will exacerbate extreme phenomena in the semi-arid and mid-latitude region (MLR), increasing the likelihood of natural disasters (IPCC et al. 2021). Natural disasters can occur anywhere, but the resulting damages vary depending on the capacity of regions and countries to predict, respond to, and recover from these events (FAO 2021). Therefore, to effectively adapt to and respond to these risks, it is necessary to provide policymakers with spatiotemporal assessments tailored to each country's circumstances under climate change (UN/ISDR 2008; CISA 2021). Environmental model should aim to reduce uncertainty and improve prediction accuracy while building interpretable models, ensuring that it can lead to effective policymaking and action plans.

With the advent of big data and artificial intelligence, the role of earth observation (EO) dataset in environmental modeling is becoming increasingly predominant. Remote sensing datasets provide valuable spatiotemporal information that can be effectively

CONTACT Woo-Kyun Lee leewk@korea.ac.kr

combined with various spatial and socioeconomic data to establish accurate decision-making criteria for risk situations such as extreme weather events and current environmental issues (Orusa, Cammareri, and Borgogno Mondino 2022). Especially, In the domain of forestry, satellite imagery is instrumental in monitoring forest dynamics, assessing deforestation and reforestation trends, and mapping and estimating forest biomass and net primary production (NPP) (Orusa and Borgogno Mondino 2021). Assessing environmental changes under climate and disaster risks requires a rapid, quantitative, and complex spatiotemporal methodology (Hassan 2018). Therefore, integration EO datasets enables the establishment of precise decision-making frameworks for addressing various environmental challenges, including extreme weather events and ongoing environmental issues.

Recently, there have been advancements have been made in the application of multiple-source datasets and data-fusion methods. E. Park et al. (2022) introduced an advanced data fusion method for drought prediction that reflects semantic and structural features of EO data, proposing a new diagnostic prediction concept. Unlike previous diagnostic models, this method uses remote sensing data to continuously calibrate regional errors in meteorological-based assessment. They developed a diagnostic drought prediction model for Kyrgyzstan, demonstrating its potential for monthly. However, they focused solely on monthly agricultural drought and did not provide detailed information on the comprehensive framework of general algorithms and data calibration within the diagnostic process. Additionally, the applicability of the diagnostic prediction model (DPM) to multi-time-series prediction targets has not yet been explored. Therefore, it is imperative to enhance the versatility of the DPM algorithm and extend its serviceability to a broad range of environmental predictions is imperative.

Forests are vital to the global carbon cycle, balancing greenhouse gases and carbon stocks in terrestrial ecosystems (Tang et al. 2010). Accurate quantification of forest production, measured using forest NPP, is essential for assessing a forest's capacity to provide ecosystem services (Kindermann et al. 2008; J. H. Park, Gan, and Park 2021; Yu, Chen, and Chen 2019). However, climate change-induced events and intricate interactions at the land surface level pose escalating challenges for quantifying forest NPP and assessing related risks (IPCC et al. 2021). Therefore, a precise estimation of forest resources should be preceded to prepare for future forest threats. Spatiotemporal modeling of forest biomass aids policymakers in developing climate adaptation strategies and provides valuable insights to local forest managers for implementing change and disaster risk reduction measures (Mickler, Earnhardt, and Moore 2002; Tang et al. 2010; Zhao and Running 2010). Various methods, including remote sensing such as Sentinel and MODIS sensors, process-based models such as Carnegie-Ames-Stanford Approach and BioGeoChemistry Management Model, and field surveys have been used to estimate global forest NPP and biomass across different scales (FAO (2020); United Nations Department of Economic and Social Affairs, United Nations Forum on Forests Secretariat 2021). As the spatiotemporal heterogeneity of forest NPP is driven by regional environmental factors and other complex interactions (Li et al. 2022), integrated modeling approaches are required for its estimation. Therefore, the development of new models and frameworks is essential for effectively estimating global NPP, enhancing forest sustainability, and promoting carbon sequestration (J. H. Park, Gan, and Park 2021). Developing data fusion methods that comprehensively integrate climate and observational data characteristics at the land surface level under predicted climate conditions is crucial for improving prediction accuracy and forest management strategies

Considering these research gaps, this study aimed to expand the theoretical concept of DPM and propose a general methodology. Specifically, we demonstrated the model's generalization by introducing methods for error prediction and calibrations within the diagnostic process of DPM. Furthermore, we conducted a case study to simulate forest NPP in MLR to access the applicability of DPM.

2. Concept of DPM

The DPM is an EO data fusion method to minimizes prediction errors by incorporating both semantic and structural features of forecastable climate and land observation data (E. Park et al. 2022). In proposing the DPM, they proposed the idea of prediction by continuously reflecting spatiotemporal regional errors through land observation into meteorological-based modeling, rather than simply adding land

observation variables to environmental modeling like previous diagnostic models. In this section, we newly introduce each step of DPM and specify functions that could be useful for estimating environmental targets.

The concept of DPM comprises two modeling steps: the prognostic prediction model (PPM) and the diagnostic process within the DPM. It has two primary characteristics. First, it captures the dynamic relationship between the most recently observed error and the prediction at the local land surface level. If fitting errors are treated as independent variables instead of omissions in a regression model, such as regression-kriging (Figure 1), the land surface information present in the error can be used (Hengl, Heuvelink, and Rossiter 2007; Piao et al. 2018). The DPM projects the latest fitting error onto the temporal axis and assumes spatiotemporal autocorrelation to predict how to differentiate the deviation of land surface environments from the general PPM prediction. This allows the calibration and optimization of future fitting errors based on temporal functions and current-fitting errors, respectively.

Therefore, we project the latest fitting error onto the temporal axis using the DPM (Figure 2) and predict the difference between the land surface environments and the general PPM prediction using the most recently observed patterns of future fitting errors. Diagnostic prediction calibrates the estimated next fitting errors of prognostic results as a function of time. Thus, the temporal equation is designed based on the characteristics of the prediction target, and the parameters are optimized using current fitting errors.

Second, the DPM differentiates between the features of meteorological impacts (which are universally applicable) and regionally varying fitting errors. Furthermore, the effects of climate change can be explained using the PPM by interpreting the effects of time-series meteorology and static variables, such as topography, which do not change over time. Regional fitting errors are differences reflecting the effects of infrastructure, policy, and other factors in the diagnostic process. These differences enhance the interpretability of each model and play a vital role in decision making regarding climate change adaptation.

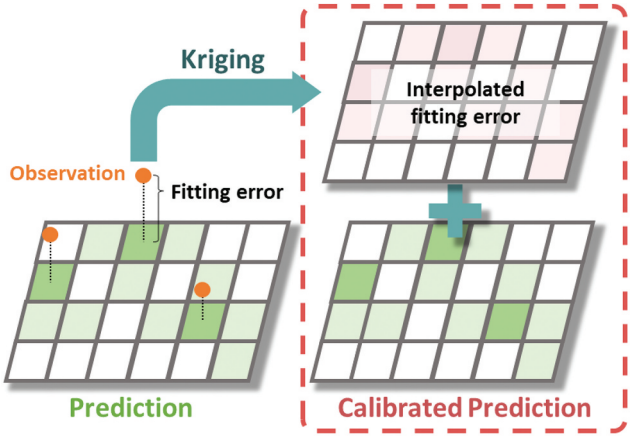

Figure 1. Concept of regression-kriging model.

The process of diagnostic prediction modeling, including prognostic prediction building and future fitting error estimation, varies depending on the prediction target, demonstrating the flexibility of the DPM (E. Park et al. 2022). However, the underlying principle is to develop a model with highly foreseeable variables and complement it with time-series monitoring. By continuously calibrating the model with the latest fitting errors, DPM improves the accuracy of future predictions over time. This calibration is particularly crucial because it allows the model to adapt to changes and reduce uncertainty in prediction, making it more reliable for long-term environmental forecasting. Based on these principles, a workflow for diagnostic prediction is proposed in this study (Figure 3). The workflow of diagnostic prediction comprises two models and validation methods. In Step 1, the PPM can be constructed using meteorological and static variables such as precipitation, temperature, topography, soil. In Step 2, the DPM diagnoses the predicted results from the PPM using land observation monitoring and subsequently calibrated them. Finally, in Step 3, the predicted result from each step can be compared with the corresponding land observations with respect to time and space using a spatiotemporal evaluation method.

2.1. Prognostic prediction model (PPM)

The PPM predicts future targets using the relationships between the meteorological variables and the target. Static environmental variables, such as topography and soil, can be also considered depending on the characteristics of the target. Land observation data, particularly satellite data, is used to extract most of the features of the prediction target. Therefore, this study generalized the prediction target of the PPM as a satellite-based index (*SI*). The *SI* at

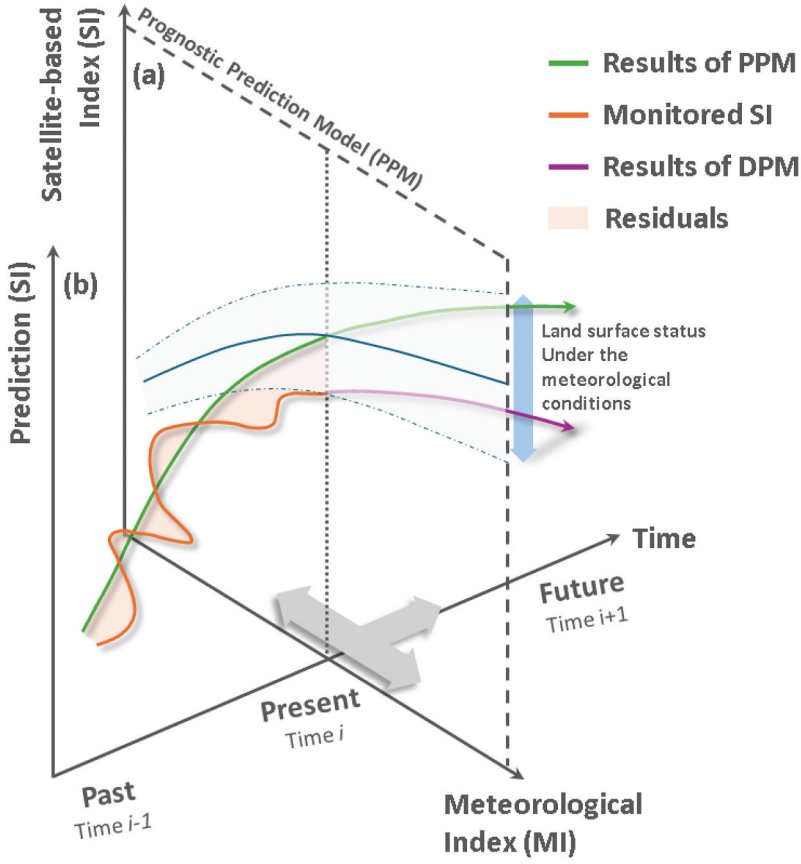

Figure 2. Generalized concept of diagnostic prediction model (DPM) (revised from E. Park et al. 2022). (a) *Meteorological Index* (MI) – *Satellite based Index* (SI) space illustrating a trace of the prognostic prediction model (PPM) through the time axis under general conditions, although the monitored *SI* may differ from the PPM under specific meteorological conditions. (b) *time* – *SI* space showing the most recent fitting errors (residuals) of the PPM. The DPM process diagnoses the current residuals and calibrates the next fitting errors of the PPM.

time i is calculated using the function of the PPM (f_{prog}) with the meteorological index (MI) and static variables (V_s), whereas the remaining variation is represented by a fitting error (e_i^{prog}) (Equation 1). The PPM provides a simple and intuitive way to understand the impacts of climate change on targets, thereby enabling short-to long-term predictions.

$$SI_i = f_{prog}(MI_i, V_s) + e_i^{prog} \quad (1)$$

where i = time.

