

Article

A Physics-Guided Deep Learning Model for 10-h Dead Fuel Moisture Content Estimation

Chunquan Fan  and Binbin He *

School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China; 201911070403@std.uestc.edu.cn

* Correspondence: binbinhe@uestc.edu.cn

Abstract: Dead fuel moisture content (DFMC) is a key driver for fire occurrence and is often an important input to many fire simulation models. There are two main approaches to estimating DFMC: empirical and process-based models. The former mainly relies on empirical methods to build relationships between the input drivers (weather, fuel and site characteristics) and observed DFMC. The latter attempts to simulate the processes that occur in the fuel with energy and water balance conservation equations. However, empirical models lack explanations for physical processes, and process-based models may provide an incomplete representation of DFMC. To combine the benefits of empirical and process-based models, here we introduced the Long Short-Term Memory (LSTM) network and its combination with an effective physics process-based model fuel stick moisture model (FSMM) to estimate DFMC. The LSTM network showed its powerful ability in describing the temporal dynamic changes of DFMC with high R^2 (0.91), low RMSE (3.24%) and MAE (1.97%). When combined with a FSMM model, the physics-guided model FSMM-LSTM showed better performance ($R^2 = 0.96$, RMSE = 2.21% and MAE = 1.41%) compared with the other models. Therefore, the combination of the physics process and deep learning estimated 10-h DFMC more accurately, allowing the improvement of wildfire risk assessments and fire simulating.

Keywords: dead fuel moisture content (DFMC); deep learning; FSMM-LSTM; LSTM; wildfires



Citation: Fan, C.; He, B. A Physics-Guided Deep Learning Model for 10-h Dead Fuel Moisture Content Estimation. *Forests* **2021**, *12*, 933. <https://doi.org/10.3390/f12070933>

Academic Editor: Alfonso Fernández-Manso and Carmen Quintano

Received: 24 May 2021
Accepted: 14 July 2021
Published: 16 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Wildfires are crucial to water and carbon cycles on earth [1–4]. On one hand, they emit CO₂ into the atmosphere, which may inflict damage to climate [5], air quality [6] and health [7]. On the other hand, human lives, property and infrastructure are vulnerable to large, high-intensity wildfires [8]. Thus, predicting forest wildfire risk counts for a great deal [9].

Fuels consumed in wildfires are composed of dead and live plant material. Many efforts have focused on live fuels [10–12], whose moisture content changes slowly throughout the day [13]. While Dead Fuel Moisture Content (DFMC) responds rapidly to atmospheric conditions, which is of great importance in affecting fire, such as ignition probability, spread rate and intensity [1,14,15]. Therefore, DFMC estimation is required for quantifying fire danger, and almost all fire models include DFMC as an input variable [16].

DFMC is a function of fuel size and atmospheric conditions [17]. It increases or decreases with the change of climate variables through water vapor sorption or desorption until it eventually reaches a stable moisture content, i.e., Equilibrium Moisture Content (EMC) [18]. The time it takes to lose or gain about 63% of the difference between its initial values and EMC is defined as time-lag [19], which is related to the diameter of fuel. With the time-lag theory, dead fuels can be divided into four categories (1-h, 10-h, 100-h and 1000-h fuels, where the number refers to the time-lag of fuel) [20]. For instance, 10-h fuel usually is related to the fuel with a diameter ranging from 0.64 to 2.54 cm [21]. The 10-h DFMC is a promising predictor since it can be automatically measured in real time at study sites. For dead fuel with a certain size, its water content is commonly modeled with meteorological variables such as air temperature, relative humidity, wind speed, rainfall, solar radiation, soil moisture

content [22]. Plenty of models have been proposed to explain the relationship between those variables and DFMC, and they can be summarized as two approaches: empirical and process-based models [23]. The empirical model relies on empirical relationships between input variables and DFMC from field observations [24–31]. Since empirical methods are fully data-driven, they lack explanations for physical processes such as heat, water vapor fluxes and vapor exchange [32]. Process-based models estimate DFMC by attempting to simulate the processes that affect the water of fuel [33]. There are three types of process-based models: bulk litter layer models, models based on Byram’s diffusion equation and complete process-based models [16]. Bulk litter layer models such as Tamai only contain interception and evaporation processes which lead to underestimating in wetter conditions [34]. The models based on Byram’s diffusion equation heavily rely on EMC equation such as the Nelson formula for EMC [18,35,36]. Complete process-based models try their best to represent fluxes of energy and water in fuel with energy and water balance conservation equations [37,38]. Due to the complexities of the physics processes related to dead fuel, it is challenging to include all these processes into a process-based model [16].

Above all, empirical models lack explanations for physical processes and process-based models are particularly complex [16]. Recently, the combination of process-based model and empirical model has been widely accepted in remote sensing research, such as foliage fuel load (FFL) monitoring using radiative transfer model and machine learning method [39], estimating crop primary productivity using machine learning methods trained with radiative transfer simulations [40], physics informed neural networks for simulating radiative transfer [41]. The combination of the process-based model and the empirical model represents another potential approach for estimating DFMC. This approach requires a process-based model and an empirical model [42]. The empirical models used in previous studies were traditional machine learning and deep learning algorithms [43,44]. Compared to traditional machine learning algorithms, deep learning methods can more effectively mine the information in the data, especially the Long Short-Term Memory (LSTM) network [45], which performs well in describing the temporal dynamic changes of the time-sequential data [46–48]. However, to our best knowledge, there are neither deep learning methods nor their combined application with physical process models in estimating DFMC of any size dead fuel. Since approaches based on the combination of physics model and empirical model have shown competitive performance in other areas, their applicability to 10-h DFMC estimation deserves to be tested.

This study firstly used the LSTM network to estimate 10-h DFMC. Secondly, we introduced an effective process-based model to guide the LSTM network for 10-h DFMC estimating. To test the performance of the LSTM network and the hybrid model, we implemented four effective models, including multiple linear regression (MLR), random forest, artificial neural network (ANN) and the fuel stick moisture model (FSMM). All six models were tested on a continuous DFMC dataset.

