Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data—A Machine Learning Approach

Iftikhar Ali, Fiona Cawkwell, Edward Dwyer, and Stuart Green

***Abstract*—More than 80% of agricultural land in Ireland is grassland, which is a major feed source for the pasture based dairy farming and livestock industry. Many studies have been undertaken globally to estimate grassland biomass by using satellite remote sensing data, but rarely in systems like Ireland's intensively managed, but small-scale pastures, where grass is grazed as well as harvested for winter fodder. Multiple linear regression (MLR), artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) models were developed to estimate the grassland biomass (kg dry matter/ha/day) of two intensively managed grassland farms in Ireland. For the first test site (Moorepark) 12 years (2001–2012) and for second test site (Grange) 6 years (2001–2005, 2007) of *in situ* measurements (weekly measured biomass) were used for model development. Five vegetation indices plus two raw spectral bands (RED=red band, NIR=Near Infrared band) derived from an 8-day MODIS product (MOD09Q1) were used as an input for all three models. Model evaluation shows that the ANFIS ($R^2_{\mathrm{Moorepark}} = 0.85$, $\mathrm{RMSE}_{\mathrm{Moorepark}} = 11.07$; $R^2_{\mathrm{Grange}} = 0.76$, $\mathrm{RMSE}_{\mathrm{Grange}} = 15.35$) has produced improved estimation of biomass as compared to the ANN and MLR. The proposed methodology will help to better explore the future inflow of remote sensing data from spaceborne sensors for the retrieval of different biophysical parameters, and with the launch of new members of satellite families (ALOS-2, Radarsat-2, Sentinel, TerraSAR-X, TanDEM-X/L) the development of tools to process large volumes of image data will become increasingly important.**

***Index Terms*—Biomass estimation, machine learning, managed grassland, remote sensing, time series.**

I. Introduction

GRASSLANDS are one of the major and crucial components of the terrestrial ecosystem [1] and most prevalent and widespread global land cover types. Grasslands cover about 40.5% of the Earth's surface [2]–[4] and after forests, grasslands are the major source (about 30%) of carbon sink [5], [6] and thus play a very important role in regulating the global carbon cycle [4]. The demand and consumption of dairy products are increasing globally [7], [8] and in order to meet this demand, an equivalent growth in livestock has to be maintained.

Grasslands are the major feed source for grazing livestock, and the amount of above ground biomass will determine the pasture's carrying capacity—the maximum number of animals that can graze a pasture for a set period without harming it. Grazed grass is the cheapest feed for livestock, and for that reason it is very important to manage grasslands because better management will result in low cost high quality grass. Based on these management approaches, grasslands can be categorized into broader management strategies: 1) unmanaged (natural) and 2) managed (agricultural pastures) grasslands [9]. The term "grassland management" in the context of this research includes weed control, removing dead plants, mowing, clipping, assessment of growth rate, grazing length, and utilization of grassland [10].

Grassland biomass can be estimated by using both ground based conventional methods and remote sensing technology. Existing ground-based methods include:

1) *Visual:* the visual assessment by human eye (expert or farmer), this method is spatially sparse with limited performance [11].
2) *Cut and dry (Clipping):* grass is harvested from the paddock and is dried and weighed to get the dry matter (DM) yield.
3) *Rising plate meter:* both mechanical and electronic plate meters work on the same principle, where the plate rises up and down the shaft taking measurements of grass height [12]–[14].
4) *Field spectrometry:* can also be used for above ground biomass estimation where collected spectra are converted into reflectance and calibration is performed from biomass samples [15], [16].

Conventional ground based methods are subjective, time consuming and are feasible (or applicable) only for small scale assessment and monitoring of grasslands [17].

More advanced and spatially extensive grassland monitoring methods include the use of remotely sensed data. Remote sensing data can be acquired from sensors (optical and/or radar) mounted on different platforms [18], for example in the case of airborne remote sensing the sensor is mounted on aircraft, helicopters or unmanned aerial vehicles while in case of satellite remote sensing the sensor(s) is mounted on a spacecraft. Airborne remote sensing is good for cloud free data acquisition at a small scale, as the aircraft can fly under the cloud cover at an optimal time for data collection. Airborne remote sensing data have been

Manuscript received September 9, 2015; revised December 14, 2015; accepted April 21, 2016. This work was supported by the Teagasc Walsh Fellowship Program.

I. Ali and F. Cawkwell are with the Department of Geography, University College Cork, Cork, Ireland (e-mail: iffi.math@gmail.com; f.cawkwell@ucc.ie).

E. Dwyer is with the European Centre for Information on Marine Science and Technology, EurOcean, Lisbon 1249-074, Portugal (e-mail: ned.dwyer@eurocean.org).

S. Green is with the Spatial Analysis Unit, Teagasc, Dublin, Ireland (e-mail: stuart.green@teagasc.ie).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2016.2561618

used for vegetation change detection [19] and discrimination [20] at a local scale. Similarly, for grassland monitoring, Curran and Williamson [21] have used airborne multispectral scanner data for the mapping of leaf area index, while in another study Darvishzadeh *et al.* [22] used hyperspectral airborne imagery. Despite the advantages (timely and flexible in acquisitions, high spatial resolution) of airborne remote sensing data, the approach is still considered expensive [23] for consistent large scale applications, and impractical for the development of operational tools. At present, in order to overcome these limitations, satellite remote sensing remains the best available alternative, where sensors with different microwave wavelengths (TerraSAR-X, Radarsat), spectral bands (Landsat, QuickBird, Hyperion), resolution and revisit time can be used operationally. Data from both optical and SAR instruments are being used for grassland related investigations [24], [25].

Since the launch of Landsat-1 in 1972, satellite remote sensing data have been used for agricultural activities e.g., biomass estimation [26], soil moisture [27], water consumption [28], discrimination of different crop types [29], and monitoring of agricultural drought [30]. With the development of satellite sensors with high spatio-temporal resolution and wide area coverage, agriculture remote sensing has moved a step further towards "precision agriculture" whereby growth rates can be monitored [31], [32], inter and intra field variability mapped [33], poor/underperforming areas identified [34], and decision support systems developed [35]–[38].

Over the past 40 years a number of methods have been developed for grassland biomass estimation based on satellite remote sensing data, and the technology is now mature enough for the monitoring of detailed grassland management activities. Based on a review of past work, satellite driven grassland biomass estimation methodologies can be categorized into three broader groups: 1) using vegetation indices (VIs), 2) biophysical simulation models, and 3) machine learning algorithms [39].

A. Use of Vegetation Indices for Grassland Biomass Estimation

The use of satellite driven VIs in combination with *in situ* measurements [40], [41] for the development of regression models for grassland biomass estimation is the most popular and well-studied approach [42]–[50]. Many researchers have investigated the application of different VIs derived from satellite imagery (e.g., QuickBird, MODIS, Landsat) and developed different regression models (e.g., linear, power, logarithmic, multiple linear) for grassland biomass estimation [17], [51]–[53]. Very high accuracies of the vegetation index based regression models for biomass estimation have been reported in the literature [17], [54]–[60], but their major limitation is that these models are site specific and do not have the capability to learn the highly non-linear and complex patterns in the data.

B. Use of Biophysical Simulation Models for Grassland Biomass Estimation

The LINGRA simulation model has been designed for the prediction of grassland (perennial rye grass) productivity in Europe [61], [62]. In a recent study, Maselli *et al.* [63] used the C-Fix parametric model for grassland gross primary production in combination with *in situ* measurements and remote sensing data. This approach of data assimilation has frequently been used for crop (i.e., wheat, rice, maize) monitoring [64]–[66], but has not been fully explored yet for grassland monitoring.

C. Use of Machine Learning Algorithms for Grassland Biomass Estimation

Unlike crops [67]–[70] and forests [71] the number of studies on the use of machine learning algorithms for remote sensing based grassland biomass estimation is limited [72]. Xie *et al.* [73] and Yang *et al.* [74] reported the successful application of an artificial neural network (ANN) approach for grassland yield estimation based on utilization of satellite driven VIs. In another study Clevers *et al.* [75] used a support vector machine approach for grassland biomass estimation based on airborne remote sensing data.

The objective of this paper is to estimate the biomass of managed grasslands where weekly grass growth (kg DM/ha/day) is recorded on a regular basis. Three different methods were used for grassland biomass estimation, namely: Multiple Linear Regression (MLR), ANNsand Adaptive-Neuro Fuzzy Inference Systems (ANFIS). To the best of our knowledge only a few studies [73], [74] have reported on the application of ANN with remote sensing data for grassland biomass estimation; and there has been no work published to date on the application of ANFIS using satellite imagery for grassland biomass estimation. After the publication of Jang [76] research, where the framework of ANFIS was introduced, this modeling approach has been used in various disciplines [77]–[84]; and some studies reported the performance comparison between ANN and ANFIS, but not previously in the context of deriving biophysical parameters from a satellite image. It is evident from a number of previous studies that in some cases ANN performs better than ANFIS [85], but in most of the cases ANFIS performs as well as/or better than the ANN [86]–[96]. This trend of model performance varies between application domains, and performance also depends upon a number of factors e.g., the quality and size of datasets and underlying problem formulation. This paper will explore some of these issues in more depth for two Irish sites with differing data inputs.