2.2. Diagnostic prediction model (DPM)

The DPM uses the current fitting error to minimize future fitting errors for reducing the uncertainty in the PPM. Although the PPM is effective for large-scale projections based on long-term meteorological changes, it inevitably comprises a fitting error (e_i^{prog}) owing to the exclusion of certain regional environmental conditions. Therefore, when considering surface-level environmental factors, such as geography and anthropogenic activities, the actual environmental response can be more heterogeneous than the PPM projection. However, considering that the impact of climate change globally is unequal across regions, the purpose of environmental modeling for climate change adaptation is to identify extreme environmental changes far from normal, which are termed as disasters. The DPM estimates the impact of the excluded variables by diagnosing the current environment using land observations at a certain time i and assuming the temporal autocorrelation of environmental factors (Shahin, Ali, and Ali 2014; Shamsnia et al. 2011), which is subsequently applied to the next time step ($i + 1$) of the PPM. The DPM specializes in short- and mid-term predictions and enhances the interpretability of models for decision-making in the climate crisis. Additionally, the process of diagnosing the current environment and

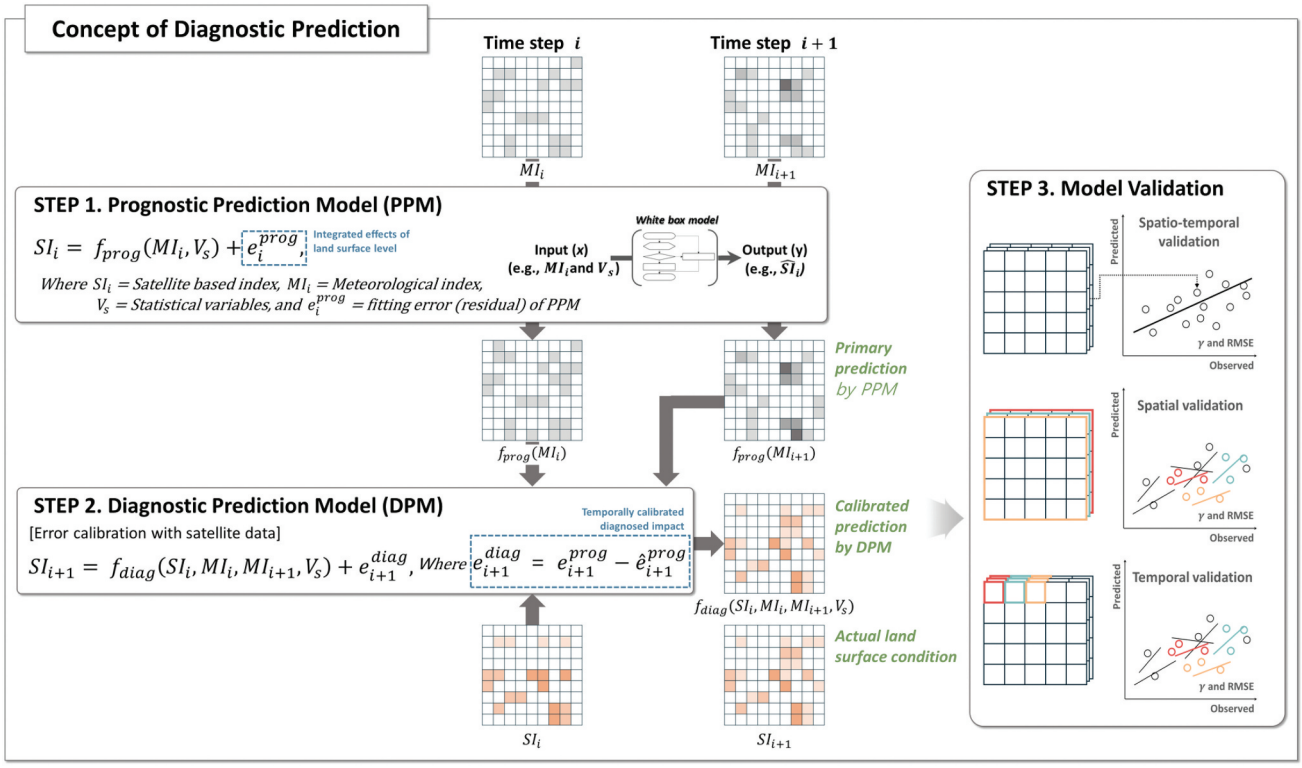

Figure 3. Workflow of diagnostic prediction modelling for the synergetic use of each earth observation data. $f_{prog}(x)$ and $f_{diag}(x)$ represent the functions of the prognostic and diagnostic prediction models, respectively. *MI*and*SI* represent the meteorological and satellite-based indices, respectively. V_s *and* e represent static variables and fitting errors (residuals), respectively.

applying it to the next time step varies based on the characteristics of the prediction target.

The DPM incorporates current and future meteorological variables, most recent land observations (SI_i), and static variables (Equation 2). Furthermore, it comprises a function of calibration (f_{calib}) that optimizes the parameters for the error prediction (f_e) using the most recent land observations (SI_i) at time i (Equation 3). Subsequently, it predicts the future fitting error ($\hat{e}_{i+1}^{prog}$) at time $i+1$using f_e, which calibrates the PPM (Equation 4). The fitting error of the DPM reflects the predicted error calculated from, aiming to be lower than the PPM's error in most cases (Equation 5). The DPM minimizes fitting errors by identifying non-meteorological effects using satellite information and subsequently incorporating them into future predictions.

$$SI_{i+1} = f_{diag}(MI_i, SI_i, MI_{i+1}, V_s) + e_{i+1}^{diag} \quad (2)$$

$$= f_{prog}(MI_{i+1}, V_s) + e_{i+1}^{prog} + f_{calib}(SI_i, MI_i) \quad (3)$$

$$= f_{prog}(MI_{i+1}, V_s) + e_{i+1}^{prog} - \hat{e}_{i+1}^{prog} \quad (4)$$

$$e_{i+1}^{diag} = e_{i+1}^{prog} - \hat{e}_{i+1}^{prog} \lessapprox e_{i+1}^{prog} \quad (5)$$

2.3. Spatiotemporal assessment

Accuracy in spatiotemporal dimensions is crucial for validating the performance of the proposed model. The prediction results of the PPM and DPM can be compared with the monitored land observations using specific aggregate window units which are set as a proper valid unit in a study area. Comparing results at scales that are extremely small or large can cause an underestimation or imprecise estimation of model performance. The root mean squared error (RMSE) and Pearson correlation coefficient (r) can be used to identify the trends and absolute errors in the prediction results (Equations 6 and 7). These metrics can be calculated for the entire study period to assess the overall performance in spatiotemporal dimensions (Figure 4a). For a comprehensive evaluation of model performance, additional spatiotemporal validations were considered depending on the characteristics of the target. To emphasize the spatial dimension, the

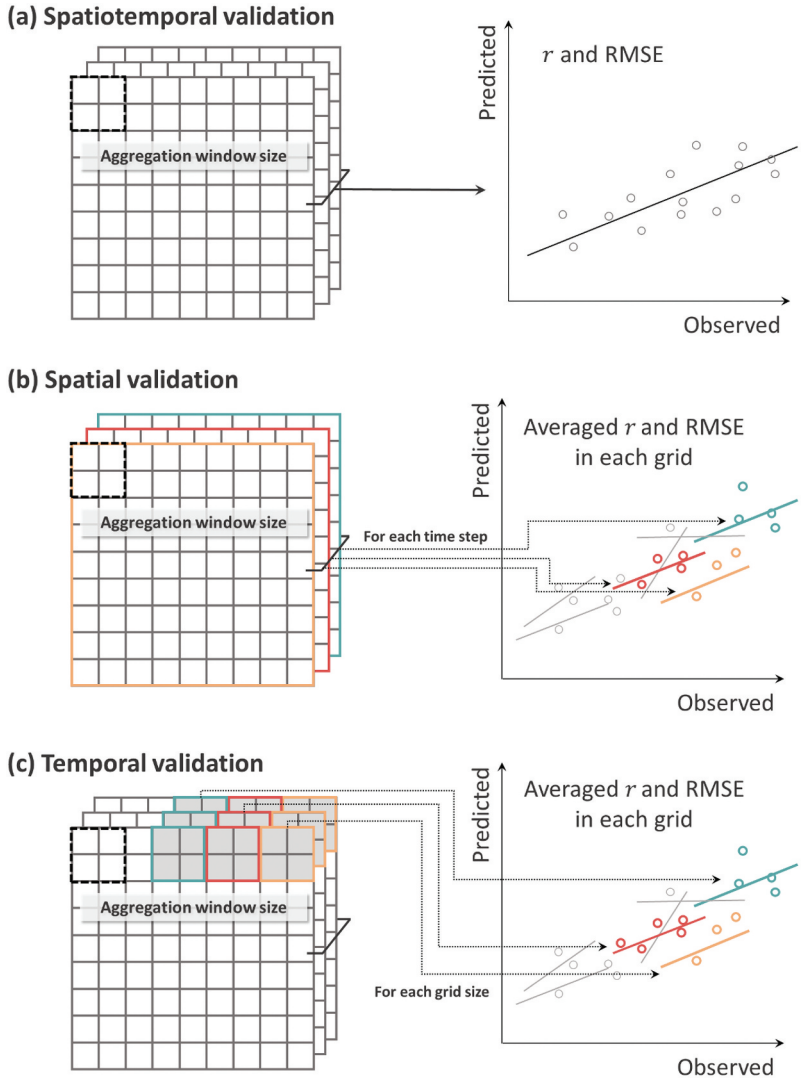

Figure 4. Multi-dimension assessment of model performance. (a) Spatiotemporal, (b) spatial, and (c) temporal validations.

aggregated values within the same time step can be presented as a line. Each line represents a specific area, as shown in Figure 4b. The validation indices were calculated and averaged for each line. Similarly, to emphasize the temporal dimension, the aggregated values within a specific grid unit over the entire study period are presented as a line, as shown in Figure 4c. Each line represents the time step within a grid unit. The validation indices were calculated and averaged for each line.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left((\hat{y}_i - y_i)^2\right)} \tag{6}$$

where $\hat{y}_i$ and y_i are predicted and observed values, respectively.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{7}$$

where x_i and y_i are the variables, and $\bar{x}$ and $\bar{y}$ are the mean values of each variable.

3. Materials

To develop a PNPM and DNPM in MLR, Meteorological data and Land observation data were utilized within the diagnostic prediction framework. Additionally, topographic and soil data were used to reflect the static environmental characteristics of NPP, and forest land cover and climate zone information were used as supplementary data (Table 1).

Table 1. List of data used in this study.