2. Materials

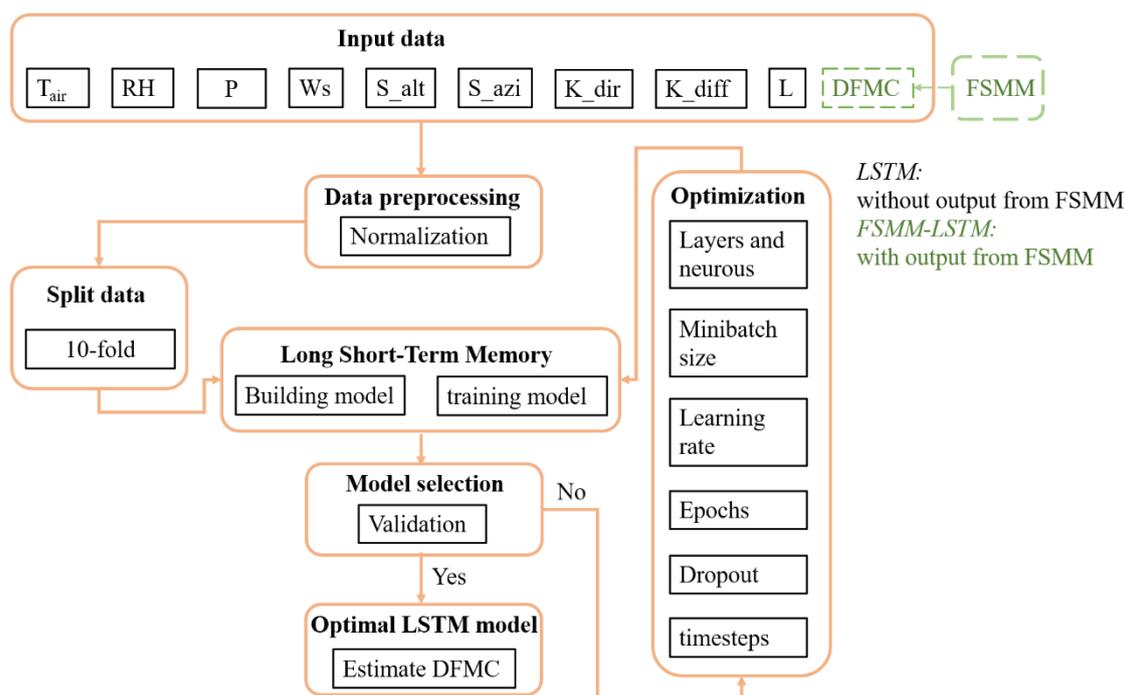
The dataset used in this study is available in D.W. van der Kamp et al. (2017) [49] and shared in Github (<https://github.com/dvdkamp/fsmm>, accessed on 20 May 2021). The observation values were collected at a forested field site located near Kamloops, British Columbia, Canada, between May and September 2014. Both the DFMC and other required variables (i.e., air temperature, relative humidity) were measured at the BC1 site in 3156 h. The 10-h DFMC observations were measured by Campbell Scientific CS506 Fuel Moisture Sensors at a standard height of 30.5 cm. This sensor consists of a time-domain reflectometer probe embedded within a standard moisture stick with a length of 50.8 cm and a radius of 0.65 cm. A Rotronic HC-S3 humidity and air temperature sensor (also at a height of 30.5 cm), a Rainwise tipping bucket rain-gauge, a Met One anemometer and a Kipp and Zonen CM3 pyranometer were used to record the air relative humidity, air temperature, rainfall, wind speed and solar radiation, respectively. Table 1 shows all the input variables.

Table 1. Overview of input drives for DFMC. These variables are recorded each hour.

Number	Input Drives	Abbreviation
1	Air temperature ($^{\circ}\text{C}$)	T_{air}
2	Relative humidity (0–100%)	RH
3	Rainfall (cm)	P
4	Wind speed (m/s)	Ws
5	Sun altitude (rad)	S_{alt}
6	Sun azimuth (rad)	S_{azi}
7	Downwelling direct shortwave radiation (W/m^2)	K_{dir}
8	Downwelling diffuse shortwave radiation (W/m^2)	K_{diff}
9	Downwelling longwave radiation (W/m^2)	L

3. Methods

We combined FSMM (process-based model) and LSTM (empirical model) to estimate 10-h DFMC (hereafter ‘FSMM-LSTM’). FSMM was selected as the fundamental process, because of its good performance across all size fuels [49]. Usually, FSMM estimates the DFMC at a certain time t , based on DFMC value at the last time $t-1$ [50]. On the other hand, the temporal dynamic information of DFMC can be mined by hourly iterations [51], and this dynamic information can be well processed by the LSTM algorithm [52]. Thus, LSTM was selected as the empirical model to estimate DFMC [52]. Figure 1 shows the flowchart of the DFMC estimation algorithms, including LSTM and FSMM-LSTM.

**Figure 1.** The flowchart of the LSTM (black) and FSMM-LSTM (black and green). Details of the input data can be found in Table 1.

3.1. Physics-Guide Model

3.1.1. FSMM Model

FSMM is a process-based model proposed by D.W. van der Kamp et al. (2017) [49]. With hourly measurements of relative humidity, air temperature, precipitation, wind speed, shortwave radiation and longwave radiation, FSMM can estimate DFMC of fuels at multiple sizes, particularly including the 10-h fuel. In this model, dead fuel was divided into two zones: an outer part that responds to atmospheric forcing, and a central core that only trades water and energy with the outer layer. The schematic of FSMM is shown in

Figure 2 and the diagram of the directed graph of FSMM is shown in Figure 3. Based on an initial value m_{t-1} , FSMM can estimate the DFMC at t (m_t) with hourly input variables and subsequently based on m_t , the value at $t + 1$ (m_{t+1}) can be estimated as well. The water content of the outer layer and the core was named m_o and m_c , respectively. In addition, the temperatures of them were named T_o and T_c . Then the average temperature, T_s and the moisture content, m_s of the fuel can be calculated as:

$$m_s = f * m_o + (1 - f) * m_c \tag{1}$$

$$T_s = f * T_o + (1 - f) * T_c \tag{2}$$

where f is the fraction of the fuel volume taken up by the outer layer and can be estimated via calibration.

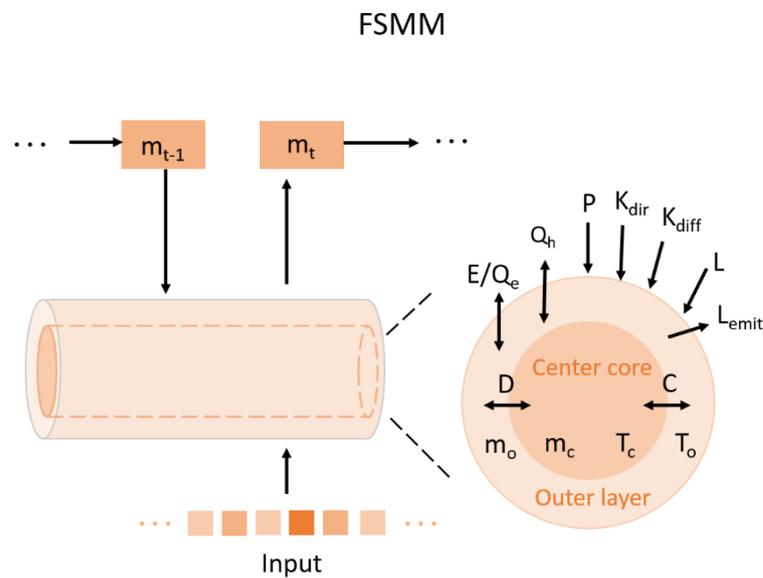


Figure 2. FSMM schematic. The m_{t-1} is refer to DFMC at time $t-1$ while m_t is the DFMC at time t . E is moisture flux between the outer layer and the atmosphere (kg/s), K is shortwave radiation (W/m^2), L is longwave radiation (W/m^2), P is precipitation (kg/s), Q is turbulent heat flux (W/m^2), Q_h is sensible heat flux (W/m^2), D is diffusion into the core (kg/s) and C is conduction into the core (W).