II. Materials and Methods

A. Study Sites

The Moorepark and Grange study sites are Teagasc (the Irish agriculture and food development authority) research farms located in the south ($50^\circ\ 07'$ N, $08^\circ\ 16'$ W) and north east ($53^\circ\ 30'$ N, $06^\circ\ 40'$ W) of Ireland respectively (see Fig. 1). Teagasc research farms in Ireland have been closely monitored for many years, providing a valuable source of grassland biomass (intensively managed grassland), meteorological and farm management data. This study uses *in situ* data of weekly biomass (kg DM/ha/day) from 2001 to 2012 for Moorepark (area: 100 ha) and from 2001 to 2005 and 2007 for Grange

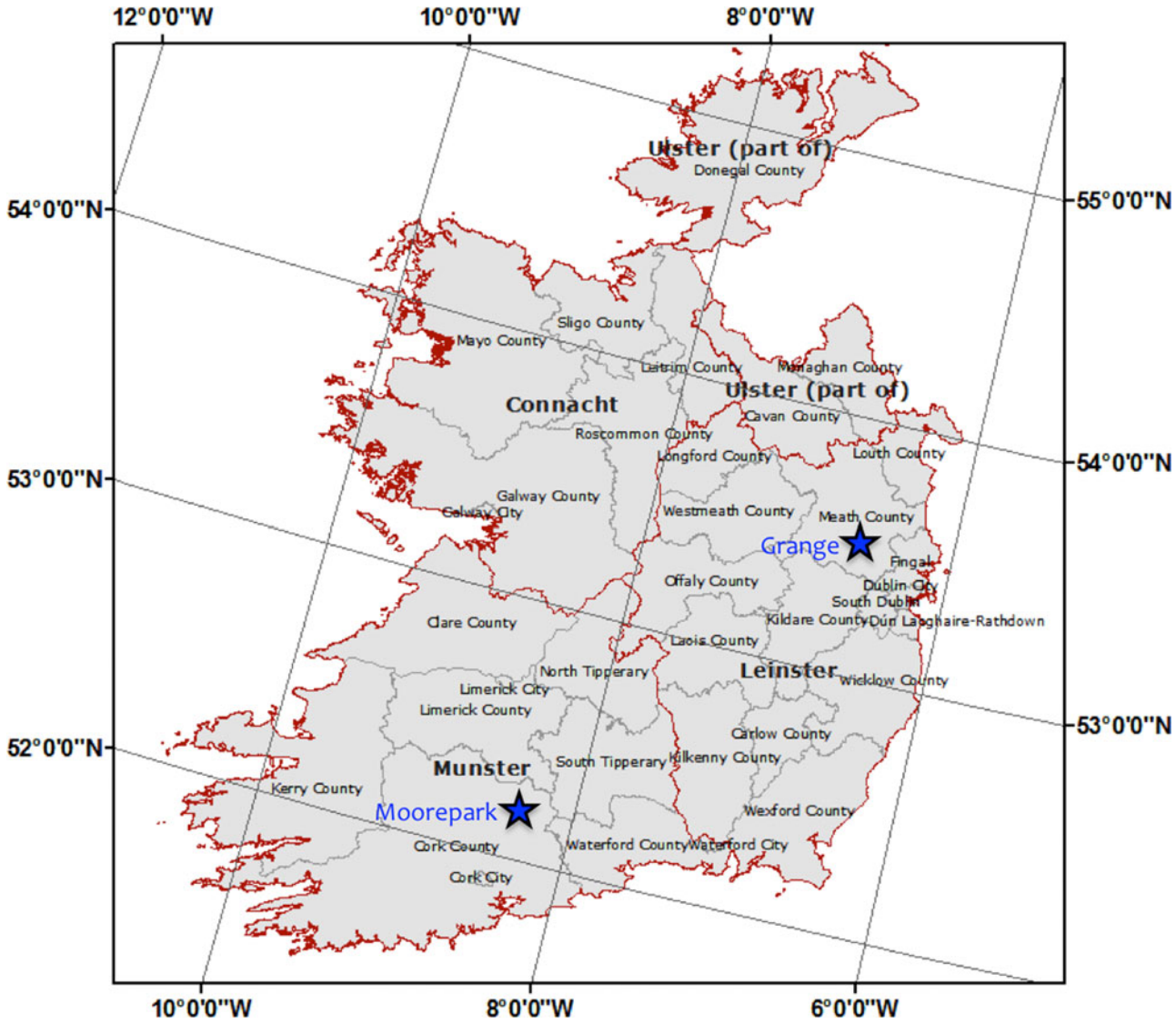

Fig. 1. Two Teagasc research farm study sites (Blue stars: Moorepark and Grange) where weekly *in situ* data are collected.

(area: 71.3 ha). For Moorepark, annual mean temperature ranges from 9.4 to 10.1°C and for Grange it is 8.8 – 11°C, while the annual average rainfall varies between 854 and 1208 mm for Moorepark and between 601.5 and 1065.8 mm for Grange study site (see Fig. 2).

B. Data Used

1) Remotely Sensed Data: A time series (46 images per year) of 250 m MODIS Terra surface reflectance 8-day composite (MOD09Q1), and 500 m MODIS Terra surface reflectance 8-day composite (MOD09A1) images were freely downloaded from the NASA Land Process Distributed Active Archive Center (https://lpdaac.usgs.gov/lpdaac/get_data/glovis) for the Moorepark study site from 2001 to 2012 and for the Grange study site from 2001 to 2007. For accurate estimation of the grass growth index based on satellite data, the date of ground truth data collection and satellite image acquisition are required in order to establish a true correlation between the observed biophysical parameters and satellite driven vegetation indicators. The day of pixel composite information was extracted from the MOD09A1 product as suggested by Guindin-Garcia *et al.* [97] and applied to the 250m product, which was used for the model development.

2) Field Data: Both the test sites consist of managed grassland pasture fields, and different grassland related biophysical parameters have been recorded for many years. In this study the grassland weekly biomass (kg DM/ha/day) values have been used. For the Moorepark test site, 12 years (2001–2012) of *in situ* measurements of grassland biomass are used, while for Grange 6 years (2001–2005 and 2007) of field data are analyzed. Biomass (DM) for each paddock is calculated by cutting and drying a grass strip of approximately 1 m wide and 3 m long (see Fig. 3) from which biomass and growth rate for the whole farm are calculated. Fig. 4 shows a summary of ground data collected for both the test sites.

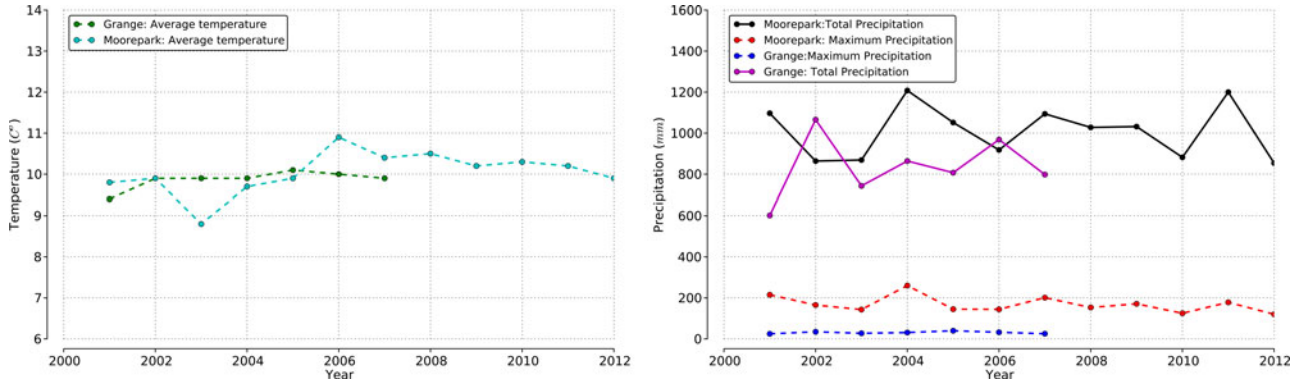

Fig. 2. Meteorological profiles of annual average temperature and annual maximum and total precipitation for Moorepark and Grange study sites.

Fig. 3. *In situ* data collection for each individual paddock using the clipped method, where a strip of grass approximately 1 metre wide and 3 metres long is cut and dried to estimate the biomass (kg DM/ha/day).

C. Data Preprocessing and Study Design

Both MODIS products were downloaded in HDF file format and a Python script was written to extract the reflectance values five VIs were calculated as shown in Table I:

Calculated VIs were filtered using the Savitzky–Golay algorithm, which is widely used to smooth high frequency variability, such as the spiky nature of the time series of VIs. This process was implemented in Python in order to smooth out noise in the time series and fill gaps resulting from cloud-induced missing data. Principal Component Analysis (PCA) was then applied to reduce the data dimensionality and variable dependencies. Fig. 5 shows the systematic workflow of this approach.

D. Model Development

1) Multiple Linear Regression Model: The MLR approach is used where there is more than one predictor variable, and to find linear relationships between the dependent and independent variables [103]. Five VIs and two raw bands (RED, NIR) were used as independent predictor variables for grassland biomass (kg DM/ha/day). The model formulation is as follows:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_k X_{ik} + \varepsilon_i \quad (1)$$

where

$Y_i =$	model reponse
$\beta_0 =$	intercept
$\beta_1, \ldots, \beta_k =$	slopes or regression coefficients
$X_{i1}, \ldots, X_{ik} =$	predictor variables
$\varepsilon_i =$	independent variables that are normally distributed with 0 mean and constant variance.

2) Artificial Neural Networks Model: ANNs belong to the family of machine learning algorithms, where the computational models have a great ability to adapt, learn and generalize the

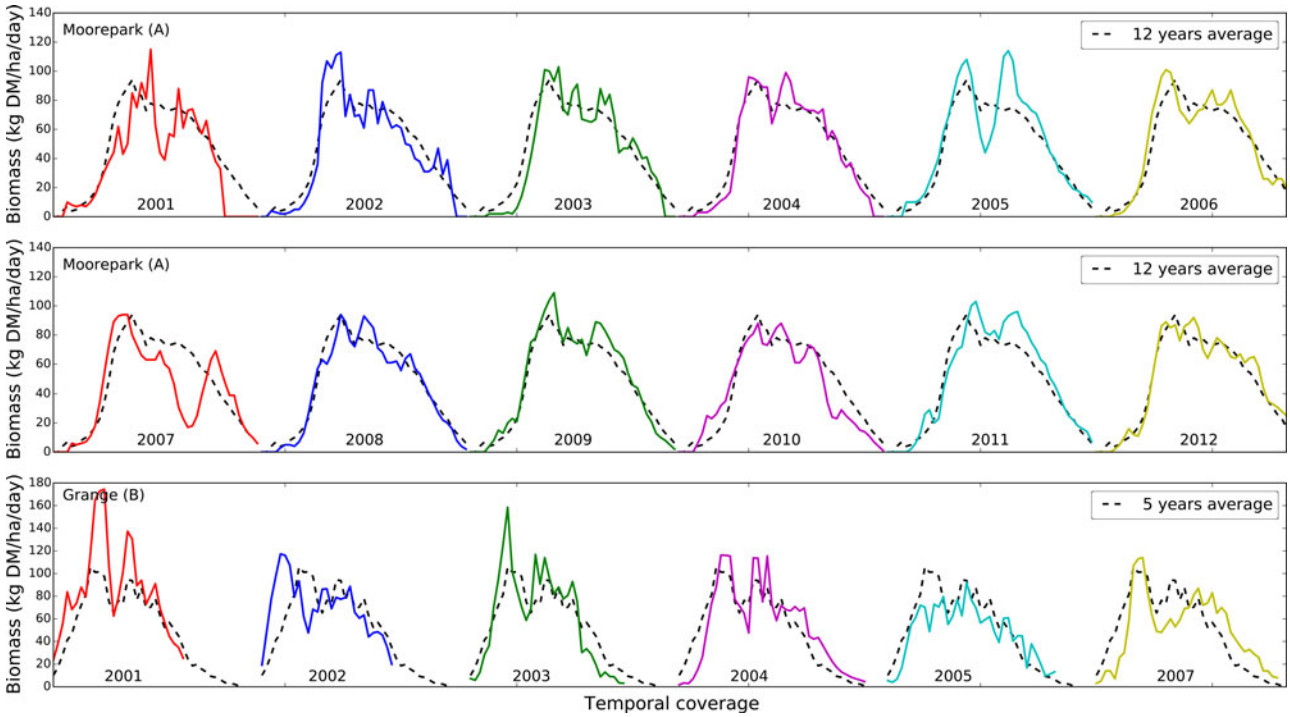

Fig. 4. Weekly biomass measurements for (A): Moorepark 12 years *in situ* measurements, (B) Grange 6 years *in situ* measurements: the black line represents the weekly biomass (kg DM/ha/day) value for each year, and red dotted lines show 12 and 6 years average biomass (kg DM/ha/day).