Dataset	Period	Original resolution	Aggregated resolution	Type of data	Source
Temperature	2000–2018 (Monthly)	2.5 arc minutes	4 km	Raster	WorldClim (Fick and Hijmans 2017)
Precipitation	2000–2018 (Monthly)	2.5 arc minutes	4 km	Raster	WorldClim (Fick and Hijmans 2017)
SPEI 3	–	–	4 km	Raster	Temperature and precipitation
DEM	–	500 m	4 km	Raster	CGIAR-CSI SRTM (Jarvis et al. 2008)
Slope	–	–	4 km	Raster	DEM
Aspect	–	–	4 km	Raster	DEM
TWI	–	–	4 km	Raster	DEM
AWC	–	1:5,000,000	4 km	Vector	FAO (Fischer et al. 2008)
TBD	–	1:5,000,000	4 km	Vector	FAO (Fischer et al. 2008)
NPP	2000–2018 (Yearly)	500 m	4 km	Raster	MODIS
Land cover map	2018	500 m	4 km	Raster	MODIS
Aridity index	–	30 arc-second	4 km	Raster	CGIAR-CSI (Trabucco and Zomer 2019)

3.1. Study area

The MLR lies between 30° N and 60° N in the Northern Hemisphere and comprises forests, barren land, grassland, savannas, and sparse vegetation. It is extremely vulnerable to climate change and has limited environmental resources owing to degradation, sand and dust storms, and droughts. Furthermore, it comprises approximately 50% of the global population and many developing countries. Therefore, concerns have arisen over its declining agricultural and forest production due to extreme seasonal variations, rising temperatures, and declining precipitation.

This study focused on MLR forests located between 18.89° N and 56.22° N latitudes and 135.79° W and 115.44° E longitudes. We used the 2018 MODIS Land Cover Map and extracted forest areas at 500 m resolution, which were subsequently aggregated to a mean value at 4 km resolution for model development (Figure 5).

3.2. Meteorological data

3.2.1. Historical climate data

Time-series precipitation and temperature were obtained from the WorldClim database (http://www.worldclim.org/). WorldClim provides historical climate data between 1960 and 2018 at 2.5 minutes spatial resolution (approximately 4.5 km from the equator) (Fick and Hijmans 2017). We collected monthly precipitation and temperature datasets at 4 km

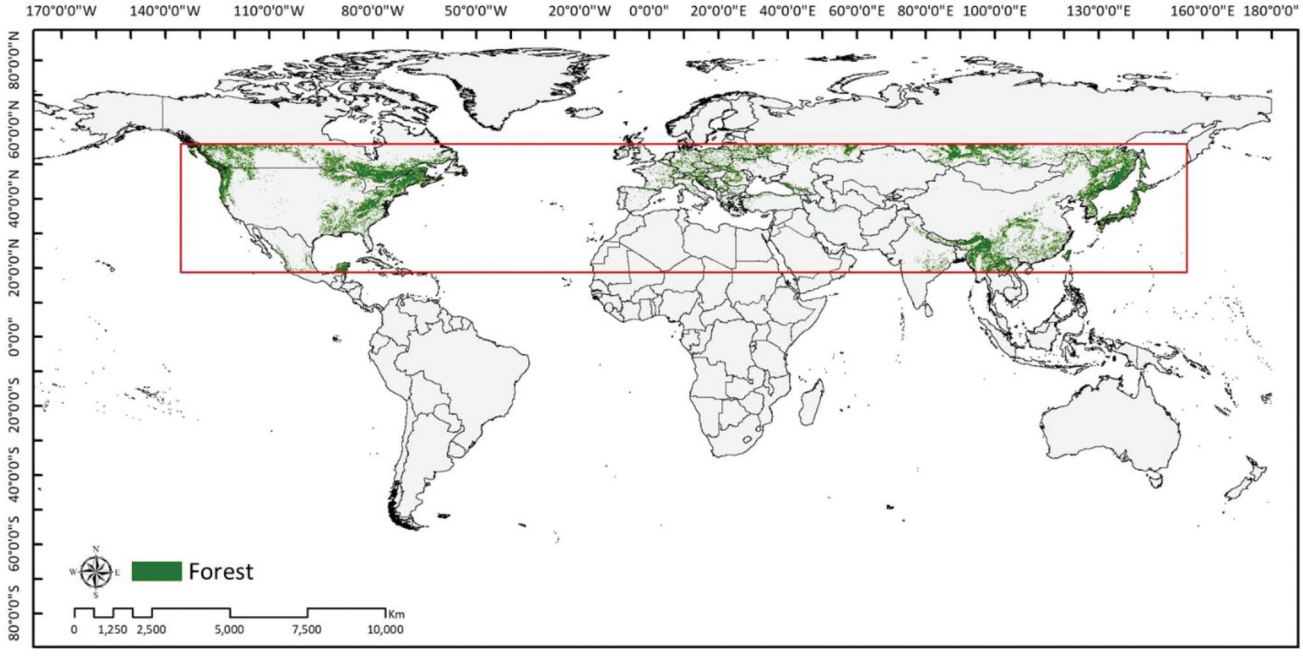

Figure 5. Study area and forests in the mid-latitude region (MLR) (EPSG: 54008).

resolution from January 1980–December 2018 to produce a meteorological drought index.

3.2.2. Standardized precipitation evapotranspiration index (SPEI)

The SPEI is a meteorological drought index that uses both Precipitation (P) and Potential Evapotranspiration (PET) to assess drought conditions in terrestrial ecosystems. It captures temperature variability; therefore, it is an important parameter for evaluating the impacts of meteorological variables on land surfaces (Chen et al. 2013; Vicente-Serrano, Beguería, and López-Moreno 2010). To calculate SPEI, the accumulated differences between precipitation and PET at different time scales (e.g. 1, 3, 6, 9, 12, or 24 months) (Equation 8) are fit to a selected distribution function, such as the log-logistic, gamma, or Pearson III distribution. The cumulative probability of the fitted distribution is then transformed using the inverse of the standard normal distribution to produce the SPEI values. Additionally, a minimum of 30 years of high-quality climate data is required to ensure the statistical significance of SPEI calculation. In this study, we produce the three-month SPEI (SPEI 3) from 1980 to 2018 by calculating PET (Thornthwaite 1984) and fitting a log-logistic distribution (Equation 9) using Deep learning & Remote sensing analYsis for Agroforesty and Drought (DRYAD) software (Jo and Lee 2024) (https://zenodo.org/doi/10.5281/zenodo.10078913) (Equation 8 and 9). To integrate with NPP from MODIS, we selected the target period from 2000 to 2018 for model development. SPEI values greater than 1.0 indicated wet conditions, whereas values less than −1.0 indicated dry conditions.

$$D^k = \sum_{i=0}^{k-1} P_{n-i} - PET_{n-i} \tag{8}$$

where n is calculation frequency and k is time scale.

$$f(x) = \frac{\beta}{a}\left(\frac{x-y}{a}\right)^{\beta-1}\left(1\left(\frac{x-y}{a}\right)^{\beta}\right)^{-2} \tag{9}$$

where α, β and γ are scale, shape and origin parameters, respectively, for D values in the range ($\gamma > D < \infty$).

3.3. Land observation data

3.3.1. Time-series NPP

Satellite-based NPP estimation is the most suitable approach for areas lacking adequate data or encompassing a wide range. Time-series NPP datasets were acquired from the MOD17A3HGF Version 6 (006) product of MODIS and prepossessed within Google Earth Engine (https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MOD17A3HGF). The MOD17A3HGF provides annual NPP data from 2000 to present at a spatial resolution of 500 m near the equator. In this study, we used annual NPP datasets from 2000 to 2018 as the dependent variable (prediction target) for forest production. We extracted by forest land cover, which were subsequently aggregated it to a mean value at 4 km resolution.

3.3.2. Static environmental dataset

In this study, static environmental datasets, comprising soil and topographic information, were used to extract static environmental features in the developing process of PPM. The Harmonized World Soil Database (HWSD) v 1.2 was obtained from the Food and Agriculture Organization (https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/). The HWSD is a global soil database that contains over 16,000 different soil properties (sol units, pH, texture, etc.) with 1 km resolution (Fischer et al. 2008). In this study, the Available Water storage Capacity (AWC) and Topsoil Bulk Density (TBD) were extracted by forest land cover, and these values were subsequently aggregated to a mean value at 4 km resolution.

To acquire topographic characteristics, the Digital Elevation Model (DEM) data was obtained from the CGIAR-CSI Shuttle Radar Topography Mission (SRTM) dataset at a resampled resolution of approximately 500 meters (Jarvis et al. 2008). (https://csidotinfo.wordpress.com/data/srtm-90m-digital-elevation-database-v4-1/). Due to its uniform resolution and accuracy over most of the Earth's surface, SRTM DEM has been widely used in the environmental modeling and monitoring. In general, DEM, Slope, and Aspect factors are utilized in vegetation modeling along with climate data, mainly to reflect regional differences. Additionally, we calculated the Topographic Wetness Index (TWI) based on general algorithms described by Sørensen, Zinko, and Seibert (2006), which combines catchment area and local slope gradients to assess topographical moisture levels and their influence on vegetation growth (Beven and Kirkby 1979; Kopecky and Cízková 2010). In this study, Slope, Aspect, and TWI were produced based on the DEM and All

variables were extracted by forest land cover, and subsequently aggregated it to a mean value at 4 km resolution.

For climate classification, Global Aridity Index, based the FAO −56 Penman Monteith equation, was obtained from the CGIAR CSI database version 3 (Trabucco and Zomer 2019). The aridity index represents the function of mean annual precipitation and mean annual potential evapotranspiration for averaged 1970–2000 period at 30 arc second spatial resolution. In this study, the index was categorized subforest areas into arid, semi-arid, dry sub-humid, and humid climate zones development for distinguish climate zone for model.

4. Methods

Figure 6 illustrates the DPM workflow for simulating forest NPP. The Prognostic NPP Prediction Model (PNPM) comprised two steps. It was developed as a multilinear regression analysis using climatic variables (SPEI 3 and annual temperature) and static environmental features. Static environmental features were extracted using a Random Forest (RF) model to analyze the impact of complex topographies on forest NPP.

We used a calibration function (see Equation 12) based on the characteristics of the fitting error time-series of the PNPM in the diagnostic process of the Diagnostic NPP Prediction Model (DNPM). The error prediction function was applied as a linear regression, which re-optimized the parameters based on the relationship between the predicted PNPM and the observed NPP results at the current time, to predict future fitting errors. We used the observed meteorological variables instead of forecasts for the next time step during the relative time-series analysis of the PNPM.

4.1. PNPM

The PNPM predicts forest production (NPP) using meteorological variables (SPEI 3 and temperature). The PNPM employs static environmental features and meteorological variables as independent variables to explain the impact of meteorology on the land surface. The development of the PNPM consisted of two steps.

First, an RF model (Breiman 2001) was used to extract static environmental features that have substantial impacts on forest NPP and remain relatively constant over time (Field, Randerson, and Malmström 1995; Running, Hunt, and Ehleringer 1993). To extract average characteristics of complex influence of topography and soil properties over the entire study period, the mean forest NPP from 2000 to 2018 was used as the dependent variable. The extracted features were then used as independent variables in the to improve the accuracy and understanding of the PNPM.