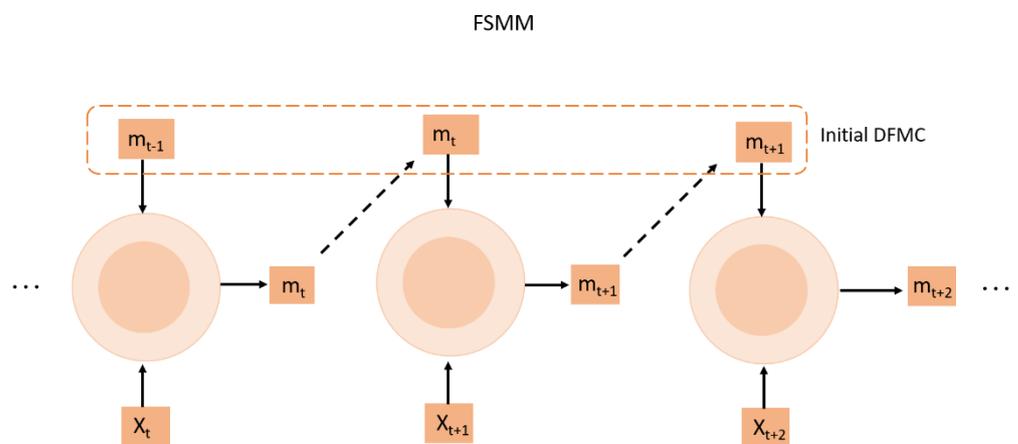


Figure 3. The operation process of FSMM model. Where m is DFMC, m_{t-1} is m at time $t-1$. X refers to the input drivers, which were shown in Table 1.

The water exchange between air and the outer layer is affected by three processes: absorbed precipitation (P), evaporation/desorption (E) and diffusion (D) into the core:

$$\frac{dm_o}{dt} = P - (a_s - 2 * \pi * r^2) * E - D \quad (3)$$

where a_s is the surface area of the entire fuel and $a_s - 2 * \pi * r^2$ means the lateral fuel surface. All these three terms are in units of kg/s. Evaporation is closely related to the latent heat flux, which connects the moisture content and the energy budgets.

The moisture content budget for the core only contains diffusion:

$$\frac{dm_c}{dt} = D \quad (4)$$

The outer layer temperature T_o , it involves multiple energy exchange processes:

$$\frac{dT_o}{dt} = \frac{1}{C_S * \rho_s * V_o} * \left(K_{dir} + K_{diff} + L - a_s * L_{emit} - a_s * Q_h - (a_s - 2 * \pi * r^2) * Q_e - C \right) \quad (5)$$

where K_{dir} and K_{diff} is the absorbed direct and diffuse shortwave radiation, respectively (W). L is the absorbed longwave radiation (W). L_{emit} is the emitted longwave radiation (W/m^2), Q_h is the sensible heat flux (W/m^2), Q_e is the latent heat flux (W/m^2) and C is the conduction into the fuel's core (W). The coefficient C_S is the fuel-specific heat ($J/(K*kg)$), ρ_s is the stick density ($400 \text{ kg}/m^3$) [38] and V_o is the volume of the outer layer.

The only energy exchange between the outer layer and core is conduction:

$$\frac{dT_o}{dt} = \frac{C}{C_S * \rho_s * V_o} \quad (6)$$

With all energy and water budget processes determined, the FSMM model can estimate the dynamic change of DFMC. More details of the FSMM model can be referenced from D.W. van der Kamp et al. (2017) [49].

3.1.2. LSTM Network

The LSTM network is developed based on the standard RNN [53], which has good performance in describing the temporal dynamic changes of the time-sequential data such as DFMC. The LSTM model was proposed to solve the problem of gradient vanishing as well as the explosion of long-term dependences in traditional the RNN networks [45]. The schematic of LSTM is shown in Figure 4. The LSTM network computed a mapping from an input sequence to an output sequence (Figure 4A).

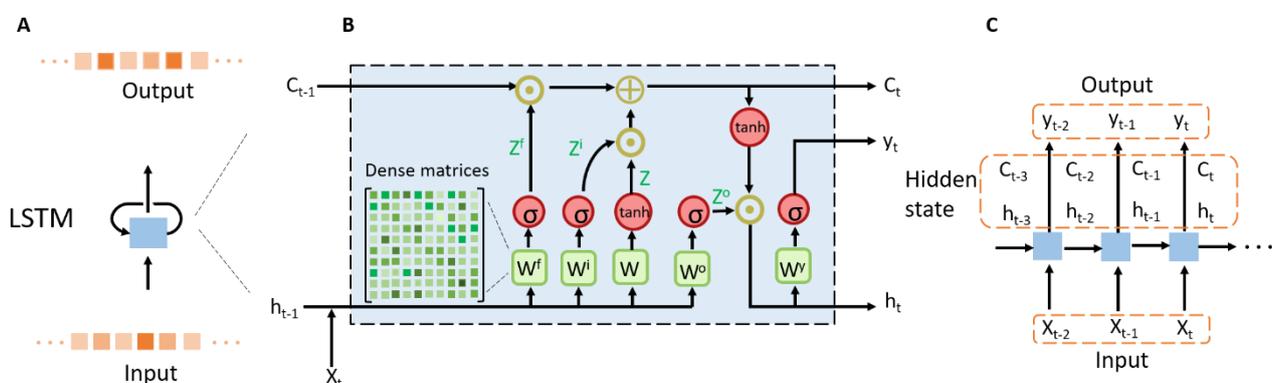


Figure 4. LSTM schematic. (A) Diagram of an LSTM cell. (B) Internal components of the LSTM cell, consisting of trainable dense matrices W^f, W^i, W, W^o and W^y . Activation functions for the hidden state and output are represented by σ and \tanh . (C) Diagram of the directed graph of the LSTM cell. x_t is the input driver introduced in Table 1.

There are three gate controllers used to determine what information should be forgotten in the LSTM unit: input (Z^i), forget (Z^f) and output gates (Z^o). Switching the gates to prevent the gradient from vanishing, the LSTM network kept the temporal memory. The basic LSTM unit requires the current input vector X_t , its previous cell state C_{t-1} and the previous hidden state h_{t-1} . These three gates are obtained as:

$$Z^i = \sigma \left(W_x^i * X_t + W_h^i * h_{t-1} + b^i \right) \quad (7)$$

$$Z^f = \sigma \left(W_x^f * X_t + W_h^f * h_{t-1} + b^f \right) \quad (8)$$

$$Z^o = \sigma \left(W_x^o * X_t + W_h^o * h_{t-1} + b^o \right) \quad (9)$$

where σ is the nonlinear activation function, usually it is set to be the sigmoid function.

Except for these three gates, another intermediate state Z is calculated as:

$$Z = \tanh(W_x * X_t + W_h * h_{t-1} + b) \quad (10)$$

where \tanh means the nonlinear tanh activation function.

Then, the memory cell (C_t) and hidden state (h_t) of this LSTM are updated as:

$$C_t = Z^f \odot C_{t-1} + Z^i \odot Z \quad (11)$$

$$h_t = Z^o \odot \tanh(C_t) \quad (12)$$

where \odot represents the pointwise multiplication operation for two vectors

Therefore, y_t can be obtained as:

$$y_t = \sigma \left(W_h^y * h_t + b^y \right) \quad (13)$$

We applied the LSTM network by following the Keras [54] package with Tensorflow backend [55]. To avoid the over-fitting issues and improve the convergence speed in the LSTM training process, we adjusted the epochs, batch size, time step, learning rate, neurons, dropout and patience, as well as early stopping (Table 2). The optimal LSTM network was determined based on the comprehensive consideration of the prediction accuracy and stability of the model. For example, a time step of 20 could successfully capture the dynamic changes of DFMC in time series. To avoid over-fitting, an early stopping procedure was employed using 10% of all data for validation, where the patience value was set to 65. The max epoch was set as 500.

Table 2. Training hyperparameters for the LSTM network and FSMM-LSTM network.