TABLE I
LIST OF VEGETATION INDICES USED

Vegetation Index	Acronyms	Formula	Description	Reference
Normalized Difference Vegetation Index	NDVI	$\frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+\mathrm{RED}}$	is widely used for the separation of green vegetation and background soil brightness with values ranging from -1 to $+1$, where, -1 represents non-vegetative and $+1$ vegetative area.	[98]
Enhanced Vegetation Index–2	EVI2	$2.5\left(\frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+2.4\mathrm{RED}+1}\right)$	is a modified form of NDVI—highly sensitive to vegetation, capable of decoupling canopy background signal, and reduces the atmospheric influence.	[99]
Soil Adjusted Vegetation Index	SAVI	$(1+L)\left(\frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+\mathrm{RED}+L}\right)$	in case of low vegetation cover soil noise causes a poor estimation of vegetation biomass. In order to overcome this limitation SAVI is used—to minimize the contribution of soil background signals by using a soil adjustment factor L. Huete (1988) suggested a value of $\mathrm{L}=0.5$ in most conditions.	[100]
Modified Soil Adjusted Vegetation Index	MSAVI	$\frac{1}{2}[(2\mathrm{NIR}+1)-\sqrt{(2\mathrm{NIR}+1)^2-8(\mathrm{NIR}-\mathrm{RED})}]$	is a modification of SAVI which has a modified soil adjustment factor L, pixels with negative values represents non-vegetative area and pixels with positive values represent vegetative area.	[101]
Optimised Soil Adjusted Vegetation Index	OSAVI	$\frac{\mathrm{NIR}-\mathrm{RED}}{\mathrm{NIR}+\mathrm{RED}+X}$	also belongs to the SAVI family of vegetation indices, in order to minimize the background soil noise, here the factor X is crucial for the minimization of background soil noise, Rondeaux et al. (1996) found an optimized value of $\mathrm{X}=0.16$.	[102]

complex and complicated patterns hidden in the data. ANN works like a biological neuron where the information flows in are processed by the neuron and the results flow out [104]. This gives the neuron an ability to react based on previously learned patterns. Scientists replicate this by creating a structure that processes information like a biological neuron does, except this approach is mathematically driven [105], [106]. Fig. 6 shows the example of a single biological (A) and artificial neuron (B).

A single processing unit (an artificial neuron) computes the weighted sum of input data sets and there is always an activation

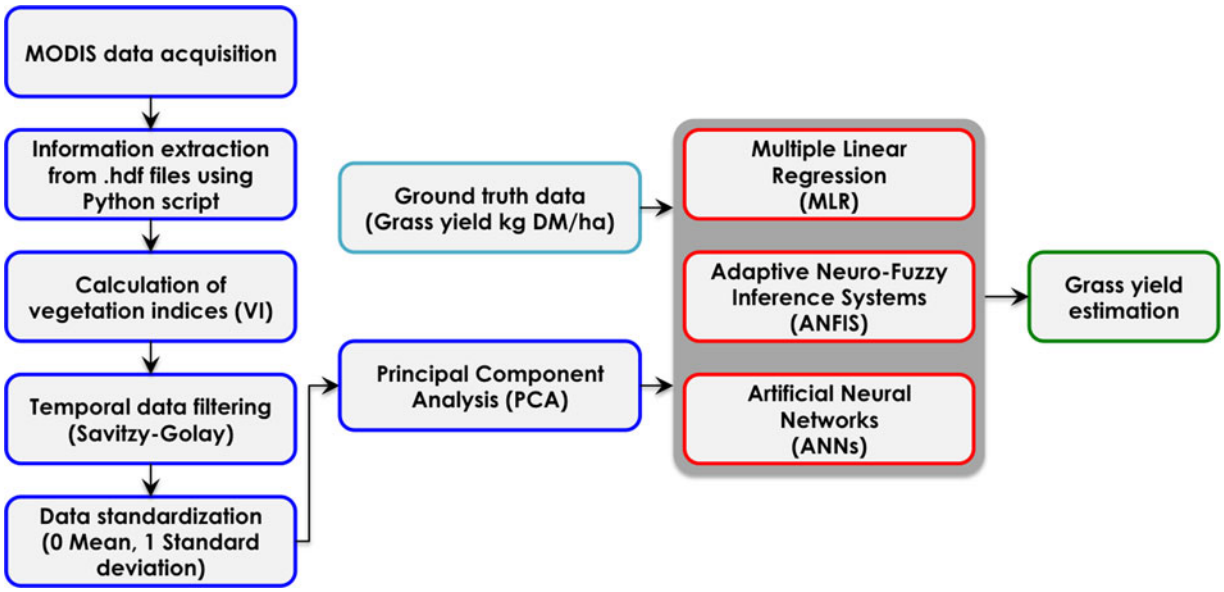

Fig. 5. Study design and methodological workflow scheme.

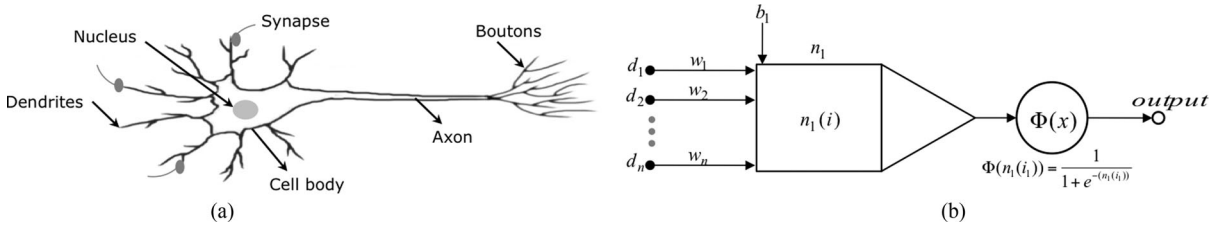

Fig. 6. (a) Biological neuron, (b) unit artificial neuron.

function, which gives the output of the unit. The mathematical representation of an artificial neuron (n_1) at an instance (i_1) and its activation function [105] are given by

$$n_1\ (i_1) = \sum_{j=1}^{n} v_{j1} d_j\ + b_1\ (i_1) \tag{2}$$

where, $d_1, \ldots, d_n$ are inputs, $v_1, \ldots, v_n$ are associated connection weights and b_1 is the bias value, with the activation function sigmoid (3)

$$\Phi\ (x) = \frac{1}{1+e^{-x}}\ . \tag{3}$$

Other possible activation functions could be linear or hyperbolic tangent functions.

For this study, a feed-forward back propagation neural network algorithm [105] was used, where individual neurons (processing units) are arranged in layers where the first layer takes inputs and last layer produces output(s). Neurons in each layer are connected to all the neurons in the next layer and information flows in the forward direction (hence "feed forward"), while there is no connection among the neurons in the same layer. Fig. 7 shows the structure of the multilayer feed-forward back propagation algorithm.

Back propagation is a form of supervised learning algorithm where the input dataset consists of training samples and desired outputs. In back propagation, learning occurs every time an input training sample is fed to the net, and the output of this exercise is compared with the desired results and an error (or deviation from original results) is calculated. The value of error is a quantitative measure, which shows how far away the output is from the desired value. Using the calculated errors, the back propagation-training algorithm then follows the backward pass through the layers from output layer to the input layer in order to adjust the weights, with the ultimate objective being to minimize the error.

3) Adaptive Neuro Fuzzy Inference Systems Model: ANNs have the power of learning patterns, while on the other hand fuzzy logic has the capabilities of reasoning. ANFIS is a fusion or hybrid model that integrates the positive aspects of both ANNs and fuzzy logic in order to construct a robust model that will associate the independent (input values) variables with the dependent (target values) variables with minimum estimation error.

A five layers ANFIS was first introduced by Jang [76], with the capability to incorporate linguistic knowledge (expert opinion) and human like reasoning based on a training data set and a set of IF-THEN fuzzy rules. A unit format for defining fuzzy

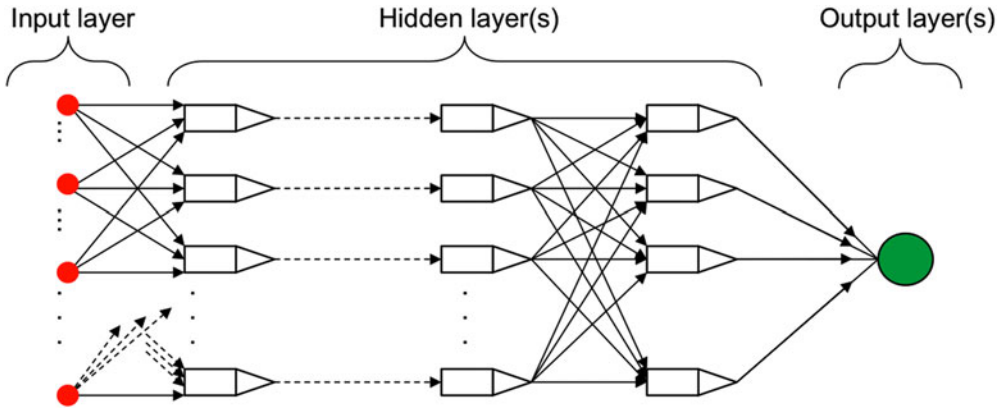

Fig. 7. Structure of multilayer feed–forward back propagation algorithm.