Second, multilinear regression was used to integrate climatic variables and extracted static environmental features into the PNPM. Regression analysis was used to determine the parameters (b_0–b_{14}), and an error term (e) was used to analyze the relationship between these variables and the NPP (Equation 10). The model explained NPP variation using simultaneous changes in meteorological conditions. Since NPP distribution varies depending on climate classification (Chen et al. 2013), the PNPM was developed for different climate zones, including arid, semi-arid, dry sub-humid, and humid zones.

$$NPP_i = f_{prog}(SPEI\ 3_i, Temp_i, V_s) + e_i^{prog} \tag{10}$$

$$= b_0 + b_1 SPEI\ 3_{i,Jan.} + \ldots + b_{12} SPEI\ 3_{i,Dec.} + b_{13} Temp_i + b_{14} V_s + e_i^{prog},$$

where $SPEI\ 3_i, Temp_i, V_s, and\ e_i^{prog}$ are yearly monthly *SPEI* 3, annual temperature, static environmental features, and fitting error of PNPM respectively, at time i.

4.2. DNPM

The DNPM aims to minimize the uncertainty and reflect the current land surface conditions to improve the accuracy and precision of the PNPM by using residuals. Although the PNPM partially incorporates regional static environmental features and meteorology, there are still residuals that were not included as independent variables. Consequently, the DNPM (Equation 11) uses a calibration function (Equation 12) to optimize the error prediction function and compensate any future fitting errors of the PNPM (Equations 13). The calibration function optimizes the error prediction based on the observed NPP at time i and returns the predicted fitting errors ($\hat{e}_{i+1}^{prog}$) at the next time $i + 1$ (Equation 14). Subsequently, the

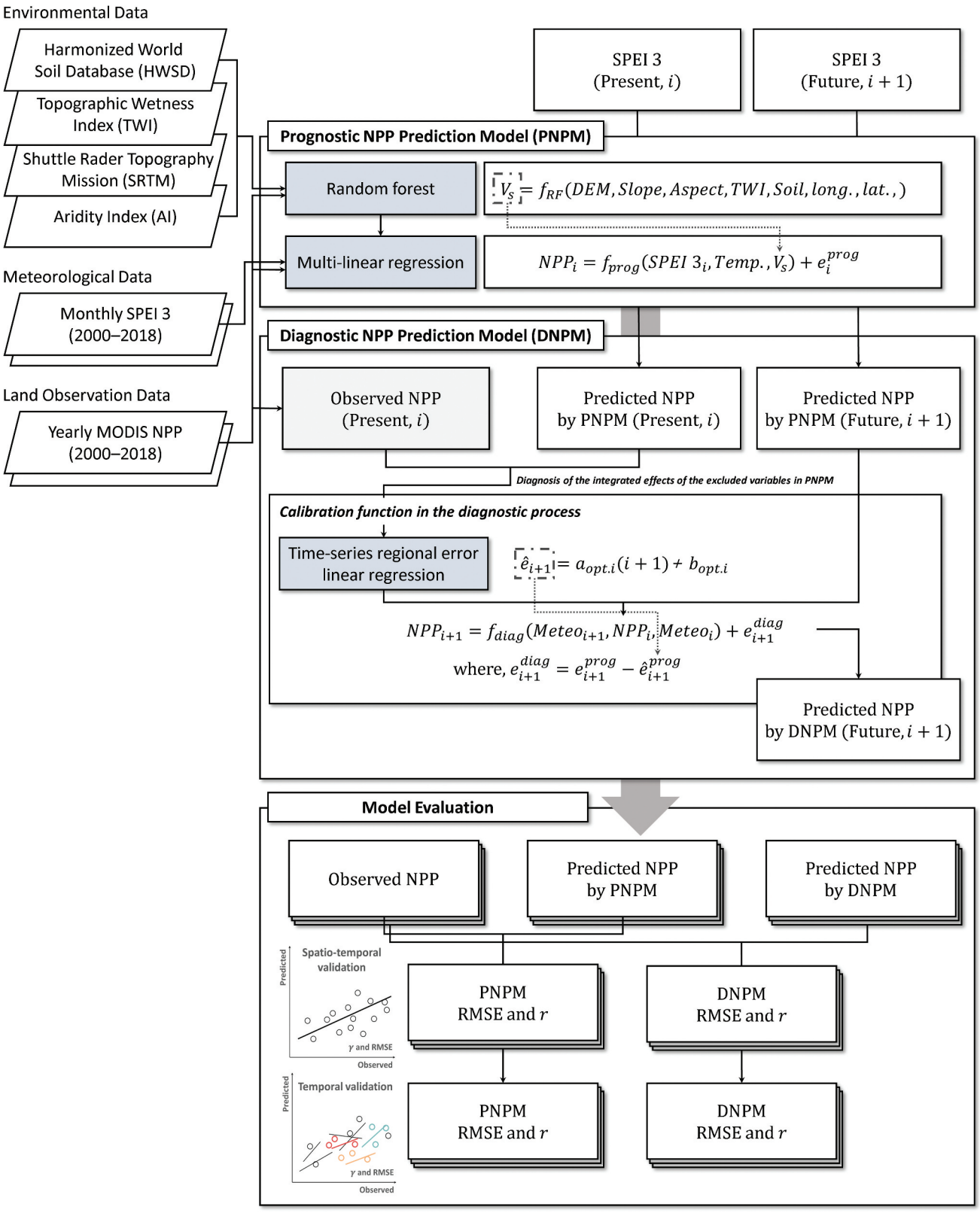

Figure 6. Workflow of the detailed DPM application for forest NPP estimation MLR.

DNPM compensates for the PNPM's future fitting error at the next time $i+1$ (Equation 13). The error prediction function used time-series residuals from 2000 to 2010 as a baseline and applied linear regression (Equation 15). During the diagnostic process, the slope and intercepts were adjusted based on the

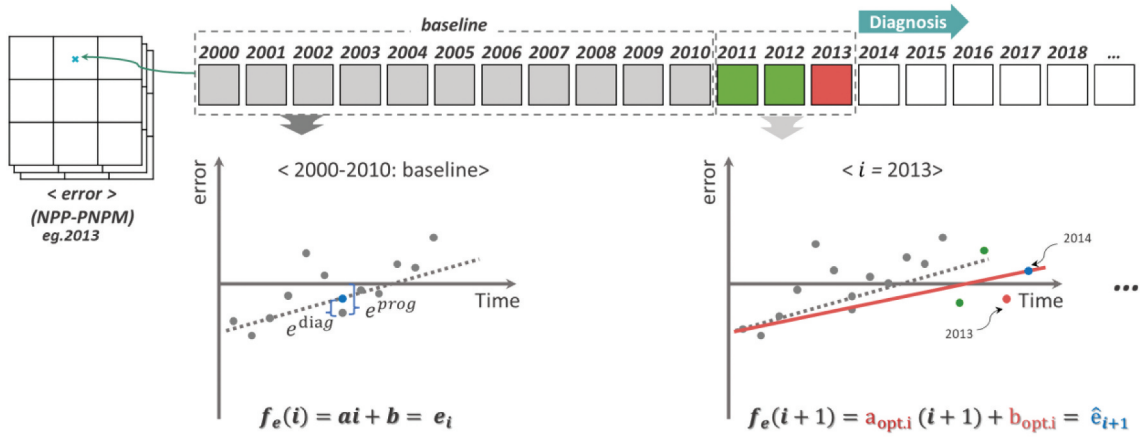

Figure 7. The process of calculating prediction error in the diagnostic process of the DNPM. Time-series fitting errors from 2000–2010 (gray box) are used. As new years are added (red and green boxes for new and previous years, respectively), the slope and intercept of the error prediction function are optimized for each year *i*, thereby predicting the fitting error value for the next year (blue point).

data added each year to forecast future fitting errors for the next period (Figure 7).

$$NPP_{i+1} = f_{diag}(Meteo_i, NPP_i, Meteo_{i+1}) + e^{diag}_{i+1} \quad (11)$$

$$= f_{prog}(Meteo_{i+1}) + e^{prog}{}_{i+1} + f_{calib.}(Meteo_i, NPP_i) \quad (12)$$

$$= f_{prog}(Meteo_{i+1}) + e^{prog}_{i+1} - \hat{e}^{prog}_{i+1}, \quad (13)$$

where $Meteo_i$ is refers $SPEI\ 3_i$, $Temp_i$, and V_s.

$$f_{calib}(Meteo_i, NPP_i) = 1)\ optimization\ of\ f_e\ with\ NPP_i \quad 2)\ and\ returen\ \hat{e}^{prog}_{i+1} \quad (14)$$

$$f_e(i+1) = a_{opt.i}(i+1) + b_{opt.i} = \hat{e}^{prog}_{i+1}, \quad (15)$$

where $a_{opt.i}$ and $b_{opt.i}$ are the adjusted slope and intercept of NPP_i at time i, respectively, and $\hat{e}^{prog}_{i+1}$ is the predicted fitting error.

4.3. Model validation in spatiotemporal dimensions

The developed models (PNPM and DNPM) were evaluated by comparing the annual predicted NPP values, which were aggregated to a 100 km^2 grid, with the observed NPP values in each climate zone. Spatial differences between the PNPM and the DNPM were not substantial owing to the inclusion of topographic soil characteristics in the PNPM. Furthermore, additional temporal validation was conducted to highlight the impact of the DNPM. Model performance was assessed using *r* and RMSE, which were calculated at each grid unit, to emphasize the temporal dimension.

5. Results

5.1. Development and validation of PNPM and DNPM

The PNPM was developed using multilinear regression based on meteorological variables and static environmental features extracted through the RF model. Initially, during PNPM development, the RF model achieved a coefficient of determination (R^2) and RMSE values of 0.96 and 51.55 gC/m^2/year. The important features identified by RF were consistent with previous studies, indicating predominant correlation among forest types, elevation, latitude, longitude, and production (Gillman et al. 2015; Kim et al. 2020; Pan et al. 1996). As a result of PNPM, predicted NPP values expanded gradually as the model shifted from the Arid to Humid zone. In the Arid zone, the values were relatively low with a narrow range, while the Humid zone encompassed a wider range of values. This difference is attributed to the majority of MLR forests being located in the Humid zone, where diverse environmental conditions support varying productivity levels. Conversely, the smallest forested areas are found in the Arid zone, characterized by high temperatures and low rainfall, limiting forest productivity. Additionally, as the distribution shifts from Arid to Humid zones, forest types transition from dry to tropical. Subsequently, the predicted NPP values from 2000 to 2018 by PNPM were calibrated using observed MODIS NPP during the same period based on DNPM's diagnostic process.

To assess the spatiotemporal error calibration effects of the DNPM, the model performance of both DNPM and PNPM was validated across spatiotemporal and temporal dimensions. Spatiotemporal evaluation results, comparing the entire year and grid-based assessment simultaneously (indicated by red lines), demonstrated high predictive performance for both PNPM and DNPM. This reflects PNPM's comprehensive development incorporating meteorological and static environmental features crucial for general NPP prediction. However, DNPM showed improved model performance by calibrating time-series PNPM's residuals obtained from satellite observation. DNPM exhibited improved validation indices across all climate zones, with *r* values of 0.98, 0.99, 0.98, and 0.98, and RMSE values of 70.27, 65.90, 47.63, and 38.99 gC/m^2/year, respectively, compared to the *r* and RMSE values of the PNPM (0.97, 0.95, 0.96, and 0.96 and 94.21, 77.24, 65.88, and 63.11 gC/m^2/year, respectively). DNPM significantly improved detailed regional errors in extreme climate zones like Arid and Humid.