Model	Parameter	Value	Parameter	Value
LSTM	Hidden units	45	Dropout	0.7
	Batch size	4	Timestep	20
	Learning rate	$5 \times 10(-3)$	Patience	65
	Optimizer	Adam	Loss	Mean square error
FSMM-LSTM	Hidden units	65	Dropout	0.1
	Batch size	16	Timestep	20
	Learning rate	$1 \times 10(-3)$	Patience	50
	Optimizer	Adam	Loss	Mean square error

3.1.3. FSMM-LSTM Network

Given our predictive learning problem where we are given a variety of input drivers, X , that are physically related to our target variable of interest, DFMC. An efficient approach is to train a data science model, e.g., a recurrent neural network, $f_{LSTM} : X \rightarrow DFMC$, over a set of training variables. Then DFMC can be estimated through this trained model.

Alternatively, a physics-based numerical model, e.g., FSMM, $f_{PHY} : X \rightarrow DFMC$, can also be used to estimate the value of our target variable, with its physical relationships with the input variables. However, process-based model may provide an incomplete description of the target variable because of simplified or missing physics in f_{PHY} [42], which may lead to inaccurate DFMC estimates and erroneous judgment of fire risk. Therefore, a hybrid model was proposed to combine f_{PHY} and f_{LSTM} as to overcome their respective shortcomings and take advantage of the information in both physics and data. The schematic of the hybrid model (FSMM-LSTM) is shown in Figure 5. This model was composed of two parts: FSMM and LSTM. All input variables need to be sent into the FSMM model to estimate a sequence of DFMC. Then the estimated DFMC becomes the input of the LSTM model, along with input variables such as air temperature, relative humidity. Finally, the final variable DFMC was estimated.

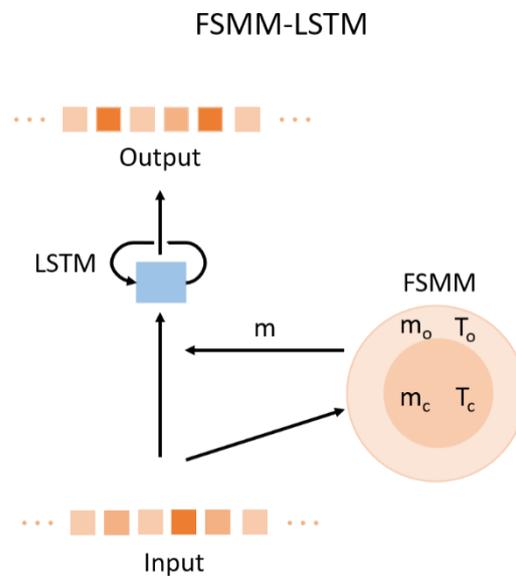


Figure 5. FSMM-LSTM schematic. Where m is DFMC, input is the input drivers.

Since the input variables of the FSMM-LSTM model contain not only the input drivers used in the LSTM network but also the output of the FSMM model, we cannot use the optimal LSTM network parameters to optimize the FSMM-LSTM model. Therefore, we adjusted the hyperparameters by multiple debugging. The training hyperparameters for the FSMM-LSTM algorithm is shown in Table 2.

3.2. Model Comparison

In previous studies, a set of models including process-based methods and empirical methods have been developed for DFMC estimating. Of them, an effective process-based model FSMM has already been introduced in Section 3.1.1. Furthermore, a simple but efficient method MLR was used. In addition, machine learning methods such as random forest were selected due to their excellent performance in DFMC estimating [25]. Another machine learning method ANN has shown better performance than random forest in regression and classification tasks [56,57]. However, to the best of our knowledge, no previous study has applied ANN to estimate DFMC of any size dead fuel, including 10-h DFMC. Therefore, ANN was select here to make sure its performance in DFMC estimating.

3.2.1. MLR

Given observation data of both the predictors for a specific sample, simple linear regression was used to estimate the relationship between a response predictor and a single explanatory predictor [58]. Adding additional explanatory predictors to a simple linear regression model developed a multiple linear regression model.

3.2.2. Random Forest

Random forest, based on the classification and regression tree (CART) [59], was one of the widely used machine-learning methods. It can build an ensemble of a lot of regression trees using generally two-thirds of the whole data as bootstrapped training data [25]. Each tree can train a predefined size part of the whole available variables. The final regression of the random forest depends on the votes of the multiple trees [60]. With the correlation among decision trees in random forest decrease, the ensemble of the trees is more reliable [61]. In this study, the number of trees was set as 100 for a balance between computational cost and reliable performance.

3.2.3. ANN

ANN is a computing system composed of a collection of connected artificial neurons [62]. Each neuron means a specific output function, named the activation function. Every connection between two neurons is a weight for the signal passing through the connection, which is equivalent to the memory of ANN [63]. It can handle nonlinear and complex problems better [64]. Due to its powerful ability to describe the relationship between inputs and outputs from the training data, it has been widely used in geophysical parameter estimations [56]. For the ANN algorithm, there was two hidden layers with 40 units and 50 units, respectively.

3.3. Model Evaluation

We have a continuous time series of DFMC data from May to September 2014. Empirical methods are prone to over-fit to the training data, rendering the calculation of the model's performance on training data unusable [9]. Therefore, 10-fold cross-validation was adopted to validate the models except for the process-based model, FSMM. In other words, the whole data were divided into 10 pieces equally. One piece was for verification in each run, and the remaining nine pieces were for training. This process should repeat 10 times until each piece was selected for verification. To fairly compare the performance of the models, the input variables of all other models were the same as the LSTM network except the FSMM-LSTM model which has an extra input from the output of the FSMM model. The determination coefficient (R^2), the root mean square error (RMSE) and the mean absolute error (MAE) were used to evaluate all model performances.

4. Results

Figure 6 provides a summary of the performance of different methods for modeling DFMC on the BC1 site. The DFMC estimated by MLR is heavily underestimated and the R^2 is 0.50, which is much lower than that of other models. Compared to MLR, the black-box data science models such as random forest and ANN, can capture the non-linear relationships between variables and DFMC without using the physics process. The random forest and ANN are much better than MLR, but still cannot reach the performance level of the process-based model, FSMM. The LSTM model achieved R^2 of 0.91, RMSE of 3.24% and MAE of 1.97%, which is significantly higher than that of MLR, random forest and ANN. Compared to FSMM, LSTM has a close R^2 , but lower RMSE and MAE. This demonstrates that knowledge gaps of the process model may be closing as long as the information of the data is fully mined and used efficiently. If the output of the process-based model along with the variables are used as input of the FSMM-LSTM model, the results are the best, with R^2 of 0.96, RMSE of 2.21% and MAE of 1.41%.