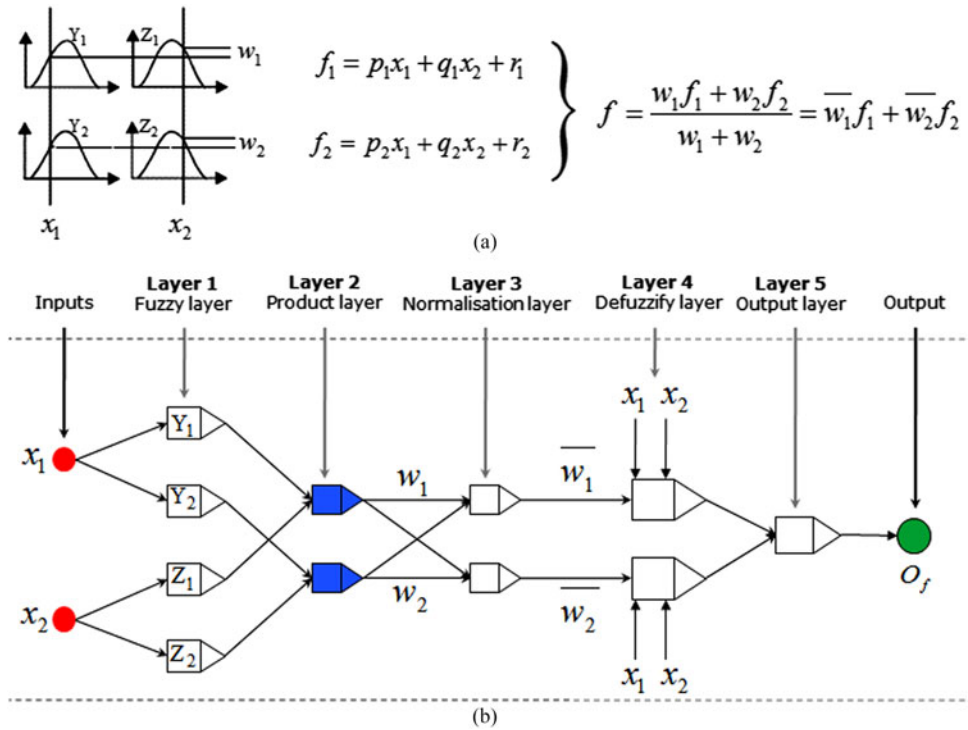

Fig. 8. (a) Type-3 fuzzy reasoning and, (b) equivalent ANFIS.

rules is:

$$\textbf{IF}\ \langle\text{Antecedent}\rangle\ \ \textbf{THEN}\ \langle\text{Consequent}\rangle$$

For illustration purpose, ANFIS architecture with two inputs (x_1, x_2) and one output (O_f) is shown in Fig. 8. The corresponding two fuzzy IF-THEN rules of Takagi and Sugeno's type [107] can be expressed as follows:

$$\begin{aligned}\text{Rule 1:}\ & \boldsymbol{IF}\ x_1\ is\ Y_1\ \text{and}\ x_2\ is\ Z_1,\\ & \boldsymbol{THEN}\ f_1 = p_1x_1 + q_1x_2 + r_1\\ \text{Rule 2:}\ & \boldsymbol{IF}\ x_1\ is\ Y_2\ \text{and}\ x_2\ is\ Z_2,\\ & \boldsymbol{THEN}\ f_2 = p_2x_1 + q_2x_2 + r_2.\end{aligned}$$

Fig. 8(a) shows the type-3 (two inputs and one output) fuzzy reasoning and Fig. 8(b) shows the corresponding ANFIS architecture.

The functionality and corresponding mathematical formulation of each layer is as follows [76]:

Layer 1: Fuzzy layer: Every node in this layer is fixed and adaptive and membership $(\mu(\circ))$ of each label $(Y_i,\ Z_i)$ is calculated by using (4) and (5)

$$U_i^1 = \mu_{Y_i}(x_1),\ \text{for}\ i = 1,\ 2 \tag{4}$$

$$U_i^1 = \mu_{Z_i}(x_2),\ \text{for}\ i = 1,\ 2 \tag{5}$$

where x_1 and x_2 are inputs and i is the node and Y_i and Z_i are the linguistic labels. $\mu_{Y_i}(x_1)$ is a membership function of Y_i which gives the degree of membership of x_1 to be part of Y_i (6).

TABLE II
FEATURES OF INPUT DATA, PERFORMANCE EVALUATION CRITERIA AND ARCHITECTURE AND PARAMETERS USED IN THE ANFIS AND ANN MODELS

Data	
Standardization:	0 Mean, 1 Std
Reduction:	PCA
Division:	70%—Training
	30%—Testing/validation
Performance evaluation criteria	
R^2:	Correlation coefficient
RMSE:	Root Mean Square Error
ANN	
Number of layers:	4
Neural net algorithm:	Feed-forward backpropagation
Number of neurons:	Input layer: 7, hidden layer neurons: 15, Output layer: 1
Initialization:	Weights: random, biases: random
Training algorithm:	Levenberg—Marquardt
Activation functions:	Log-sigmoid
ANFIS	
Number of layers:	5
Type:	Sugeno-type
Input membership function type:	Generalized bell-shaped membership function
Learning rule	Hybrid learning algorithm

The parameters $\{a_i,\ b_i,\ c_i\}$, referred to as premise parameters, determine the shape of the membership function

$$\mu_{Y_i}\ (x_1) = \frac{1}{1 + \left[\left(\frac{x_1 - c_i}{a_i}\right)^2\right]^{b_i}}\ . \tag{6}$$

Layer 2: Product layer: Every node in this layer is labeled N_2, the outcome of this layer is the product of incoming signals and is given by

$$w_i = \mu_{Y_i}\ (x_1)\ \times\ \mu_{Z_i}\,(x_2)\,,\ \textit{for}\ i\ =\ 1,\ 2 \tag{7}$$

where w_i is the output of layer 2.

Layer 3: Normalization layer: The third layer, labeled as N_3, is called the normalization layer

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}\ ,\ \text{for}\ i\ =\ 1,\ 2 \tag{8}$$

where the ratio of each weight to the total weight is calculated, i.e., *i*th node calculates the ratio of the *i*th rule's firing strength (8).

Layer 4: Defuzzify layer: Every node in this layer is adaptive and it is called the defuzzification layer (labeled as N_4 (9))

$$D_i^4 = \bar{w}_i\ \ f_i = \bar{w}_i\ \ (p_i x_1 + q_i x_2 + r_i) \tag{9}$$

where $\bar{w}_i$ is the output of layer 3, and the set of parameters $\{p_i, q_i, r_i\}$ is referred to as consequent parameters.

Layer 5: Output layer: All the incoming signals are summed in order to compute the overall output (10), i.e.,

$$O_f\ =\ \text{output}\ =\sum_{i=1}^{2}\bar{w}_i f_i\ =\frac{\sum_{i=1}^{2} w_i f_i}{\sum_{i=0}^{2} w_i}\ =\frac{w_1 f_1 + w_2 f_2}{w_1 + w_2}\ . \tag{10}$$

Table II shows the list of parameters used for in ANN and ANFIS models.

4) Performance Evaluation Criteria: Root mean square error (RMSE) and coefficient of determination (R^2) were used as bench marks for the performance assessment of all models (MLR, ANN and ANFIS). The mathematical formulations of these statistical error/performance criteria are as follows [108]:

TABLE III
MODELS DEVELOPMENT AND EVALUATION

Model development				
	Moorepark (R^2)		Grange (R^2)	
	Training	Testing	Training	Testing
MLR	0.31	0.21	0.39	0.29
ANN	0.65	0.54	0.71	0.54
ANFIS	0.88	0.78	0.80	0.74
Model evaluation on entire data set				
	Moorepark		Grange	
	R^2	$RMSE$	R^2	$RMSE$
MLR	0.29	25.08	0.38	24.02
ANN	0.63	18.05	0.59	20.43
ANFIS	0.85	11.07	0.76	15.35

$$R^2\ =\ \frac{\sum_{i\,=\,1}^{n}\left(O_i\ -\ \bar{O}_i\right)^2\ -\ \sum_{i\,=\,1}^{n}\left(O_i\ -\ P_i\right)^2}{\sum_{i\,=\,1}^{n}\left(O_i\ -\bar{O}_i\right)^2} \tag{11}$$

$$\text{RMSE}\ =\ \sqrt{\frac{1}{n}\sum_{i\,=\,1}^{n}\left(P_i\ -\ O_i\right)^2} \tag{12}$$

where n is the number of observations; P_i is predicted/estimated value; O_i is actual/observed value and $\overline{O_i}$ is the mean of observed values. The ideal performance of the underlying model gives a value of RMSE close to zero and the value of R^2 should be close to 1.

III. RESULTS AND DISCUSSION

Three different biomass estimation models, including both statistical (MLR) and machine learning (ANN and ANFIS)

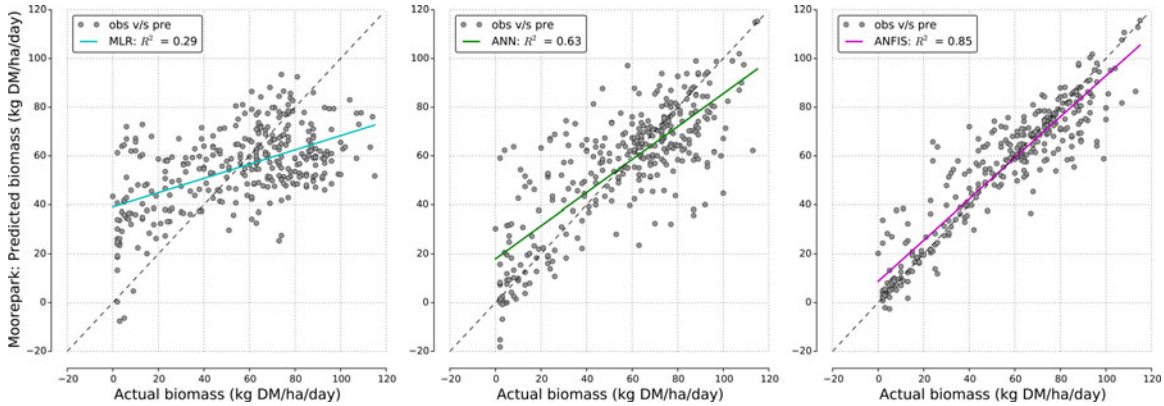

Fig. 9. Scatter plots for the accuracy comparison of MLR, ANN and ANFIS estimated grassland biomass versus *in situ* biomass for the Moorepark test site.