Temporal evaluation results revealed that while the PNPM represented horizontal time-series patterns within the same region (illustrated by the blue lines) with *r* values of 0.33, 0.41, 0.36, and 0.19, and RMSE values of 89.31, 63.79, 57.23, and 55.23 gC/m^2/year, respectively in Figures 8a–, DNPM displayed relatively upward patterns with low variations (depicted in Figures 8f–), characterized by *r* values of 0.34, 0.49, 0.41, and 0.23, and RMSE values of 70.75, 46.01, 48.52, and 50.52 gC/m^2/year, respectively. These improvements in the evaluation indices were particularly pronounced in the semi-arid and dry sub-humid zones, as shown in Figures 8b,f,,, respectively. This can be attributed to the higher sensitivity to climate change and significant land surface changes in these regions, which are relatively more responsive to temporal variations within the same grid. In contrast, the improvements were relatively modest in the arid and humid zones. In the arid zone, degraded and low vegetation health primarily caused consistently low production, regardless of environmental conditions, resulting in relatively low improvement. The humid region, characterized by a concentration of diverse forest structures in the MLR, exhibited relatively lower temporal improvement. This is likely due to the mix of multiple age structures, tree species, and site productivity within the 100 km of spatial validation grid size. Compared with PNPM, DNPM demonstrated relative performance improvements, underscoring the potential of residual calibration to enhance prediction accuracy across all dimensions.

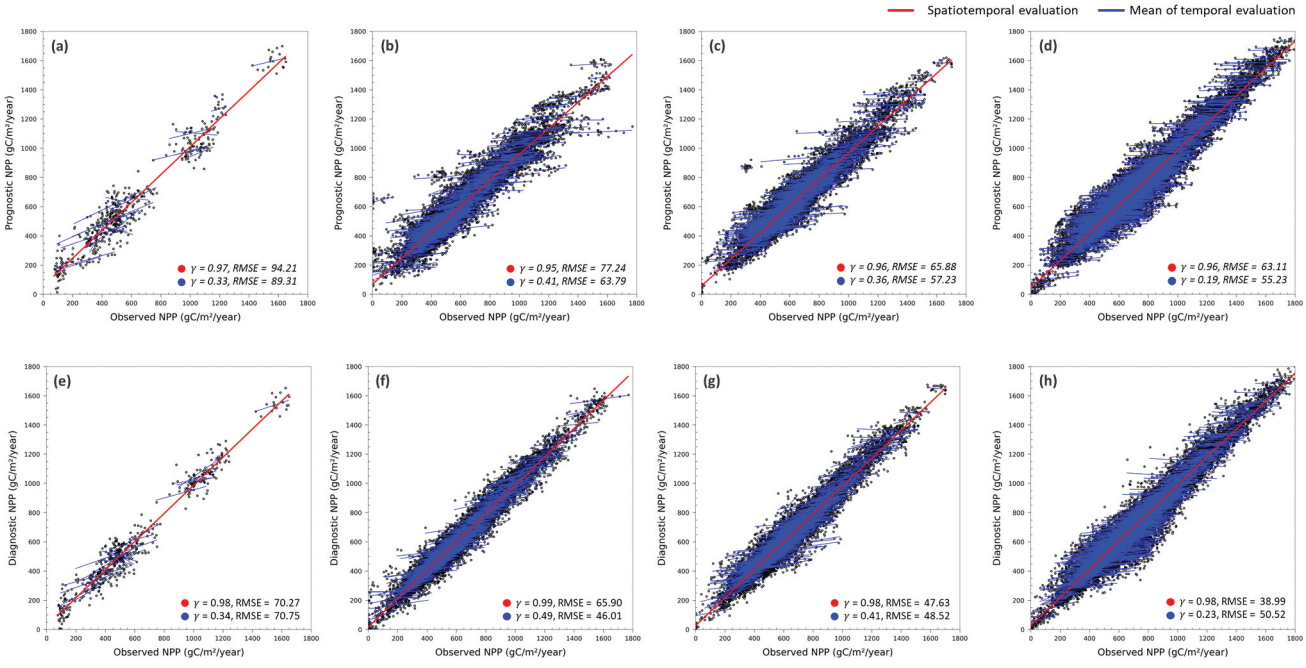

Figure 8. Comparison of the spatiotemporal prediction patterns for the entire period (red line) and temporal prediction for a specific region (blue line) using the prognostic NPP prediction model (PNPM; (a) arid, (b) semi-arid, (c) dry sub-humid, and (d) humid), diagnostic NPP prediction model (DNPM; (e) arid, (f) semi-arid, (g) dry sub-humid, and (h) humid) and observed NPP.

5.2. Calibration effects and characteristics of DNPM

Figures 9–12 compare the predicted NPP using the PNPM and DNPM and observed NPP in MLR at both regional and local scales. At the regional scale (large scale), the prediction from both the PNPM and DNPM were similar spatial pattern to the observed NPP (Figure 9). PNPM, reflecting the empirical relationships between static environmental features and meteorological variables, generally estimates well and shows similarity to the observed NPP over the same period. However, at the zoomed-in local scale (Figures 10a, 11a, and 12a), PNPM occasionally showed over- or under-estimations which constrained the detection of detailed differences. This is due to smooth patterns of meteorological properties and broad spatial autocorrelation, particularly observed in the Korean Peninsula and Japan, which can reduce PNPM's spatial accuracy during model evaluation.

In contrast, the DNPM effectively performed at the local scale by calibrating regional differences using the latest satellite images based on the PNPM predictions (Figures 10b–12b). Although DNPM still showed an underestimated result in some regions (Figure 12b), it showed the potential DNPM's ability by calibrating residual patterns closer to the observed value. This is because DNPM learned the fitting error patterns of the PNPM and calibrated them for future predictions.

The DPM's effectiveness lies in its ability to adjust for regional to local differences between expected impacts from the PPM and the latest land observations, especially noticeable at the small scale, particularly in extreme weather or scenarios with varied vegetation responses under similar climates (Figure 4.13). These findings highlight the unique advantages of the DPM. The PNPM demonstrated broad-scale prediction areas under general meteorological condition, learned from historical weather events, with relatively limited coverage in extreme conditions that are difficult to predict in the extreme future condition, Conversely, The DNPM's sporadic enhancements introduced spatiotemporal variation to the PNPM

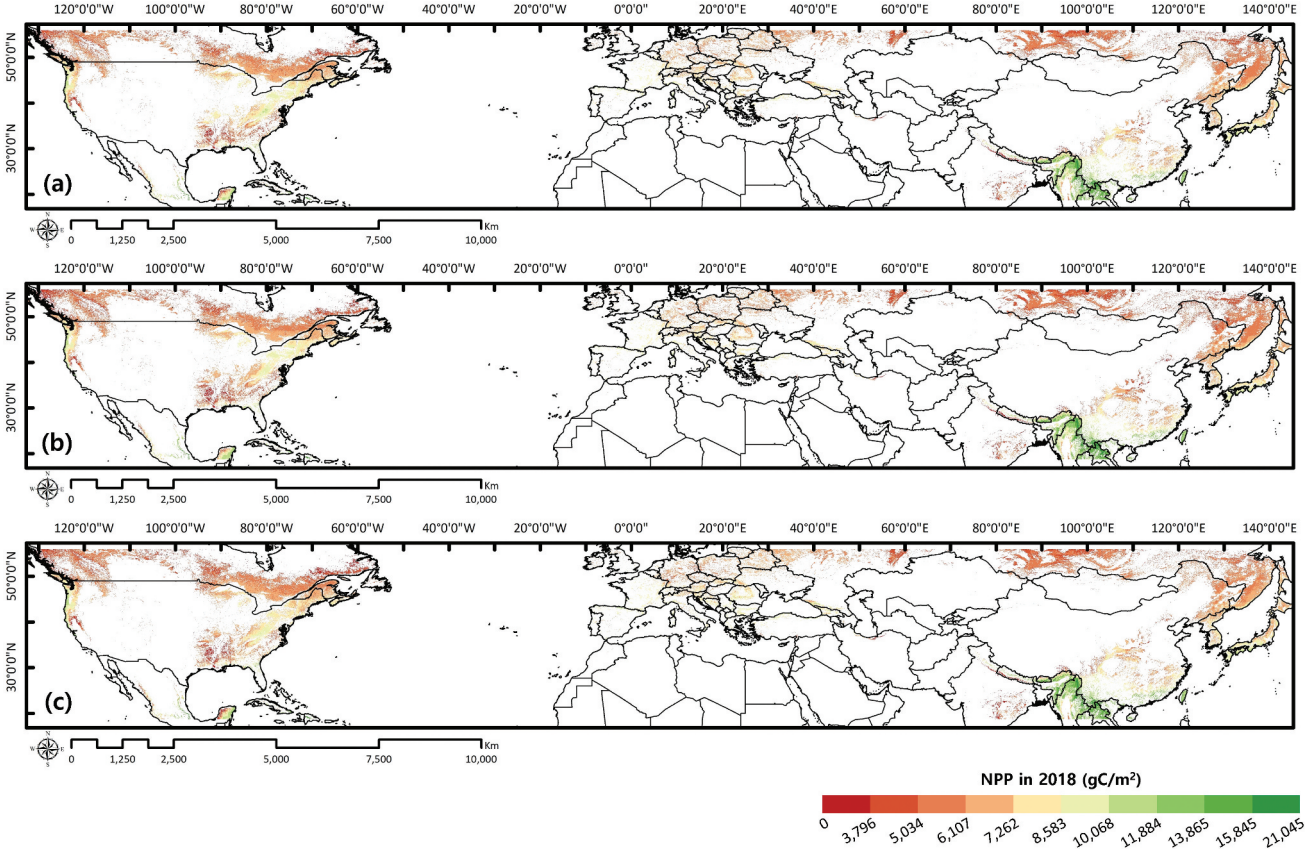

Figure 9. Comparison between the observed and predicted NPP on the regional scale for 2018. (a) PNPM NPP, (b) DNPM NPP, and (c) observed NPP (EPSG: 54008).