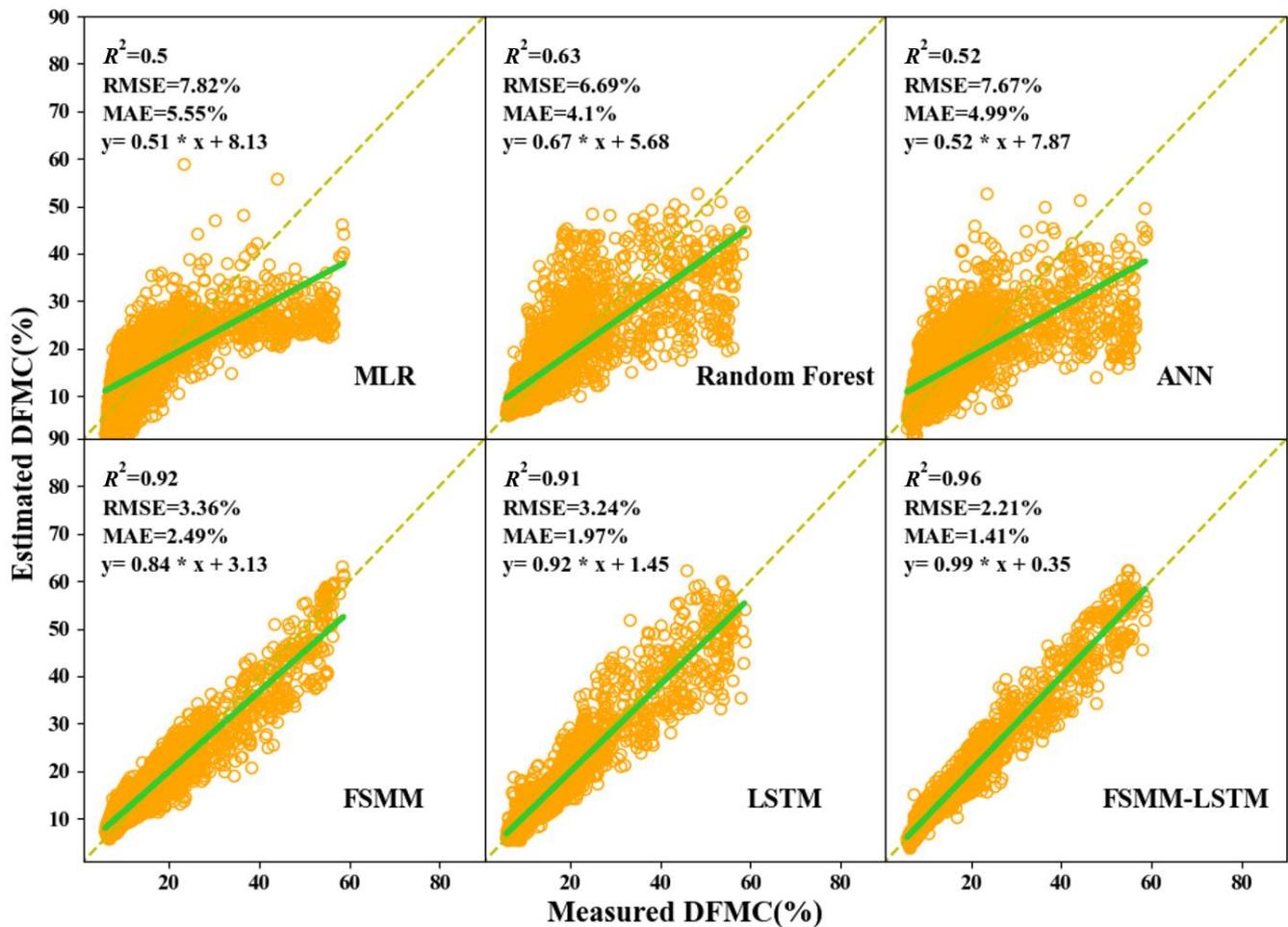


Figure 6. Performances of all models, including MLR, Random Forest, ANN, FSMM, LSTM, FSMM-LSTM.

To compare the performance of the six models in time-series more accurately, we provide their time series pattern in Figure 7. The MLR model is still the worst, which will overestimate or underestimate the DFMC. The results, estimated by random forest and ANN algorithms, can basically agree with measurements in the time series, but there are still plenty of significant overestimations and underestimations. Furthermore, the process-based FSMM model, generates a satisfying result, except for the underestimates from May 20 to May 28 and August 14 to 16. The LSTM model generally shares a similar underestimation as FSMM, but it is more minor. The FSMM-LSTM algorithm still achieved the best results, that the estimated results are in perfect agreement with the measured values, with no overestimations and underestimations over the entire time series.

In the practical application of the model, a non-negligible aspect is computation efficiency. Thus, we listed the calibration time and the test time for all models on our dataset in Table 3. When the process-based model FSMM was used, the computing time is enormous on both calibration (53.75 h) and test (7.53 h). However, all the data-driven models including the LSTM network did not take much time to train and test. Of all data-driven models, the LSTM network showed a comparative result with FSMM with higher efficiency. When the FSMM model and the LSTM network were combined, the total time cost is the same as FSMM, but the test time cost is very tiny.

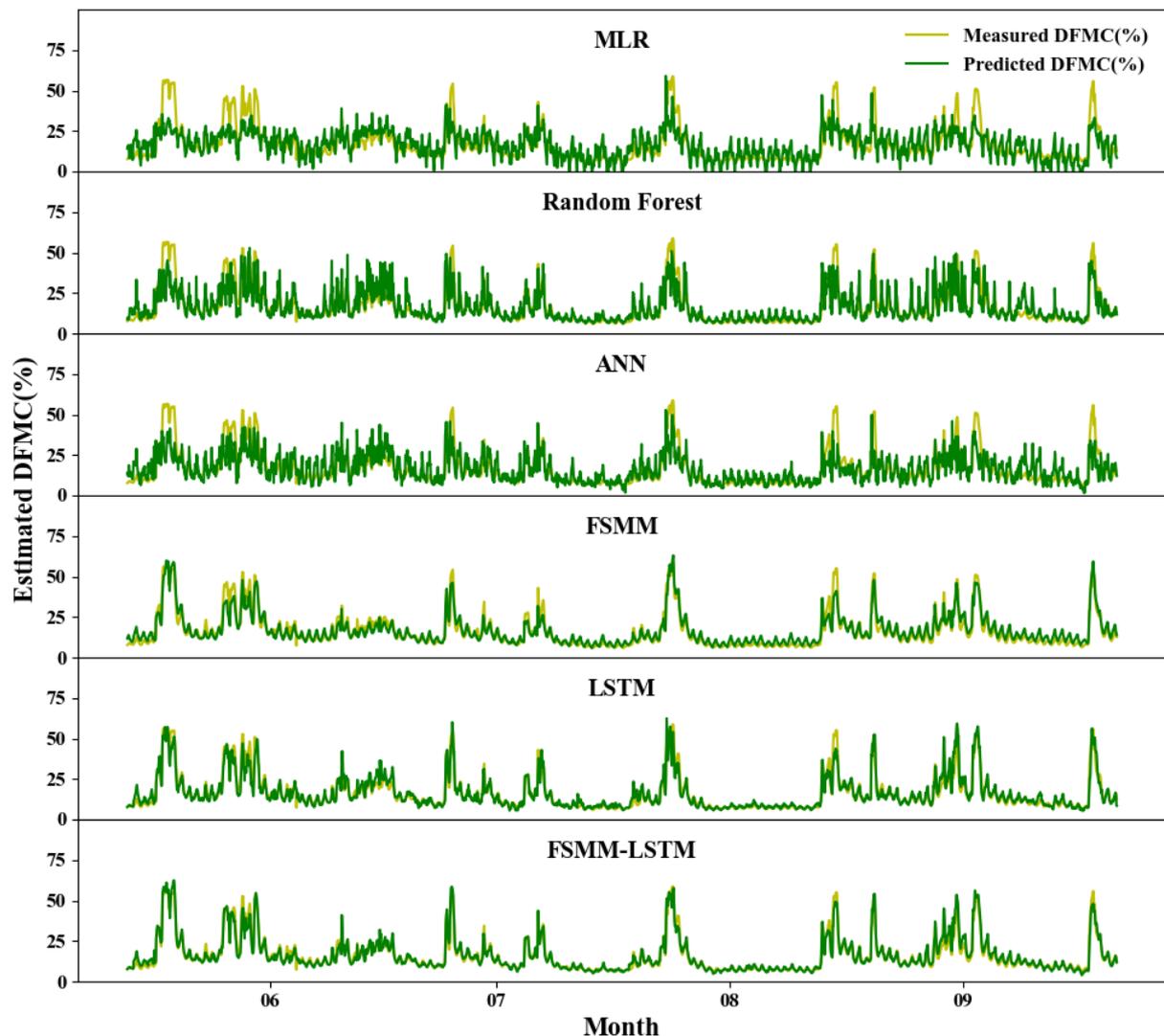


Figure 7. The time series patterns of DFMC predicted by six models, the gold line is measured DFMC and the green line represents predicted DFMC.