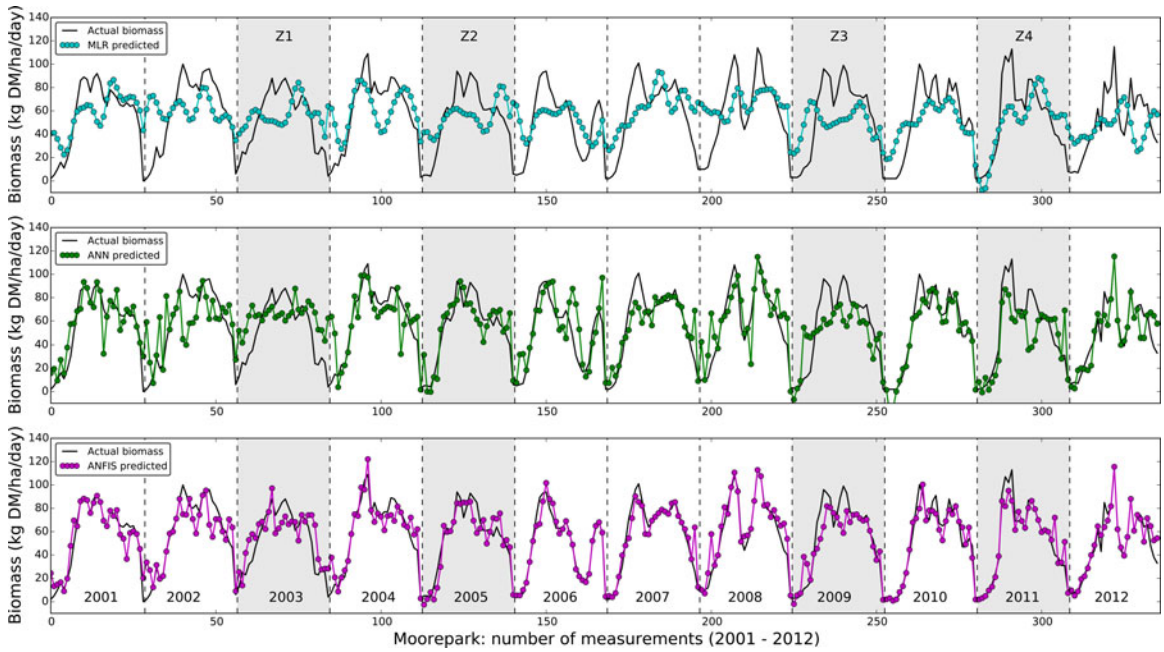

Fig. 10. Plots of observed and modeled time series by using MLR, ANN and ANFIS for Moorepark study site. The shaded regions Z1 (2003), Z2 (2005), Z3 (2009) and Z4 (2011) were selected for more detailed analysis (see Fig. 11).

approaches, were used to estimate intensively managed grassland biomass. These three models were used for both aforementioned study sites (see Section 2.1); and grassland biomass estimation models were developed where five VIs plus two spectral bands (RED, NIR) were used as input features. Firstly, the 12-year time series for Moorepark was used for model (MLR, ANN and ANFIS) development, with the dataset randomly divided into training (70%) and testing (30%) subsets (see Table II). The evaluation of the models was performed on the entire datasets (see Table III).

Fig. 9 shows the results for Moorepark study site. The first approach to estimating grassland biomass in this study was with the MLR, which has been demonstrated to be very robust when the relationship between datasets is linear. However, as shown in Fig. 9, the value of coefficient of determination for MLR is very low ($R^2 = 0.29$) and the value of RMSE ($\mathrm{RMSE} = 25.08$) is high compared to the ANN model ($R^2 = 0.63$, $\mathrm{RMSE} = 18.05$), suggesting a non-linear relationship between the variables. To date the use of machine learning algorithms for grassland biomass estimation is not very widespread. Two studies [73], [74] have compared the performance of the MLR and ANN, and in every case ANN outperformed the MLR; and the results generated by this study endorsed these findings.

The literature review suggests that the application of ANFIS is very powerful for estimation and prediction tasks [77]–[79], [81], but the use of ANFIS for spaceborne earth observation applications is only in its infancy [109], [110] and in these studies a high overall accuracy of ANFIS against ANN was reported. This outcome can also be seen here, as the ANFIS model gave better estimation results ($R^2 = 0.85$, $\mathrm{RMSE} = 11.07$) than both MLR and ANN (see Fig. 9).

Fig. 10, which gives an overview of the performance of the three models, shows that the ANN model was able to identify the

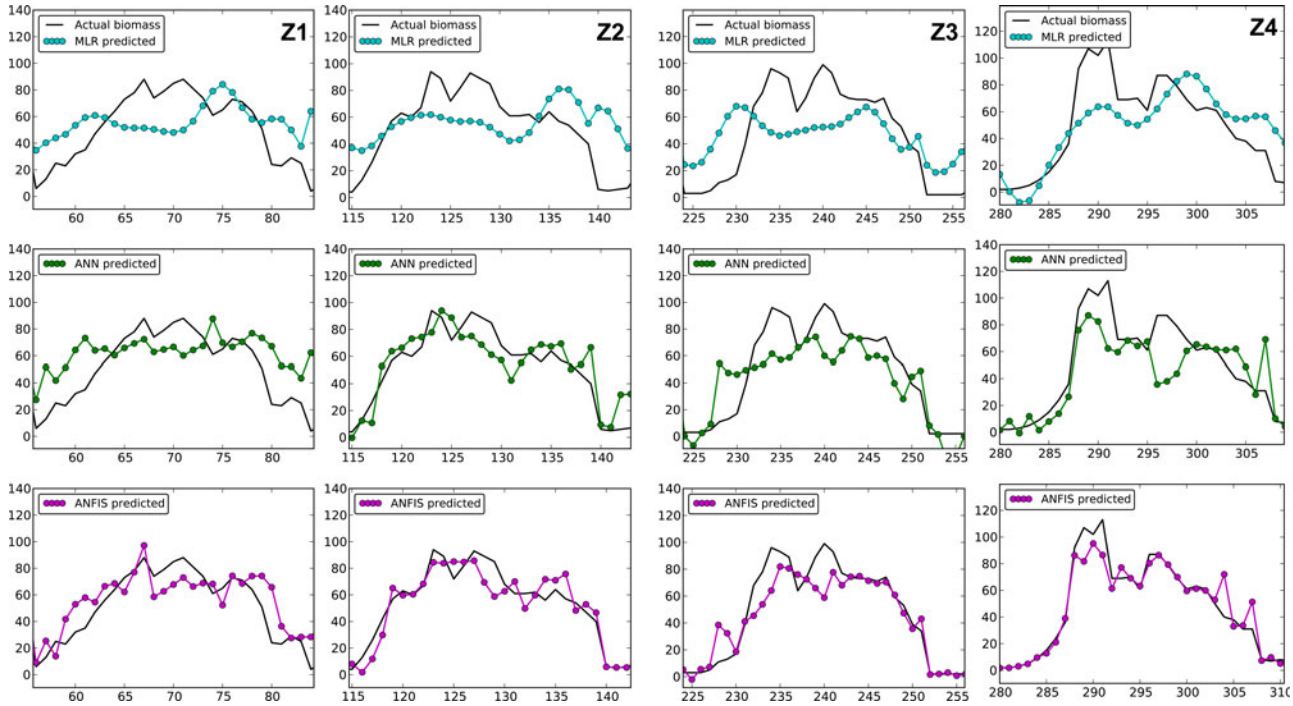

Fig. 11. Zoomed view of Fig. 10 highlighted parts (Z1, Z2, Z3 and Z4).

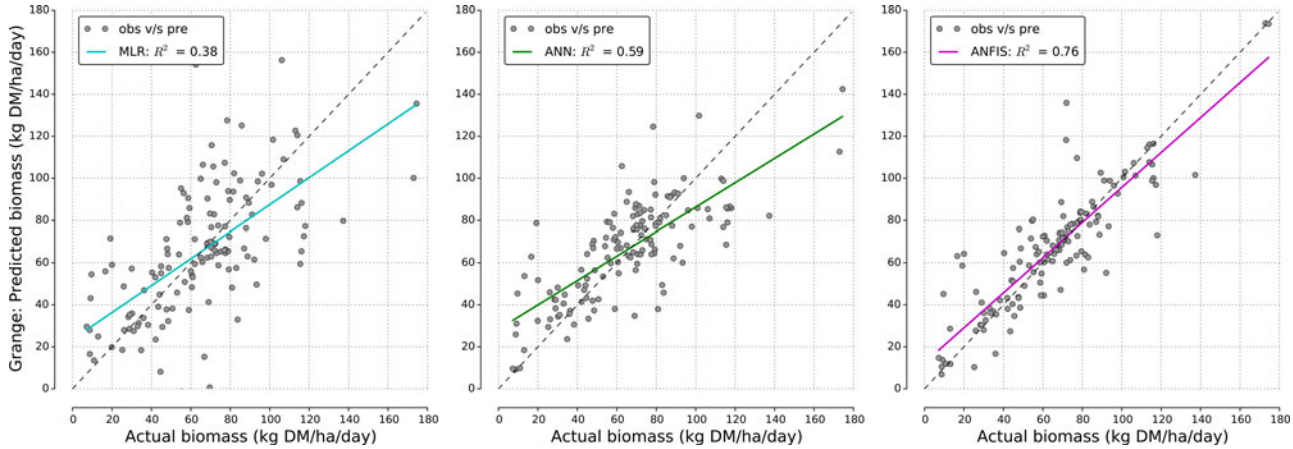

Fig. 12. Scatter plots for the accuracy comparison of MLR, ANN and ANFIS estimated grassland biomass verses *in situ* biomass for Grange test site.

start of the season more reliably than the MLR, this was further improved by the ANFIS model. The ANFIS also produced a closer seasonal curve fit with minimum residuals compared to the MLR and ANN, but there are still some spurious spikes that are not present in the field data and some features are not replicated.