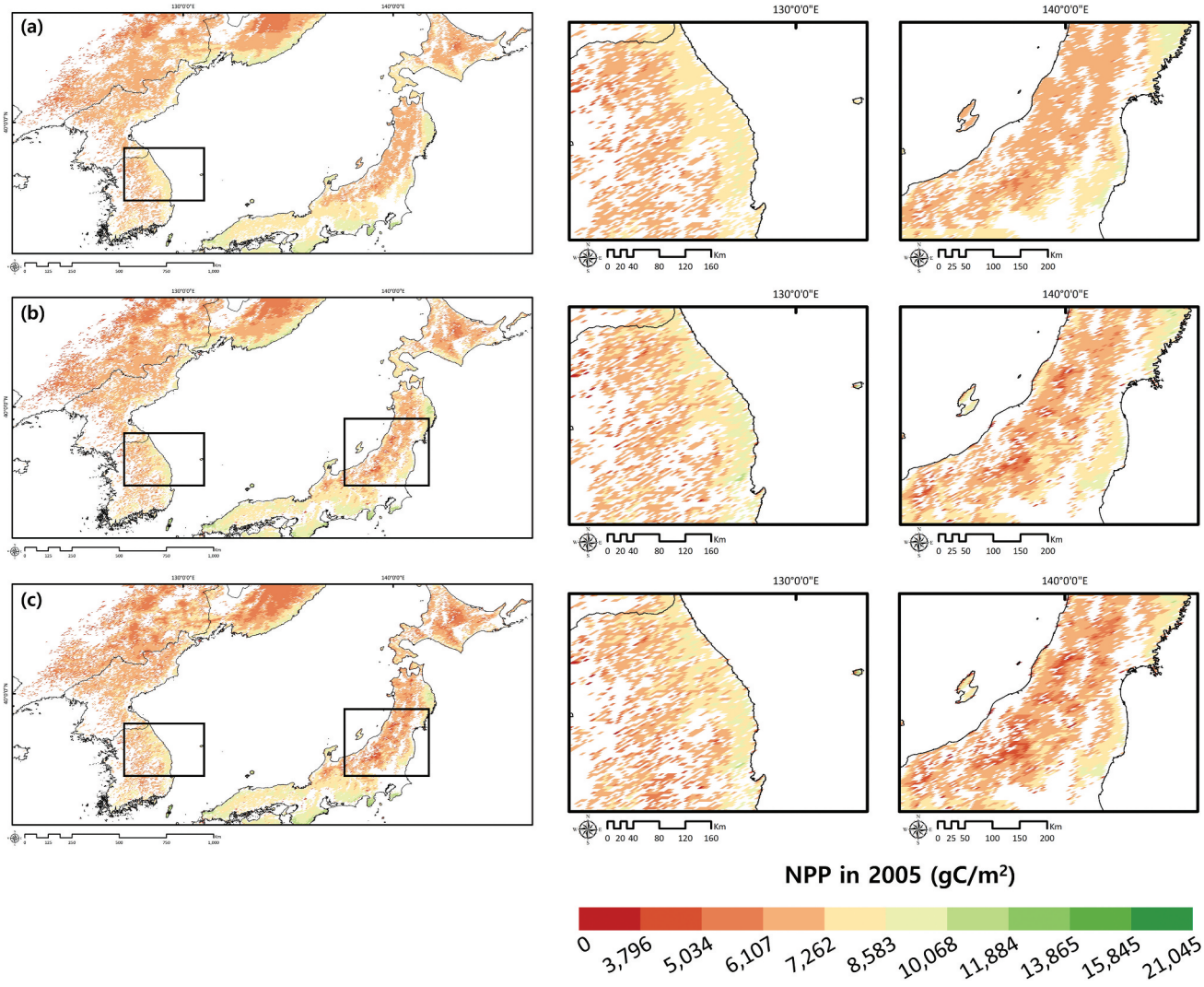

Figure 10. Comparison between the observed and predicted NPP of the eastern part (Korean Peninsula and Japan) on the local scale for 2005. (a) PNPM NPP, (b) DNPM NPP, and (c) observed NPP (EPSG: 54008).

outcomes, albeit in relatively modest proportions (depicted by shaded areas in Figure 13). Despite their slight variation, these changes are significant for formulating adaptation strategies, as it highlights areas with varying susceptibility to meteorological impacts and identifies locations with extreme values prone to disasters.

6. Discussion

6.1. Advantages of DPM concept

In this study, the versatility of the DPM approach was emphasized, suggesting its potential applicability to various environmental targets and its utility in making informed policy decisions across different scales. One of the key advantages of the DPM is its ability to be flexibly edited to estimate prediction targets using only meteorological and satellite data, making it suitable for application even in data-scarce regions. This flexibility allows for the continuous integration of spatiotemporal regional errors, enhancing prediction accuracy and applicability across diverse environmental contexts. Additionally, by presenting the diagnostic methodology within the DPM and guidelines for the PPM, it was demonstrated that predictive performance could be improved in various environmental sectors, such as drought and forest productivity. Unlike previous models, the DPM concept provides more accurate future predictions by continually diagnosing current environmental conditions and regional response characteristics, effectively calibrating residual from meteorological based prediction using the latest land observation.

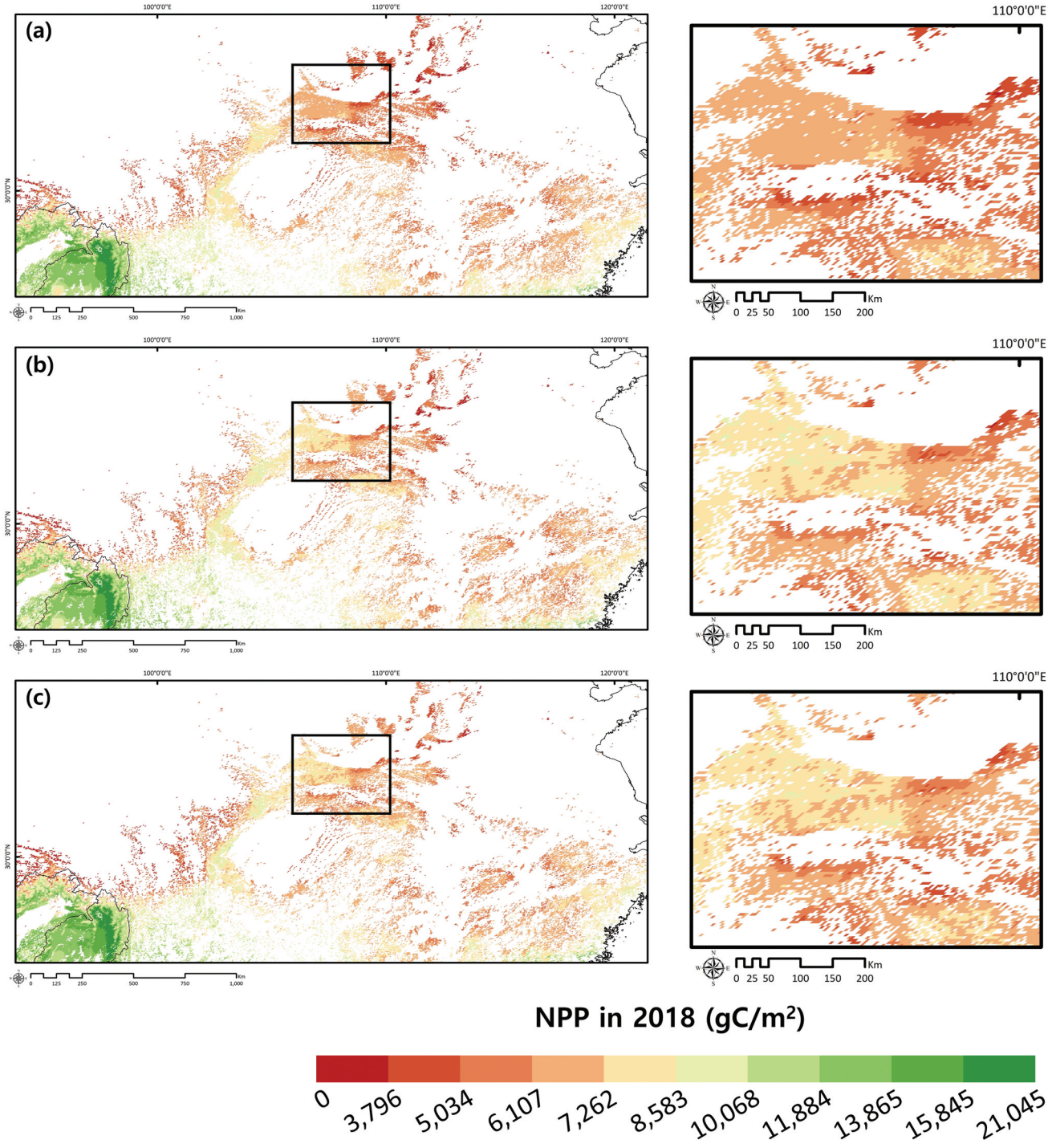

Figure 11. Comparison between the observed and predicted NPP of the eastern part (China) on the regional scale for 2018. (a) PNPM NPP, (b) DNPM NPP, and (c) observed NPP (EPSG: 54008).

The significance of DPM lies in its utilization of interpretable and predictable PPM as a foundation, augmented by accurate land observations to compensate for prediction errors. By incorporating spatiotemporal variation into meteorological based PPM prediction, DPM is able to detect local differences which can be the starting point of disaster even under the same weather conditions. This capability position not only to provide accurate predictive information but also to offer essential insight for establishing adaptation strategies in disaster-prone regions.

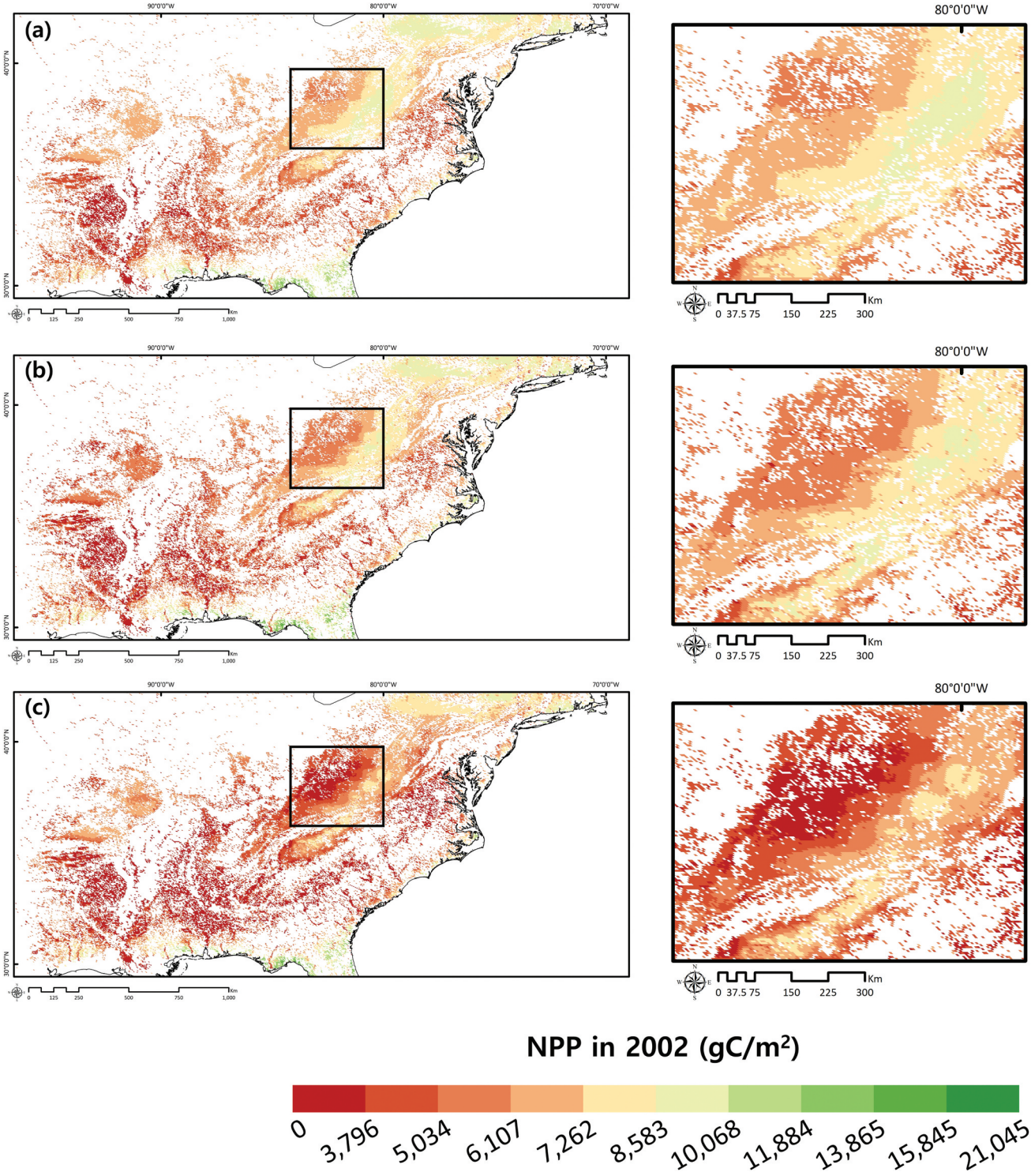

Figure 12. Comparison between the observed and predicted NPP of the western part (the US) on the regional scale for 2002. (a) PNPM NPP, and (b) DNPM NPP, and (c) observed NPP (EPSG: 54008).