Table 3. The performance and the time cost of each model.

Model	MAE (%)	RMSE (%)	R^2	Calibration Time	Test Time
MLR	5.55	7.82	0.5	<2 (second)	<1 (second)
Random Forest	4.1	6.69	0.63	<4 (second)	<1 (second)
ANN	4.99	7.67	0.52	<4 (second)	<1 (second)
FSMM	2.49	3.36	0.92	53.75 (hour)	7.53 (hour)
LSTM	1.97	3.24	0.91	<2 (minute)	<2 (second)
FSMM-LSTM	1.41	2.21	0.96	60.31 (hour)	<2 (second)

5. Discussion

The classical deep learning network-LSTM and its combination with a physics model FSMM were introduced to estimate DFMC. Furthermore, the effective process-based model FSMM and excellent empirical methods (MLR, random forest and ANN) were implemented for comparison.

Our results suggest that the MLR model has the worst performance, this may be because the linear model is insufficient to characterize the non-linear relationship between the input variables and the DFMC. Except for this, when black-box data science models

such as random forest and ANN were used, the result did not become much better. The reason is that although random forest and ANN try to learn the non-linear relationships between drivers and DFMC, they cannot capture the information on a time series, that is what the LSTM network performs better in. The results of the LSTM network showed that the information in the data if it is fully mined, may help in closing knowledge gaps of the process model. Even though the process-based model FSMM performed well as the LSTM network, it is at the expense of enormous time (total about 61.28 h), which may be due to the continuous iteration hour by hour and the 3600 times incessant iteration in each hour. When we combined LSTM and FSMM (FSMM-LSTM), we can achieve even better results. This is because the output of FSMM contains important physical information about the dynamics of DFMC which when coupled with powerful data science frameworks such as the LSTM network, can result in great improvements in R^2 and RMSE (Table 3).

The LSTM network is a classic deep learning algorithm that performs well in capturing the temporal relationship in data. Furthermore, there are plenty of variants of the LSTM network. For example, SLSTM [53], in which the quality variable vector is additionally regarded as the input of the intermediate cell and the three gates. In addition, when the residual connection was applied to the LSTM network, a more efficient deep learning algorithm was developed [65]. It was beyond the scope of our study to investigate whether variants of the LSTM network performed best. We introduced the classic LSTM network only to show the superior performance of the deep learning network in estimating DFMC.

This study focuses on 10-h DFMC while there are four sizes of fuel in total: 1-h, 10-h, 100-h, 1000-h. The results showed that both the LSTM network and FSMM-LSTM model worked excellently in 10-h DFMC estimating. Since LSTM is a data-driven algorithm, it is not difficult to figure out its excellent performance on 1-h DFMC, 100-h DFMC and 1000-h DFMC. The FSMM-LSTM model performed better than the LSTM network, based on the hybrid of the LSTM network and a complex process-based model. Since the FSMM model has shown outstanding performances for all-size fuel [49], it seems that the FSMM-LSTM model also applies to all types of fuel.

In addition, some other methods are effective at estimating DFMC, such as Aguado et al. (2007) [66], Matthews et al. (2010) [23]. Aguado et al. (2007) focused on the lack of calibration of moisture codes to different ecosystem or climate characteristics. They built the relationship between moisture and DFMC based on regression analysis. On the contrary, Matthews et al. (2010) is a complete process-based model which simulated the processes that occur in the fuel with energy and water balance conservation equations. These two methods are the typical example of the empirical method and the process-based model. However, empirical models lack explanations for physical processes, and process-based models may provide an incomplete representation of DFMC. Our study provides a kind of idea to combine them. Going forward, there are some directions that can be exploited future as a continuation of this work. Firstly, the LSTM network used in this study is not a state-of-the-art recurrent neural network and can be replaced. Second, for the specific problem of DFMC estimating, given its temporal and spatial nature, a promising extension would be to explore the temporal and spatial dependencies in DFMC. Third, we present a simple way for constructing a hybrid physics process and data information model by using the output of the process-based model as an input of the data science model, more complex ways of constructing models need to be explored to make the process-based and data science parts are tightly coupled.

6. Conclusions

In this study, a widely used deep learning method LSTM was introduced in DFMC estimation. Furthermore, we proposed a novel approach FSMM-LSTM to estimate DFMC, by using the outputs of process-based model estimates to guide the learning of the LSTM network. By anchoring the LSTM network with a priori knowledge, we found that the proposed physics-guided method was superior across all conditions relative to the other

models, which may help acquire more accurate DFMC estimates to address wildfire risk assessments and fire simulation.

Author Contributions: Study conception and design, B.H.; methodology refining, C.F.; acquisition of data, C.F.; analysis and interpretation of data, B.H. and C.F.; the first draft of the manuscript, C.F.; critical revisions, B.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was founded by the Sichuan Science and Technology Program (Contract No. 2020YFS0058) and the National Natural Science Foundation of China (Contract No. U20A2090).

Data Availability Statement: The data that support the findings of this study are openly available in Github at <https://github.com/dvdkamp/fsmm>, the last visit was on the 20 of May 2021.

Acknowledgments: We remain indebted to Jianpeng Yin, Rui Chen, Tengfei Xiao, Gangqiang An, Gengke Lai and Yanxi Li for invaluable comments in the early stages of this study.