Fig. 11 shows examples where peaks (higher biomass values) were not reached and underestimation is observed in four cases (Z1, Z2, Z3 and Z4). The reason for these anomalies is not yet clear, but one potential cause could be saturation of the satellite data, as in all four cases (Z1, Z2, Z3 and Z4) the overall general behavior of the estimated/modeled biomass curve is comparable for the three models (MLR, ANN, ANFIS). For example, in the case of Z1 all three models have overestimated the biomass during the start of the season, underestimated the higher biomass values (during summer) and again at the end of the season overestimation is observed (see Fig. 11 (Z1)). A similar trend is shown in Fig. 11 (Z2 and Z3) where higher biomass values are under estimated, while at the end of the season ANFIS has improved (reduced) the overestimation as compared to the MLR and ANN. In the case of Z4 the overall trend of the estimated pattern of MLR, ANN and ANFIS is comparable and ANN and ANFIS have identified the start of the season quite well, although ANFIS has minimized the estimation error, but still it has failed to reach the peak (higher biomass values) during the mid of the season, and the anomalies at the end of the season are similar for ANN and ANFIS. Fig. 11 shows that ANFIS has produced an improved estimation as compared to the MLR

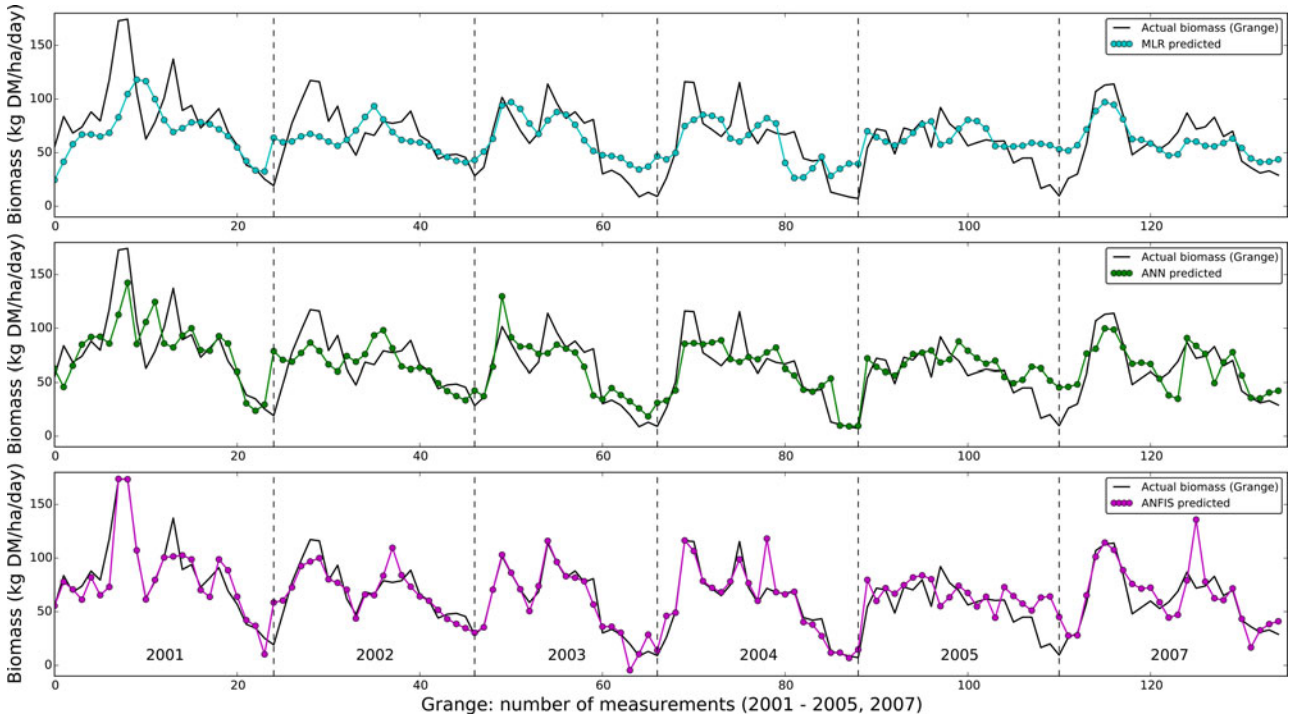

Fig. 13. Plots of observed and modeled time series for Grange study site.

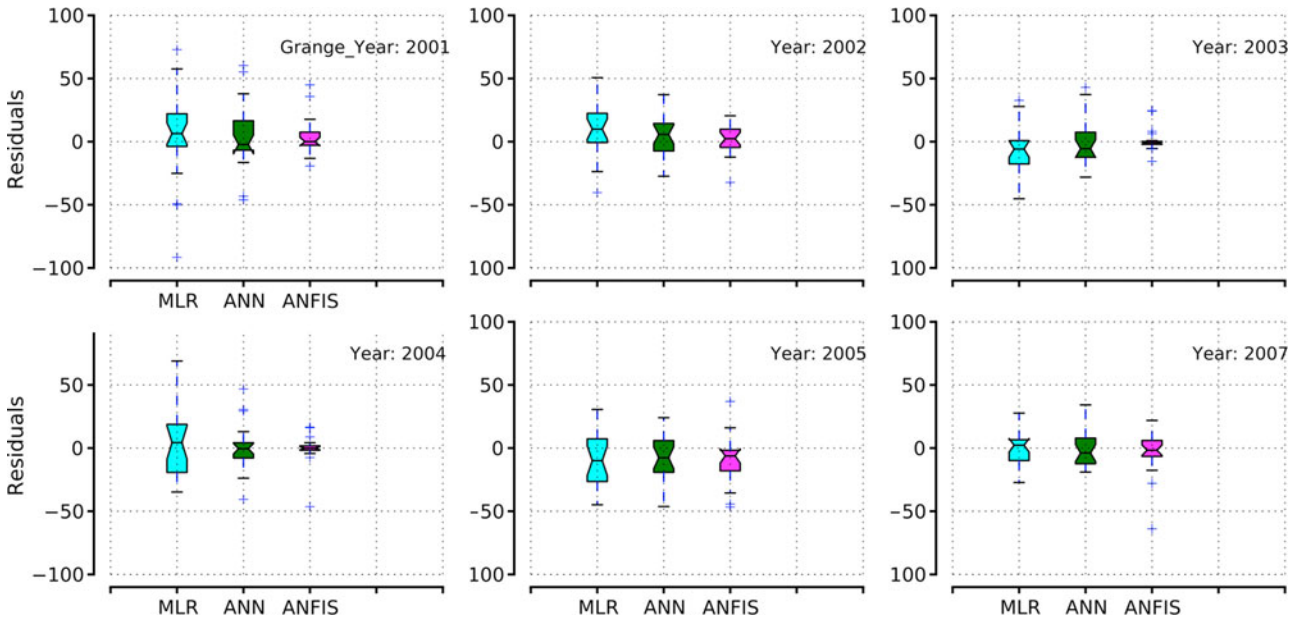

Fig. 14. Year wise residual plots for Grange study site.

and ANN but still in some cases it has underestimated the high biomass values.

To further explore the functional relationship between the input feature space and the *in situ* data, and also to explore the performance of the models, the same approach was applied to the Grange study site. Again the results show that the ANFIS model was the most accurate among the three models, with a higher value of coefficient of determination ($R^2 = 0.76$) and low RMSE ($\mathrm{RMSE} = 15.35$), followed by ANN ($R^2 = 0.59$, $\mathrm{RMSE} = 20.43$) and MLR ($R^2 = 0.38$, $\mathrm{RMSE} = 24.02$) (see Table III; Fig. 12).

Machine learning methods require large data sets in order to better understand the patterns hidden inside the data, which could be a reason for the higher accuracy achieved for the Moorepark study site. Another reason for the lower accuracy at the Grange test site could be that it is under more intense grazing practices, indicated by the biomass curves for Grange (see Fig. 4(B)) being more complex and variable, with

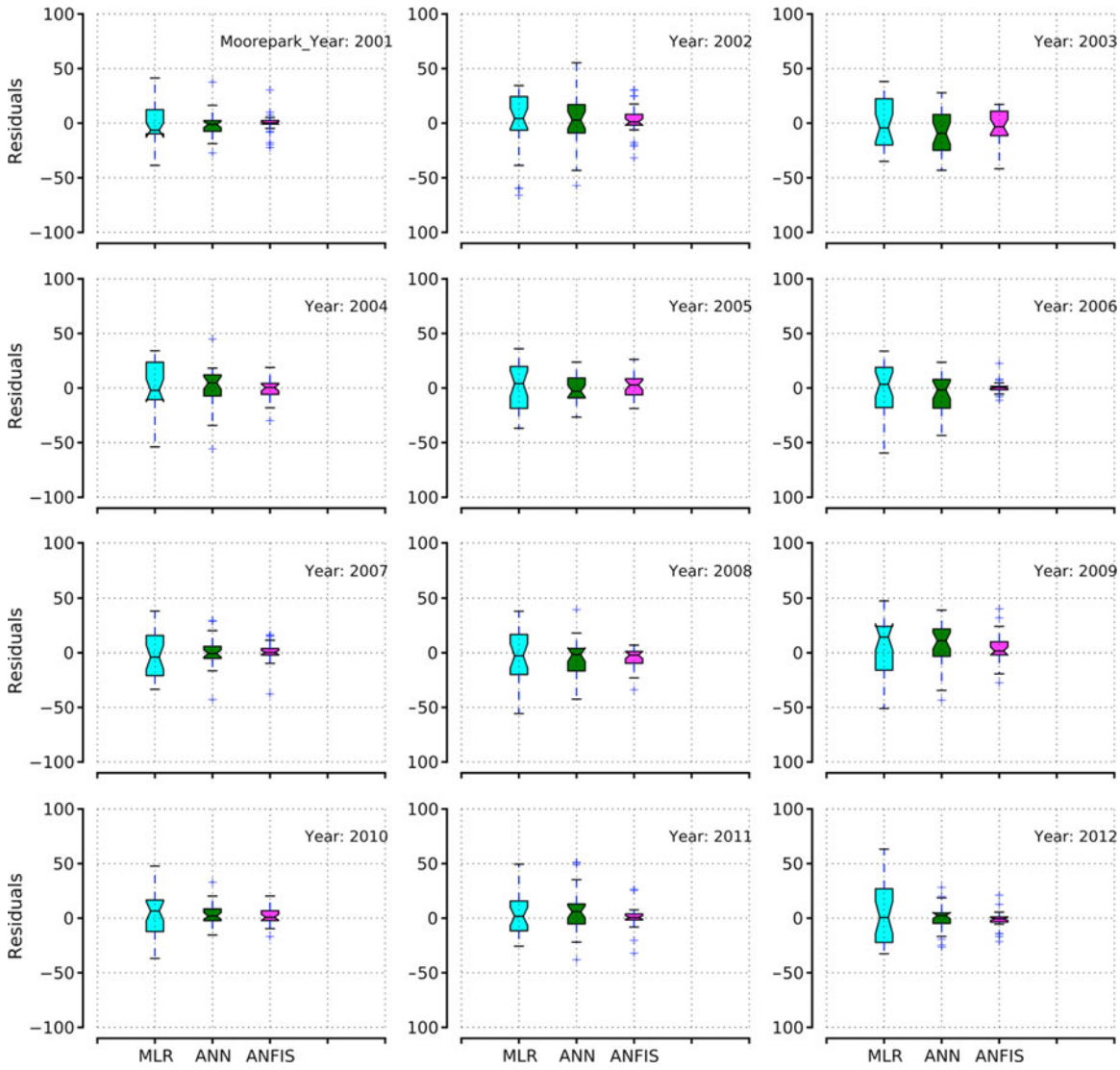

Fig. 15. Year wise residual plots for Moorepark study site.

considerable inter annual variation compared to the Moorepark *in situ* data (see Fig. 4(A)).

The issue of underestimation at higher biomass values was also observed at the Grange study site in some time periods, although better estimation at the start of the season can be seen in Fig. 13. In order to further analyze the effect of complexity on the models' performance, residual boxplots for each year were created for the Grange study site as shown in Fig. 14. The 2002 and 2005 residual boxplots show the highest variability and wide spread, especially for 2005 which is the most complex and nonlinear part of the Grange time series. By contrast, the Moorepark 12-year average and individual yearly biomass curves are quite consistent and similar in data range (min–max values; see Figs. 4 and 15).