6.2. DPM application for estimating forest NPP

Through the applicability assessment of DPM on forest productivity, this study quantitatively identified forest productivity by only using physical environmental factors based on meteorological and satellite data in MLR where forest resource data is insufficient. In terms of forest modeling, this study was designed as an interpretable regression model for climate reflection and a complex machine learning model for static environmental reflection to provide valuable information to decision makers for establishing future

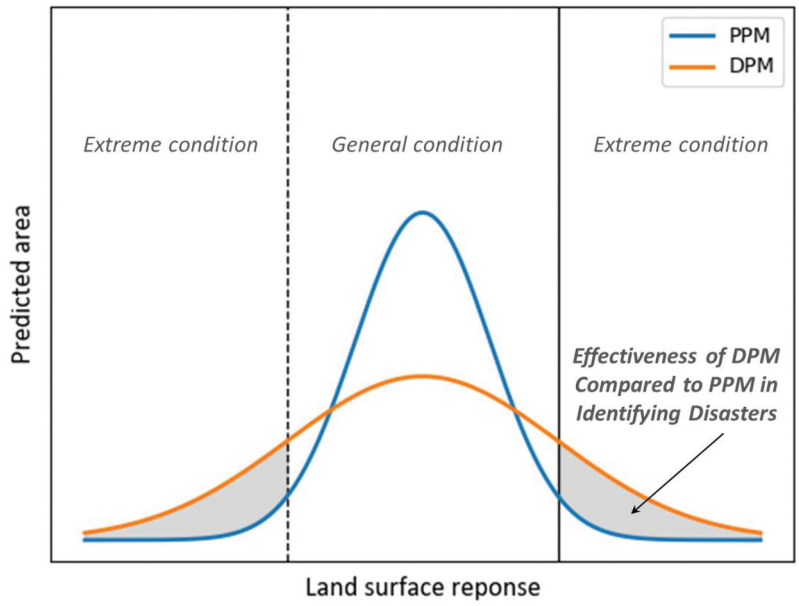

Figure 13. Prediction patterns for the predicted surface response using the PPM and DPM under different climate conditions. The DPM adds spatiotemporal variation (shaded area) to the PPM results by reflecting land surface response at the local level.

climate change and forest policy. The results of the PNPM were consistent with previous studies, indicating that topographic environments and climatic factors are pivotal for NPP simulation and can influence forest productivity over the medium and long term (Eum, Kang, and Lee 2005; Gillman et al. 2015; Kim et al. 2020; Shvidenko, Nilsson, and Buluy 2008; Song et al. 2019).

While PNPM relies on historical patterns within the diagnostic prediction concept, it allows for the interpretation of variables and long-term projections, thus playing a vital role in the quantification of forest resources. Since the PNPM already accounted for some regional spatial differences by static environmental features in this study, DNPM focused on the evolving land surface over time. In the diagnostic process of DNPM, the residuals were calculated following the premise of the DPM, and the PNPM results were temporally calibrated using linear regression, Consequently, DNPM exhibited accurate and effective model performance in both spatiotemporal and temporal validation. These results can be inferred that the DNPM captured the effects of changing socioenvironmental conditions, including factors, such as forest growth, land denudation progress, and forest management over time, which were not explicitly included in the PNPM. This was achieved on a regional spatial scale by using the time-series pattern of residuals at the annual prediction level. Therefore, the effectiveness of the DPM in compensating for spatiotemporal variations underscores the advantage of local-level prediction based on regional land surface responses, as opposed to the PPM, which relies solely on historical patterns. These results and approach can be linked to the existing forest models such as Global Forest Model (Kindermann et al. 2013) for assessing carbon sequestration and developing various scenarios of reforestation and forest management. This DPM approach is highly significant as it opens avenues for addressing complex environmental issues and informing evidence-based policy interventions through collaboration with forest stakeholders, such as national entities or MLR consortia, and interdisciplinary cooperation.

6.3. Limitations and outlook for future studies

The scalability and transferability of the DPM approach can be further refined through future research. DPM has shown potential in predicting various phenomena. Its effectiveness could be further demonstrated by applying it to a broader range of geographic regions and additional targets, including wildfires and landslides. The principle behind DPM is that PPM serves not only as the foundation of DPM but also as a dynamic element responsive to current conditions, with fitting errors exhibiting temporal

autocorrelation within a specific time range (E. Park et al. 2022). Consequently, high-quality datasets and various process-based models with diverse objectives can either replace PPM or enhance its interpretability. In the diagnostic process, sophisticated temporal algorithms could be used to estimate the impact of the current residuals on the subsequent predictions (E. Park et al. 2022). Additionally, identifying the range of temporal autocorrelation is crucial for establishing a flexible prediction period. Given that DPM assumes temporally autocorrelated fitting errors in PPM, applying a semi-variogram could help determine the range of autocorrelation with respect to temporal distance. As climate variability intensifies, the temporal autocorrelation range on the time axis is expected to change in the future. Thus, future studies could refine DPM by applying a semi-variogram after adjusting the time axis according to the rate of climate change.

Meanwhile, several measures need to be taken to further improve the DNPM, which has demonstrated applicability of DPM to forest productivity. While previous studies have utilized various meteorological data to predict forest production, this study used only SPEI 3 and annual temperature as climate variables. Future research should consider additional climatic variables, such as solar radiation, humidity, and CO_2 concentration. Given the potential complexity in the relationship among meteorological factors considered in the PNPM, selecting the most effective and relevant indicators is crucial. This study used multilinear regression in the PNPM to explore the relationship between the NPP and both meteorological and static environmental features across different climate zones. Additionally, although this study examined the comprehensive impact of environmental features on NPP, dynamic factors such as soil nitrogen and phosphorus can change due to forest management and climate change, affecting forest productivity. Future research should account for these dynamic variabilities. Due to the multifaceted and nonlinear nature of the natural environment, future PNPM improvements might involve using nonlinear regression or machine learning algorithms (Xue et al. 2022; Yang et al. 2017). These advancements should be directly compared and validated against existing NPP-related models in future research. Continuous collaboration with decision-makers is essential for refining the model by systematically comparing data, methodologies, and data fusion techniques. This collaborative effort aims to provide decision-makers with intuitive and efficient information for informed decision-making strategies. Finally, to assess model performance effectively across large MLR and ensure accurate validation without under- or over-estimation, future research should discuss appropriate validation grid sizes. Moreover, while the prediction error was reduced through time-series calibration in the DNPM, only simple linear regression was used. Additionally, satellite-based vegetation indices related to NPP, such as normalized difference vegetation index and leaf area index, should be considered in the diagnostic process. Therefore, future research should consider time-series models and various algorithms that can predict highly complex residual patterns.

7. Conclusions

Environmental spatial modeling for disaster risk reduction must balance model interpretability and accuracy in decision-making. This study generalized a diagnostic prediction concept, ensuring its applicability tailored to various environmental prediction modeling techniques. By introducing additional diagnosis algorithms and integrating topographic and soil features, the detailed development of the PPM and the diagnosis process within the DPM improved spatiotemporal accuracy, which is a crucial task for establishing adaptation strategies. The improved DPM concept, especially the diagnostic process that considered time-series residual patterns, was validated through the case study on forest NPP in MLR, where forest resource data is often insufficient. In the spatiotemporal validation, the DNPM (r: 0.975–0.992 and RMSE: 38.99–70.23gC/m^2/year) showed improved model performance compared to the PNPM (r: 0.948–0.967 and RMSE: 63.11–94.21 gC/m^2/year). Furthermore, in the temporal validation, the DNPM (r: 0.233–0.494 and RMSE: 46.01–70.75 gC/m^2/year) exhibited better accuracy than the PNPM (r: 0.192–0.406 and RMSE: 55.23–89.31 gC/m^2/year). These results demonstrate that the DNPM model consistently improves upon the performance of the PNPM model across all climate zones, highlighting the advantages of DNPM in emphasizing areas with varying susceptibility to meteorological impacts. From this perspective, DNPM can provide essential quantitative data on forest resources in MLR, addressing data limitations and the absence of detailed forest

models. This enhanced diagnostic prediction concept can aid in finalizing policy decisions and strategies at various scales, from local to global, by offering accurate future predictions. This study contributes to the developing globally applicable models and frameworks, facilitating accurate forest NPP estimation and supporting sustainable forest management. Additionally, it provides new insights for achieving nationally determined contributions and forming a consultative forest group in the MLR.

Acknowledgments

This work is part of the Ph.D. dissertation of the first author (E.P.) at Department of Environmental Science and Ecological Engineering, Korea University, under the supervision of W.K.L.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported under the framework of international cooperation program managed by the National Research Foundation of Korea [2021K2A9A1A02101519 and RS-2023-00270057].

ORCID

Eunbeen Park http://orcid.org/0000-0002-0442-7621
Hyun-Woo Jo http://orcid.org/0000-0001-6127-883X
Gregory Scott Biging http://orcid.org/0000-0001-5869-543X
Jong Ahn Chun http://orcid.org/0000-0001-8047-1811
Seong Woo Jeon http://orcid.org/0000-0001-5928-8510
Yowhan Son http://orcid.org/0000-0001-5621-9894
Florian Kraxner http://orcid.org/0000-0003-3832-6236
Woo-Kyun Lee http://orcid.org/0000-0002-2188-359X

Data availability statement

Historical climate data was downloaded from "CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations" (http://doi.org/10.15780/G2RP4Q. and https://www.chc.ucsb.edu/data/chirps). The DRYAD software was downloaded from Zenodo (https://zenodo.org/doi/10.5281/zenodo.10078913). The NPP data was downloaded and processed within Google Earth Engine (https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MOD17A3HGF). The soil and topographic data were downloaded from the Harmonized World Soil Database (https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/.) and STRM (https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm), respectively.

References

Beven, K. J., and M. J. Kirkby. 1979. "A Physically Based, Variable Contributing Area Model of Basin Hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant." *Hydrological Sciences Bulletin* 24 (1): 43–69. https://doi.org/10.1080/02626667909491834.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Chen, T., G. R. Werf, R. A. M. Jeu, G. Wang, and A. J. Dolman. 2013. "A Global Analysis of the Impact of Drought on Net Primary Productivity." *Hydrology and Earth System Sciences* 17 (10): 3885–3894. https://doi.org/10.5194/hess-17-3885-2013.

CISA. 2021. *Drought and Infrastructure: A Planning Guide*. Virginia, US: Cybersecurity and Infrastructure Security Agency. https://www.cisa.gov/sites/default/files/publications/Drought_and_Infrastructure_A_Planning_Guide_508c.pdf.