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

References

- Sullivan, A.L. Wildland surface fire spread modelling, 1990–2007. 2: Empirical and quasi-empirical models. *Int. J. Wildland Fire* **2009**, *18*, 369–386. [[CrossRef](#)]
- Shakesby, R.A.; Doerr, S.H. Wildfire as a hydrological and geomorphological agent. *Earth Sci. Rev.* **2006**, *74*, 269–307. [[CrossRef](#)]
- Van Der Werf, G.R.; Randerson, J.T.; Giglio, L.; Van Leeuwen, T.T.; Chen, Y.; Rogers, B.M.; Mu, M.; Van Marle, M.J.E.; Morton, D.C.; Collatz, G.J.; et al. Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data* **2017**, *9*, 697–720. [[CrossRef](#)]
- Wei, X.; Hayes, D.J.; Fraver, S.; Chen, G. Global pyrogenic carbon production during recent decades has created the potential for a large, long-term sink of atmospheric CO₂. *J. Geophys. Res. Biogeosci.* **2018**, *123*, 3682–3696. [[CrossRef](#)]
- Randerson, J.T.; Liu, H.; Flanner, M.; Chambers, S.; Jin, Y.; Hess, P.G.; Pfister, G.; Mack, M.C.; Treseder, K.; Welp, L.; et al. The Impact of Boreal Forest Fire on Climate Warming. *Science* **2006**, *314*, 1130–1132. [[CrossRef](#)] [[PubMed](#)]
- Crutzen, P.J.; Andreae, M.O. Biomass burning in the tropics: Impact on atmospheric chemistry and biogeochemical cycles. *Science* **1990**, *250*, 1669–1678. [[CrossRef](#)] [[PubMed](#)]
- Lelieveld, J.; Evans, J.S.; Fnais, M.; Giannadaki, D.; Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **2015**, *525*, 367–371. [[CrossRef](#)]
- Moritz, M.A.; Batllori, E.; Bradstock, R.A.; Gill, A.M.; Handmer, J.; Hessburg, P.F.; Leonard, J.; McCaffrey, S.; Odion, D.C.; Schoennagel, T.; et al. Learning to coexist with wildfire. *Nature* **2014**, *515*, 58–66. [[CrossRef](#)]
- Rao, K.; Williams, A.P.; Flefil, J.F.; Konings, A.G. SAR-enhanced mapping of live fuel moisture content. *Remote Sens. Environ.* **2020**, *245*, 111797. [[CrossRef](#)]
- Quan, X.; Yebra, M.; Riaño, D.; He, B.; Lai, G.; Liu, X. Global fuel moisture content mapping from MODIS. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *101*, 102354. [[CrossRef](#)]
- Wang, L.; Quan, X.; He, B.; Yebra, M.; Xing, M.; Liu, X. Assessment of the dual polarimetric sentinel-1A data for forest fuel moisture content estimation. *Remote Sens.* **2019**, *11*, 1568. [[CrossRef](#)]
- Yebra, M.; Scortechini, G.; Badi, A.; Beget, M.E.; Boer, M.M.; Bradstock, R.; Chuvieco, E.; Danson, F.M.; Dennison, P.; De Dios, V.R.; et al. Globe-LFMC, a global plant water status database for vegetation ecophysiology and wildfire applications. *Sci. Data* **2019**, *6*, 1–8.
- Quan, X.; He, B.; Yebra, M.; Liu, X.; Liu, X.; Zhung, X.; Cao, H. Retrieval of Fuel Moisture Content from Himawari-8 Product: Towards Real-Time Wildfire Risk Assessment. In Proceedings of the 2018 IEEE International Geoscience & Remote Sensing Symposium (IGARSS), Valencia, Spain, 22–27 July 2018; pp. 7660–7663.
- Rothermel, R.C. *How to Predict the Spread and Intensity of Forest and Range Fires*; US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station: Ogden, UT, USA, 1983; Volume 143.
- Viney, N.R. A review of fine fuel moisture modelling. *Int. J. Wildland Fire* **1991**, *1*, 215–234. [[CrossRef](#)]
- Matthews, S. Dead fuel moisture research: 1991. *Int. J. Wildland Fire* **2014**, *23*, 78–92. [[CrossRef](#)]
- Nolan, R.H.; de Dios, V.R.; Boer, M.M.; Caccamo, G.; Goulden, M.L.; Bradstock, R.A. Predicting dead fine fuel moisture at regional scales using vapour pressure deficit from MODIS and gridded weather data. *Remote Sens. Environ.* **2016**, *174*, 100–108. [[CrossRef](#)]
- Catchpole, E.; Catchpole, W.; Viney, N.; McCaw, W.; Marsden-Smedley, J. Estimating fuel response time and predicting fuel moisture content from field data. *Int. J. Wildland Fire* **2001**, *10*, 215–222. [[CrossRef](#)]
- Nieto, H.; Aguado, I.; Chuvieco, E.; Sandholt, I. Dead fuel moisture estimation with MSG–SEVIRI data. Retrieval of meteorological data for the calculation of the equilibrium moisture content. *Agric. For. Meteorol.* **2010**, *150*, 861–870. [[CrossRef](#)]
- Walding, N.G.; Williams, H.T.; McGarvie, S.; Belcher, C.M. A comparison of the US National Fire Danger Rating System (NFDRS) with recorded fire occurrence and final fire size. *Int. J. Wildland Fire* **2018**, *27*, 99–113. [[CrossRef](#)]
- Bradshaw, L.S. *The 1978 National Fire-Danger Rating System: Technical Documentation*; US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station: Ogden, UT, USA, 1984. [[CrossRef](#)]
- Matthews, S. A process-based model of fine fuel moisture. *Int. J. Wildland Fire* **2006**, *15*, 155–168. [[CrossRef](#)]