A. Selection of Input Variables

As machine-learning models are generally data driven and require a large amount of data for better performance all five VIs along with two spectral bands (RED, NIR) were used as input variables to the PCA, and resulting principal component features were used as an input to the models. Various different combinations of inputs were tested and it was shown that the best accuracy was achieved by using all VIs as input variables for both statistical and ANN models.

B. Comparison of Three Models Performance on Both Study Sites

In terms of performance evaluation of the three models (see Fig. 16), it is evident that the distribution of interquartile range (IQR) of MLR ($\mathrm{IQR}_{\mathrm{Moorepark}} = 39.98$, $\mathrm{IQR}_{\mathrm{Grange}} = 26.72$) and ANN ($\mathrm{IQR}_{\mathrm{Moorepark}} = 17.99$, $\mathrm{IQR}_{\mathrm{Grange}} = 22.37$) residuals is quite large for both the test sites as compared to the distribution of IQRof ANFIS ($\mathrm{IQR}_{\mathrm{Moorepark}} = 7.78$, $\mathrm{IQR}_{\mathrm{Grange}} = 10.2$). For both the study sites ANFIS has less variability than the MLR and ANN, and the overall spread (min–max whisker range) of ANFIS for both the sites is small and

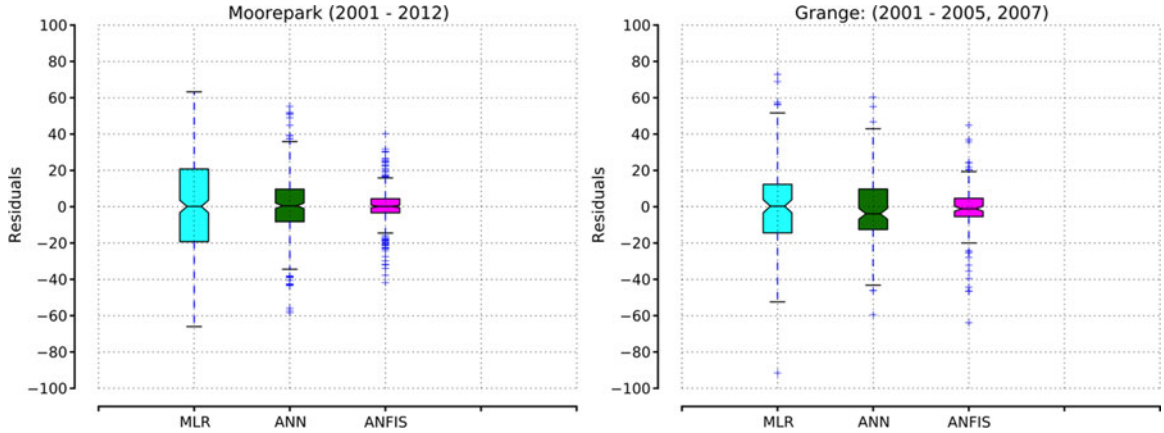

Fig. 16. Variations in residual for all three models (MLR, ANN, ANFIS) estimations. The boxplots show the spread, lower quartiles, medians and upper quartiles. The lines are drawn from the box (1.5 times the interquartile range from the nearer quartile).

symmetrical around zero. However the ANFIS scatter plots for both Moorepark (Fig. 9) and Grange (Fig. 12) are more unreliable for the higher biomass values. For example, for Moorepark underestimation is evident for the values > 60 kg DM/ha/day, similarly for Grange, large over and underestimation errors can be seen for the values > 100 kg DM/ha/day.

Studies show that whenever ANN and MLR are used for grassland biomass estimation ANN has always out-performed the traditional statistical approach. For example, Xie *et al.* [73] used a single Landsat ETM + image for above ground grassland biomass estimation and showed the superior performance of ANN ($R^2 = 0.817$) against MLR ($R^2 = 0.591$). Similar findings were reported by Yang *et al.* [74] where MODIS driven VIs from July–September 2005 were used to model the grass yield estimation, and ANN models were found to be more accurate ($R^2 = 0.56 - 0.71$) compared to the statistical models ($R^2 = 0.54 - 0.68$). The results of the current study have endorsed this trend of high performance for ANN against MLR. With respect to ANFIS and ANN, it has also been established in both non-remote sensing and remote sensing applications that the former generates more reliable results e.g., Rajesh *et al.* [110] indicated the higher classification performance of ANFIS (overall accuracy: 86.01%) compared to the ANN (overall accuracy: 83.62%).

C. Advantages and Limitations of Proposed Methodology

ANFIS not only integrates the strengths of ANN and fuzzy logic, but also overcomes some of the disadvantages of each applied separately and produces better results in terms of smoothness and adaptability. The presented framework of ANFIS modeling allows multiple inputs to produce a single output, however to achieve a higher level of accuracy a larger amount of data might be required to drive the model, with the model performance also dependent on the data quality and study design [111], [112].

IV. Conclusion

In this paper, the estimation capabilities of the ANFIS approach are compared against the ANN and more commonly used MLR modeling techniques. Although well established in other scientific fields (engineering, expert systems) the potential of ANFIS modeling in remote sensing is not yet fully explored, although as demonstrated by this research it is a technique that holds promise for future studies. Five MODIS derived VIs and two spectral bands (RED, NIR) along with the *in situ* measurements were used for model training and testing; and their performance was evaluated using R^2 and $RMSE$. For both the study sites, ANFIS ($R^2_{\text{Moorepark}} = 0.85$, $\text{RMSE}_{\text{Moorepark}} = 11.07$; $R^2_{\text{Grange}} = 0.76$, $\text{RMSE}_{\text{Grange}} = 15.35$) produced better estimations of biomass compared to the ANN ($R^2_{\text{Moorepark}} = 0.63$, $\text{RMSE}_{\text{Moorepark}} = 18.05$; $R^2_{\text{Grange}} = 0.59$, $\text{RMSE}_{\text{Grange}} = 20.43$) and MLR ($R^2_{\text{Moorepark}} = 0.29$, $\text{RMSE}_{\text{Moorepark}} = 25.08$; $R^2_{\text{Grange}} = 0.39$, $\text{RMSE}_{\text{Grange}} = 24.02$). However, there are some occasions when the model data underestimates the actual biomass peak (a common feature of VI driven biomass models); one potential reason for this underestimation could be the effect of saturation of the satellite signal or vegetation index value and further work is required to understand these anomalies. Nevertheless, these results show significant promise for the use of a hyper-temporal time series of satellite imagery as input to modeling for an effective tool for grassland monitoring and management.

With the launch of members of satellite families (ALOS-2, Radarsat-2, Sentinel, TerraSAR-X, TanDEM-X/L) the volume of data for such modeling studies will increase markedly, and concepts of big data are becoming more relevant in the remote sensing domain. As machine-learning models are considered to be data driven models, more data heralds higher accuracy. To date, grassland-modelling activities over 12 years have not been reported in the literature, but the scope for such long term studies will increase significantly over the coming years. In addition to demonstrating the potential of such long time series studies, this work has also highlighted the potential for complex modeling approaches such as ANFIS in the field of remote sensing. With the passage of time and availability of high quality spectral, spatial and temporal resolution data, these models will get further refined, more robust and applicable to other biophysical parameter retrieval tasks.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable suggestions and comments.

REFERENCES

[1] J. M. O. Scurlock, K. Johnson, and R. J. Olson, "Estimating net primary productivity from grassland biomass dynamics measurements," *Global Change Biol.*, vol. 8, no. 8, pp. 736–753, Aug. 2002.

[2] FAO, "The role of livestock in climate change," *Food Agriculture Org.*, 2014. [Online]. Available: http://www.fao.org/agriculture/lead/themes0/climate/en/.

[3] H. Lieth, Ed., *Patterns of Primary Production in the Biosphere*. London, U.K.: Hutchinson Ross Publishing Company, 1978.

[4] J. M. O. Scurlock and D. O. Hall, "The global carbon sink: a grassland perspective," *Global Change Biol.*, vol. 4, no. 2, pp. 229–233, 1998.

[5] J. M. Anderson, "The effects of climate change on decomposition processes in grassland and coniferous forests," *Ecol. Appl.*, vol. 1, no. 3, pp. 326–347, Aug. 1991.

[6] J. D. Derner and G. E. Schuman, "Carbon sequestration and rangelands: A synthesis of land management and precipitation effects," *J. Soil Water Conservation*, vol. 62, no. 2, pp. 77–85, Mar. 2007.

[7] C. Delgado, M. Rosegrant, H. Steinfeld, S. Ehui, and C. Courbois, "Livestock to 2020: The next food revolution," IFPRI, Working Paper, 1999.

[8] J. Kearney, "Food consumption trends and drivers," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 365, no. 1554, pp. 2793–2807, Sep. 2010.

[9] I. Ali, F. Cawkwell, E. Dwyer, B. Barrett, and S. Green, "Satellite remote sensing of grasslands: From observation to management—A review," *J. Plant Ecol.*, p. rtw005, Feb. 2016. DOI: 10.1093/jpe/rtw005.

[10] M. P. W. Hybu Cig Cymru, "Grassland management: http://hccmpw.org.uk/publications/farming_and_industry_development/grassland_management/," 2008.

[11] G. Newnham, I. Grant, D. Martin, and S. A. J. Anderson, "Improved Methods for Assessment and Prediction of Grassland Curing Satellite Based Curing Methods and Mapping - final report," 2010. Available online at: http://www.bushfirecrc.com/publications/citation/bf-2555.

[12] M. E. Castle, "A simple disc instrument for estimating herbage yield," *J. Brit. Grassland Soc.*, vol. 31, no. 1, pp. 37–40, 1976.

[13] J. Hakl, Z. Hrevušová, M. Hejcman, and P. Fuksa, "The use of a rising plate meter to evaluate lucerne (Medicago sativa L.) height as an important agronomic trait enabling yield estimation," *Grass Forage Sci.*, vol. 67, no. 4, pp. 589–596, Dec. 2012.

[14] M. Hejcman, L. Sochorová, V. Pavlů, J. Štrobach, M. Diepolder, and J. Schellberg, "The Steinach grassland experiment: Soil chemical properties, sward height and plant species composition in three cut alluvial meadow after decades-long fertilizer application," *Agriculture Ecosyst. Environ.*, vol. 184, pp. 76–87, Feb. 2014.

[15] E. S. Flynn, C. T. Dougherty, and O. Wendroth, "Assessment of pasture biomass with the normalized difference vegetation index from active ground-based sensors," *Agron. J.*, vol. 100, no. 1, p. 114–121, 2008.