Eum, S., S. Kang, and D. Lee. 2005. "A Simulation Study to Investigate Climatic Controls on Net Primary Production (NPP) of a Rugged Forested Landscape in the Mid-Western Korean Peninsula." *Korean Journal of Agricultural and Forest Meteorology* 7 (1): 66–77.

FAO. 2020. *Global Forest Resources Assessment 2020: Main Report*. Rome. https://doi.org/10.4060/ca9825en.

FAO. 2021. *The Impact of Disasters and Crises on Agriculture and Food Security: 2021*. Rome. https://doi.org/10.4060/cb3673en.

Fick, S. E., and R. J. Hijmans. 2017. "WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37 (12): 4302–4315. https://doi.org/10.1002/joc.5086.

Field, C. B., J. T. Randerson, and C. M. Malmström. 1995. "Global Net Primary Production: Combining Ecology and Remote Sensing. Remote Sens." *Remote Sensing of Environment* 51 (1): 74–88. https://doi.org/10.1016/0034-4257(94)00066-V.

Fischer, G., F. Nachtergaele, S. Prieler, H. T. van Velthuizen, L. Verelst, and D. Wiberg. 2008. *Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008)*. Rome, Italy: IIASA, Laxenburg, Austria and FAO.

Gillman, L. N., S. D. Wright, J. Cusens, P. D. Mcbride, Y. Malhi, and R. J. Whittaker. 2015. "Latitude, Productivity and Species Richness." *Global Ecology & Biogeography* 24 (1): 107–117. https://doi.org/10.1111/geb.12245.

Hassan, Q. K. 2018. "Introduction to Environmental Modelling." In *Environmental Modelling*, edited by Q. K. Hassan. Calgary, AB, Canada. https://doi.org/10.11575/PRISM/5245.

Hengl, T., G. B. M. Heuvelink, and D. G. Rossiter. 2007. "About Regression-Kriging: From Equations to Case Studies." *Computers & Geosciences* 33 (10): 1301–1315. https://doi.org/10.1016/j.cageo.2007.05.001.

Jarvis, A., H. I. Reuter, A. Nelson, and E. Guevara. 2008. "Hole-Filled SRTM for the Globe Version 4." *Available from the CGIAR-CSI SRTM 90m Database* 15 (25–54): 5.

Jo, H. W., and W. K. Lee. 2024. *Deep Learning & Remote Sensing Analysis for Agroforesty and Drought (DRYAD)*. Zenodo. https://doi.org/10.5281/zenodo.10687417.

Kim, M., B.-Y. Ham, F. Kraxner, A. Shvidenko, D. Schepaschenko, A. Krasovskii, T. Park, and W.-K. Lee. 2020. "Species- and Elevation-Dependent Productivity Changes in East Asian Temperate Forests." *Environmental Research Letters* 15 (3): 034012. https://doi.org/10.1088/1748-9326/ab71a2.

Kindermann, G. E., I. McCallum, S. Fritz, and M. Obersteiner. 2008. "A Global Forest Growing Stock, Biomass and Carbon Map Based on FAO Statistics." *Silva Fennica* 42 (3): 387–396. https://doi.org/10.14214/sf.244.

Kindermann, G. E., S. Schörghuber, T. Linkosalo, A. Sanchez, W. Rammer, R. Seidl, and M. J. Lexer. 2013. "Potential Stocks and Increments of Woody Biomass in the European Union Under Different Management and Climate Scenarios." *Carbon Balance and Management* 8 (1): 1–20. https://doi.org/10.1186/1750-0680-8-2.

Kopecky, M., and S. Cízková. 2010. "Using Topographic Wetness Index in Vegetation Ecology: Does the Algorithm Matter?" *Applied Vegetation Science* 13 (4): 450–459. https://doi.org/10.1111/j.1654-109X.2010.01083.x.

Li, H., D. Yang, C. Gao, and W. Liu. 2022. "Effects of Built Area on Net Primary Productivity Provided by Different Landscape Characters and the Promotion Potential." *Frontiers in Environmental Science* 10 (September): 1–14. https://doi.org/10.3389/fenvs.2022.988362.

Mickler, R. A., T. S. Earnhardt, and J. A. Moore. 2002. "Regional Estimation of Current and Future Forest Biomass." *Environmental Pollution* 116:S7–S16. https://doi.org/10.1016/S0269-7491(01)00241-X.

Orusa, T., and E. Borgogno Mondino. 2021. "Exploring Short-Term Climate Change Effects on Rangelands and Broad-Leaved Forests by Free Satellite Data in Aosta Valley (Northwest Italy)." *Climate* 9 (3): 47. https://doi.org/10.3390/cli9030047.

Orusa, T., D. Cammareri, and E. Borgogno Mondino. 2022. "A Possible Land Cover EAGLE Approach to Overcome Remote Sensing Limitations in the Alps Based on Sentinel-1 and Sentinel-2: The Case of Aosta Valley (NW Italy)." *Remote Sensing* 15 (1): 178. https://doi.org/10.3390/rs15010178.

Pan, Y., A. D. Mcguire, D. W. Kicklighter, and J. M. Meillo. 1996. "The Importance of Climate and Soils for Estimates of Net Primary Production: A Sensitivity Analysis with the Terrestrial Ecosystem Model." *Global Change Biology* 2 (1): 5–23. https://doi.org/10.1111/j.1365-2486.1996.tb00045.x.

Park, E., H.-W. Jo, W.-K. Lee, S. Lee, C. Song, H. Lee, S. Park, W. Kim, and T.-H. Kim. 2022. "Development of Earth Observational Diagnostic Drought Prediction Model for Regional Error Calibration: A Case Study on Agricultural Drought in Kyrgyzstan." *GIScience & Remote Sensing* 59 (1): 36–53. https://doi.org/10.1080/15481603.2021.2012370.

Park, J. H., J. Gan, and C. Park. 2021. "Discrepancies Between Global Forest Net Primary Productivity Estimates Derived from Modis and Forest Inventory Data and Underlying Factors." *Remote Sensing* 13 (8): 1441. https://doi.org/10.3390/rs13081441.

Piao, D., M. Kim, G.-M. Choi, J. Moon, H. Yu, W.-K. Lee, S. W. Wang, et al. 2018. "Development of an Integrated DBH Estimation Model Based on Stand and Climatic Conditions." *Forests* 9 (3): 155. https://doi.org/10.3390/F9030155.

IPCC, Pirani, A., S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen. 2021. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group 1 to Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. edited by Masson-Delmotte, V., P. Zhai, Maycock, T. W, O. Yelekçi, R. Yu, and B. Zhou. 2391. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. https://doi.org/10.1017/9781009157896.

Running, S. W., and E. R. Hunt Ehleringer, J.R., eds. 1993. *Generalization of a Forest Ecosystem Process Model for Other Biomes, BIOME-BGC, and an Application for Global-Scale Models in Scaling Physiological Processes: Leaf to Globe*. Field, C.B. San Diego: Academic Press, Inc. https://doi.org/10.1016/B978-0-12-233440-5.50014-2.

Shahin, M. A., M. A. Ali, and A. B. M. S. Ali. 2014. "Vector Autoregression (VAR) Modeling and Forecasting of Temperature, Humidity, and Cloud Coverage." In *Computational Intelligence Techniques in Earth and Environmental Sciences*, 29–51. Springer Netherlands. https://doi.org/10.1007/978-94-017-8642-3_2.

Shamsnia, S. A., N. Shahidi, A. Liaghat, A. Sarraf, and S. F. Vahdat. 2011. "Modeling of Weather Parameters Using Stochastic Methods (ARIMA Model)(case Study: Abadeh Region, Iran)." *International Conference on Environment and Industrial Innovation*, 12: 282–285. IPCBEE. http://www.ipcbee.com/vol12/55-C30028.pdf.

Shvidenko, A. S., D. G. Nilsson, and Y. I. Buluy. 2008. "Tables and Models of Growth and Productivity of Forests of Major Forest Forming Species of Northern Eurasia (Standard and Reference Materials)." *Federal Agency of Forest Management*. http://webarchive.iiasa.ac.at/Research/FOR/forest_cdrom/Articles/THR.pdf.

Song, C., S. A. Pietsch, M. Kim, S. Cha, E. Park, A. Shvidenko, D. Schepaschenko, F. Kraxner, and W.-K. Lee. 2019. "Assessing Forest Ecosystems Across the Vertical Edge of the Mid-Latitude Ecotone Using the BioGeochemistry Management Model (BGC-MAN)." *Forests* 10 (6): 523. https://doi.org/10.3390/f10060523.

Sörensen, R., U. Zinko, and J. Seibert. 2006. "On the Calculation of the Topographic Wetness Index: Evaluation of Different Methods Based on Field Observations." *Hydrology and Earth System Sciences* 10 (1): 101–112. https://doi.org/10.5194/hess-10-101-2006.

Tang, G., B. Beckage, B. Smith, and P. A. Miller. 2010. "Estimating Potential Forest NPP, Biomass and Their Climatic Sensitivity in New England Using a Dynamic

Ecosystem Model." *Ecosphere* 1 (6): 1–20. https://doi.org/10.1890/ES10-00087.1.

Thornthwaite, C. W. 1948. "An Approach Toward a Rational Classification of Climate." *Geographical Review* 38 (1): 55–94. https://doi.org/10.2307/210739.

Trabucco, A., and R. J. Zomer. 2019. "Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database V2." *CGIAR Consortium for Spatial Information (CGIAR-CSI)* 10. https://doi.org/10.6084/m9.figshare.7504448.v3.

UN/ISDR. 2008. "Climate Change and Disaster Risk Reduction." *UN/ISDR Briefing Note* 1. https://eird.org/publicaciones/Climate-Change-DRR.pdf.

United Nations Department of Economic and Social Affairs, United Nations Forum on Forests Secretariat. 2021. "The Global Forest Goals Report 2021." https://www.un.org/esa/forests/wp-content/uploads/2021/04/Global-Forest-Goals-Report-2021.pdf.

Vicente-Serrano, S. M., S. Beguería, and J. I. López-Moreno. 2010. "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index." *Journal of Climate* 23 (7): 1696–1718. https://doi.org/10.1175/2009JCLI2909.1.

Xue, P., H. Liu, M. Zhang, H. Gong, and L. Cao. 2022. "Nonlinear Characteristics of Npp Based on Ensemble Empirical Mode Decomposition from 1982 to 2015—A Case Study of Six Coastal Provinces in Southeast China." *Remote Sensing* 14 (1): 15. https://doi.org/10.3390/rs14010015.

Yang, J., X. C. Zhang, Z. H. Luo, and X. J. Yu. 2017. "Nonlinear Variations of Net Primary Productivity and Its Relationship with Climate and Vegetation Phenology, China." *Forests* 8 (10): 361. https://doi.org/10.3390/f8100361.

Yu, B., F. Chen, and H. Chen. 2019. "NPP Estimation Using Random Forest and Impact Feature Variable Importance Analysis." *Journal of Spatial Science* 64 (1): 173–192. https://doi.org/10.1080/14498596.2017.1367331.

Zhao, M., and S. W. Running. 2010. "Drought-Induced Reduction in Global Terrestrial Net Primary Production from 2000 Through 2009." *Science* 329 (5994): 940–943. https://doi.org/10.1126/science.1192666.