23. Matthews, S.; Gould, J.; McCaw, L. Simple models for predicting dead fuel moisture in eucalyptus forests. *Int. J. Wildland Fire* **2010**, *19*, 459–467. [[CrossRef](#)]
24. de Dios, V.R.; Fellows, A.W.; Nolan, R.H.; Boer, M.M.; Bradstock, R.A.; Domingo, F.; Goulden, M.L. A semi-mechanistic model for predicting the moisture content of fine litter. *Agric. For. Meteorol.* **2015**, *203*, 64–73. [[CrossRef](#)]
25. Lee, H.; Won, M.; Yoon, S.; Jang, K. Estimation of 10-Hour Fuel Moisture Content Using Meteorological Data: A Model Inter-Comparison Study. *Forests* **2020**, *11*, 982. [[CrossRef](#)]
26. Sharples, J.; McRae, R.; Weber, R.; Gill, A.M. A simple index for assessing fuel moisture content. *Environ. Model. Softw.* **2009**, *24*, 637–646. [[CrossRef](#)]
27. Alves, M.; Batista, A.; Soares, R.; Ottaviano, M.; Marchetti, M. Fuel moisture sampling and modeling in *Pinus elliottii* Engelm. plantations based on weather conditions in Paraná-Brazil. *iFor. Biogeosci. For.* **2009**, *2*, 99. [[CrossRef](#)]
28. Anderson, H.E. Moisture diffusivity and response time in fine forest fuels. *Can. J. For. Res.* **1990**, *20*, 315–325. [[CrossRef](#)]
29. Simard, A. *The Moisture Content of Forest Fuels—I. A Review of the Basic Concepts*; Information Report FF-X-14; Canadian Department of Forest and Rural Development, Forest Fire Research Institute: Ottawa, ON, Canada, 1968.
30. Van Wagner, C. *Equilibrium Moisture Contents of Some Fine Forest Fuels in Eastern Canada*; Information Report PS-X-36; Canadian Forestry Service, Petawawa Forest Experiment Station: Chalk River, ON, Canada, 1972.
31. Van Wagner, C.E. *Development and Structure of the Canadian Forest Fire Weather Index System*; Forestry Technical Report 35; Canadian Forest Service, Petawawa National Forestry Institute: Chalk River, ON, Canada, 1987.
32. Bilgili, E.; Coskuner, K.A.; Usta, Y.; Saglam, B.; Kucuk, O.; Berber, T.; Goltas, M. Diurnal surface fuel moisture prediction model for Calabrian pine stands in Turkey. *iFor. Biogeosci. For.* **2019**, *12*, 262. [[CrossRef](#)]
33. Venäläinen, A.; Heikinheimo, M. Meteorological data for agricultural applications. *Phys. Chem. Earth* **2002**, *27*, 1045–1050. [[CrossRef](#)]
34. Tamai, K. Estimation of model for litter moisture content ratio on forest floor. In *Soil-Vegetation-Atmosphere Transfer Schemes and Large-Scale Hydrological Models: Proceedings of an International Symposium (Symposium S5) Held During the Sixth Scientific Assembly of the International Association of Hydrological Sciences (IAHS), Maastricht, The Netherlands, 18–27 July 2001*; Volume 270, pp. 53–58.
35. Byram, G.; Nelson, R. *An Analysis of the Drying Process in Forest Fuel Material*; US Department of Agriculture Forest Service, Southern Research Station: Asheville, NC, USA, 2015; Volume 200, pp. 1–45.
36. Nelson, R.M., Jr. A method for describing equilibrium moisture content of forest fuels. *Can. J. For. Res.* **1984**, *14*, 597–600. [[CrossRef](#)]
37. Tanskanen, H.; Venäläinen, A. The relationship between fire activity and fire weather indices at different stages of the growing season in Finland. *Boreal Environ. Res.* **2008**, *13*, 285–302.
38. Nelson, R.M., Jr. Prediction of diurnal change in 10-h fuel stick moisture content. *Can. J. For. Res.* **2000**, *30*, 1071–1087. [[CrossRef](#)]
39. Quan, X.; Li, Y.; He, B.; Cary, G.; Lai, G. Application of Landsat ETM+ and OLI Data for Foliage Fuel Load Monitoring Using Radiative Transfer Model and Machine Learning Method. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 5100–5110. [[CrossRef](#)]
40. Wolanin, A.; Camps-Valls, G.; Gómez-Chova, L.; Mateo-García, G.; Guanter, L. Estimating Crop Primary Productivity with Sentinel-2 and Landsat 8 using Machine Learning Methods Trained with Radiative Transfer Simulations. *Remote Sens. Environ.* **2019**, *225*, 441–457. [[CrossRef](#)]
41. Mishra, S.; Molinaro, R. Physics Informed Neural Networks for Simulating Radiative Transfer. *J. Quant. Spectrosc. Radiat. Transf.* **2020**, *270*, 107705. [[CrossRef](#)]
42. Karpatne, A.; Watkins, W.; Read, J.; Kumar, V. Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling. *arXiv* **2017**, arXiv:1710.11431. Available online: <https://arxiv.org/pdf/1710.11431.pdf> (accessed on 15 May 2021).
43. Daw, A.; Thomas, R.Q.; Carey, C.C.; Read, J.S.; Appling, A.P.; Karpatne, A. Physics-Guided Architecture (PGA) of Neural Networks for Quantifying Uncertainty in Lake Temperature Modeling. In Proceedings of the 2020 SIAM International Conference on Data Mining, Society for Industrial & Applied Mathematics (SIAM), Cincinnati, OH, USA, 7–9 May 2020; pp. 532–540.
44. Bolderman, M.; Lazar, M.; Butler, H. Physics-Guided Neural Networks for Inversion-based Feedforward Control applied to Linear Motors. *arXiv* **2021**, arXiv:2103.060922021.
45. Greff, K.; Srivastava, R.K.; Koutnik, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A Search Space Odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2222–2232. [[CrossRef](#)]
46. Zhanga, J.; Yan, Z.; Zhanga, X.; Ming, Y.; Yangb, J. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *J. Hydrol.* **2018**, *561*, 918–929. [[CrossRef](#)]
47. Ata, A.A.; Tiantian, Y.; Hsu, K.; Sorooshian, S.; Lin, J.; Peng, Q. Short-Term Precipitation Forecast Based on the PERSIANN System and LSTM Recurrent Neural Networks. *J. Geophys. Res. Atmos.* **2018**, *123*, 12543–12563.
48. Hosman, T.; Vilela, M.; Milstein, D.; Kelemen, J.N.; Brandman, D.M.; Hochberg, L.R.; Simeral, J.D. BCI decoder performance comparison of an LSTM recurrent neural network and a Kalman filter in retrospective simulation. In Proceedings of the 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER), San Francisco, CA, USA, 20–23 March 2019; pp. 1066–1071.
49. Van, D.; Moore, R.D.; Mckendry, I.G. A model for simulating the moisture content of standardized fuel sticks of various sizes. *Agric. For. Meteorol.* **2017**, *236*, 123–134.
50. Tamai, K.; Goto, Y. The Estimation Of Temporal And Spatial Fluctuations In A Forest Fire Hazard Index—The Case Of A Forested Public Area In Japan. *WIT Trans. Ecol. Environ.* **2008**, *119*, 397–404.

51. Carlson, J.D.; Bradshaw, L.S.; Nelson, R.M.; Bensch, R.R.; Jabrzemski, R. Application of the Nelson model to four timelag fuel classes using Oklahoma field observations: Model evaluation and comparison with National Fire Danger Rating System algorithms. *Int. J. Wildland Fire* **2007**, *16*, 204–216. [[CrossRef](#)]
52. Sahin, S.O.; Kozat, S.S. Nonuniformly Sampled Data Processing Using LSTM Networks. *EEE Trans. Neural Netw. Learn. Syst.* **2018**, *30*, 1452–1461.
53. Yuan, X.; Li, L.; Wang, Y. Nonlinear Dynamic Soft Sensor Modeling with Supervised Long Short-Term Memory Network. *IEEE Trans. Ind. Inform.* **2020**, *16*, 3168–3176. [[CrossRef](#)]
54. Ketkar, N. Introduction to Keras. In *Deep Learning with Python*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 97–111.
55. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. TensorFlow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
56. Liu, H.; Zhou, Q.; Zhang, S.; Deng, X. Estimation of Summer Air Temperature over China Using Himawari-8 AHI and Numerical Weather Prediction Data. *Adv. Meteorol.* **2019**, *2019*, 2385310. [[CrossRef](#)]
57. Ahmad, M.W.; Mourshed, M.; Rezgui, Y. Trees vs. Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy Build.* **2017**, *147*, 77–89. [[CrossRef](#)]
58. Tranmer, M.; Elliot, M. *Multiple Linear Regression*; Cathie Marsh Centre for Census and Survey Research: Manchester, UK, 2008.
59. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
60. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: New York, NY, USA, 2013; Volume 103.
61. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
62. Jenkins, B.; Tanguay, A. *Handbook of Neural Computing and Neural Networks*; MIT Press: Boston, MA, USA, 1995.
63. Bulsari, A. Some analytical solutions to the general approximation problem for feedforward neural networks. *Neural Netw.* **1993**, *6*, 991–996. [[CrossRef](#)]
64. Wu, Y.-C.; Feng, J.-W. Development and application of artificial neural network. *Wirel. Pers. Commun.* **2018**, *102*, 1645–1656. [[CrossRef](#)]
65. Gui, T.; Zhang, Q.; Zhao, L.; Lin, Y.; Peng, M.; Gong, J.; Huang, X. Long short-term memory with dynamic skip connections. *Proc. AAAI Conf. Artif. Intell.* **2019**, *33*, 6481–6488. [[CrossRef](#)]
66. Aguado, I.; Chuvieco, E.; Borén, R.; Nieto, H. Estimation of dead fuel moisture content from meteorological data in Mediterranean areas. Applications in fire danger assessment. *Int. J. Wildland Fire* **2007**, *16*, 390–397. [[CrossRef](#)]