[16] A. Psomas, M. Kneubühler, S. Huber, K. Itten, and N. E. Zimmermann, "Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of grassland habitats," *Int. J. Remote Sens.*, vol. 32, no. 24, pp. 9007–9031, 2011.

[17] B. Xu *et al.*, "MODIS-based remote sensing monitoring of grass production in China," *Int. J. Remote Sens.*, vol. 29, no. 17/18, pp. 5313–5327, 2008.

[18] G. Joseph, *Fundamentals of Remote Sensing*. Universities Press, Hyderabad India, 2005. Hyderabad India

[19] D. Stow, Y. Hamada, L. Coulter, and Z. Anguelova, "Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1051–1061, Mar. 2008.

[20] M. Lewis, V. Jooste, and A. A. de Gasparis, "Discrimination of arid vegetation with airborne multispectral scanner hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1471–1479, Jul. 2001.

[21] P. J. Curran and H. D. Williamson, "Estimating the green leaf area index of grassland with airborne multispectral scanner data," *Oikos*, vol. 49, no. 2, pp. 141–148, Jun. 1987.

[22] R. Darvishzadeh, C. Atzberger, A. Skidmore, and M. Schlerf, "Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models," *ISPRS J. Photogrammetry Remote Sens.*, vol. 66, no. 6, pp. 894–906, Nov. 2011.

[23] L. G. Olmanson, P. L. Brezonik, and M. E. Bauer, "Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: The Mississippi River and its tributaries in Minnesota," *Remote Sens. Environ.*, vol. 130, pp. 254–265, Mar. 2013.

[24] I. Ali, C. Schuster, M. Zebisch, M. Forster, B. Kleinschmit, and C. Notarnicola, "First results of monitoring nature conservation sites in alpine region by using very high resolution (VHR) X-band SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 5, pp. 2265–2274, Oct. 2013.

[25] C. Schuster, I. Ali, P. Lohmann, A. Frick, M. Förster, and B. Kleinschmit, "Towards detecting swath events in TerraSAR-X time series to establish NATURA 2000 grassland habitat swath management as monitoring parameter," *Remote Sens.*, vol. 3, no. 7, pp. 1308–1322, Jun. 2011.

[26] M. Claverie *et al.*, "Maize and sunflower biomass estimation in southwest France using high spatial and temporal resolution remote sensing data," *Remote Sens. Environ.*, vol. 124, pp. 844–857, Sep. 2012.

[27] M. E. Holzman, R. Rivas, and M. C. Piccolo, "Estimating soil moisture and the relationship with crop yield using surface temperature and vegetation index," *Int. J. Appl. Earth Observ. Geoinform.*, vol. 28, pp. 181–192, May 2014.

[28] D. K. Sari, I. H. Ismullah, W. N. Sulasdi, and A. B. Harto, "Estimation of water consumption of lowland rice in tropical area based on heterogeneous cropping calendar using remote sensing technology," *Proc. Environ. Sci.*, vol. 17, pp. 298–307, 2013.

[29] T. G. Van Niel and T. R. McVicar, "Determining temporal windows for crop discrimination with remote sensing: A case study in south-eastern Australia," *Comput. Electron. Agriculture*, vol. 45, no. 1–3, pp. 91–108, Dec. 2004.

[30] J. Rhee, J. Im, and G. J. Carbone, "Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data," *Remote Sens. Environ.*, vol. 114, no. 12, pp. 2875–2887, Dec. 2010.

[31] G. Donald, S. Gherardi, A. Edirisinghe, S. Gittins, D. Henry, and G. Mata, "Pasture growth rate for individual paddocks can be accurately predicted real-time from MODIS imagery, climate and soil data," in *Proc. Aust. Soc. Animal Prod.*, Jul. 2010, vol. 50, pp. 611–615.

[32] M. J. Hill, G. E. Donald, M. W. Hyder, and R. C. G. Smith, "Estimation of pasture growth rate in the south west of Western Australia from AVHRR NDVI and climate data," *Remote Sens. Environ.*, vol. 93, no. 4, pp. 528–545, Dec. 2004.

[33] Y. P. Dang, M. J. Pringle, M. Schmidt, R. C. Dalal, and A. Apan, "Identifying the spatial variability of soil constraints using multi-year remote sensing," *Field Crops Res.*, vol. 123, no. 3, pp. 248–258, Sep. 2011.

[34] Y. Gu and B. K. Wylie, "detecting ecosystem performance anomalies for land management in the upper colorado river basin using satellite observations, climate data, and ecosystem models," *Remote Sens.*, vol. 2, no. 8, pp. 1880–1891, Jul. 2010.

[35] S. Kalluri, P. Gilruth, and R. Bergman, "The potential of remote sensing data for decision makers at the state, local and tribal level: Experiences from NASA's Synergy program," *Environ. Sci. Policy*, vol. 6, no. 6, pp. 487–500, Dec. 2003.

[36] M. Qingyuan, Z. Chao, C. Zhenghua, and Y. Zhen, "Management decision-making support system of percision agriculture based on CNCS," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 819–822.

[37] S. K. Seelan, S. Laguette, G. M. Casady, and G. A. Seielstad, "Remote sensing applications for precision agriculture: A learning community approach," *Remote Sens. Environ.*, vol. 88, no. 1/2, pp. 157–169, Nov. 2003.

[38] M. Turker and E. H. Kok, "Field-based sub-boundary extraction from remote sensing imagery using perceptual grouping," *ISPRS J. Photogrammetry Remote Sens.*, vol. 79, pp. 106–121, May 2013.

[39] I. Ali, F. Greifeneder, J. Stamenkovic, M. Neumann, and C. Notarnicola, "Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data," *Remote Sens.*, vol. 7, no. 12, pp. 16398–16421, Dec. 2015.

[40] G. . Anderson, J. . Hanson, and R. . Haas, "Evaluating landsat thematic mapper derived vegetation indices for estimating above-ground biomass on semiarid rangelands," *Remote Sens. Environ.*, vol. 45, no. 2, pp. 165–175, Aug. 1993.

[41] B. K. Wylie, J. A. Harrington, S. D. Prince, and I. Denda, "Satellite and ground-based pasture production assessment in Niger: 1986–1988," *Int. J. Remote Sens.*, vol. 12, no. 6, pp. 1281–1300, 1991.

[42] P. Dusseux, L. Hubert-Moy, T. Corpetti, and F. Vertès, "Evaluation of SPOT imagery for the estimation of grassland biomass," *Int. J. Appl. Earth Obs. Geoinform.*, vol. 38, pp. 72–77, Jun. 2015.

[43] N. Huang, J.-S. He, and Z. Niu, "Estimating the spatial pattern of soil respiration in Tibetan alpine grasslands using Landsat TM images and MODIS data," *Ecological Indicator*, vol. 26, pp. 117–125, Mar. 2013.

[44] S. Itano and H. Tomimatsu, "Reflectance spectra for monitoring green herbage mass in Zoysia-dominated pastures," *Grassland Sci.*, vol. 57, no. 1, pp. 9–17, Mar. 2011.

[45] L. Jianlong, L. Tiangang, and C. Quangong, "Estimating grassland yields using remote sensing and GIS technologies in China," *New Zealand J. Agricultural Res.*, vol. 41, no. 1, pp. 31–38, Jan. 1998.

[46] Y. Jin *et al.*, "Remote sensing-based biomass estimation and its spatio-temporal variations in temperate grassland, Northern China," *Remote Sens.*, vol. 6, no. 2, pp. 1496–1513, Feb. 2014.

[47] M. Shen *et al.*, "Estimation of aboveground biomass using in situ hyperspectral measurements in five major grassland ecosystems on the Tibetan Plateau," *J. Plant Ecology*, vol. 1, no. 4, pp. 247–257, Dec. 2008.

[48] W. Song *et al.*, "A remote sensing based forage biomass yield inversion model of alpine-cold meadow during grass-withering period in Sanjiangyuan area," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 17, no. 1, Mar. 2014, Art no. 012042.

[49] B. Xu, X. C. Yang, W. G. Tao, Z. H. Qin, H. Q. Liu, J. M. Miao, and Y. Y. Bi, "MODIS-based remote sensing monitoring of grass production in China," *Int. J. Remote Sens.*, vol. 29, no. 17/18, pp. 5313–5327, Sep. 2008.

[50] F. Zhao *et al.*, "Remote sensing estimates of grassland aboveground biomass based on MODIS net primary productivity (NPP): A case study in the Xilingol grassland of Northern China," *Remote Sens.*, vol. 6, no. 6, pp. 5368–5386, Jun. 2014.

[51] M. Boschetti, S. Bocchi, and P. A. Brivio, "Assessment of pasture production in the Italian Alps using spectrometric and remote sensing information," *Agriculture Ecosyst. Environ.*, vol. 118, no. 1–4, pp. 267–272, Jan. 2007.

[52] S. Ullah, Y. Si, M. Schlerf, A. K. Skidmore, M. Shafique, and I. A. Iqbal, "Estimation of grassland biomass and nitrogen using MERIS data," *Int. J. Appl. Earth Obs. Geoinform.*, vol. 19, pp. 196–204, Oct. 2012.

[53] Y. Zha, J. Gao, S. Ni, Y. Liu, J. Jiang, and Y. Wei, "A spectral reflectance-based approach to quantification of grassland cover from Landsat TM imagery," *Remote Sens. Environ.*, vol. 87, no. 2/3, pp. 371–375, Oct. 2003.

[54] D. Bella *et al.*, "Remote sensing capabilities to estimate pasture production in France," *Int. J. Remote Sens.*, vol. 25, no. 23, pp. 5359–5372, 2004.

[55] J. B. Bradford, J. A. Hicke, and W. K. Lauenroth, "The relative importance of light-use efficiency modifications from environmental conditions and cultivation for estimation of large-scale net primary productivity," *Remote Sens. Environ.*, vol. 96, no. 2, pp. 246–255, May 2005.

[56] L.-C. Han, "A method of modifying error for non-synchronicity of grass yield remote sensing estimation and measurement," *Int. J. Remote Sens.*, vol. 22, no. 17, pp. 3363–3372, 2001.

[57] V. Loris and G. Damiano, "Mapping the green herbage ratio of grasslands using both aerial and satellite-derived spectral reflectance," *Agriculture Ecosyst. Environ.*, vol. 115, no. 1–4, pp. 141–149, Jul. 2006.

[58] L. Vescovo and D. Gianelle, "Using the MIR bands in vegetation indices for the estimation of grassland biophysical parameters from satellite remote sensing in the Alps region of Trentino (Italy)," *Adv. Space Res.*, vol. 41, no. 11, pp. 1764–1772, 2008.

[59] B. Wylie, D. Meyer, L. Tieszen, and S. Mannel, "Satellite mapping of surface biophysical parameters at the biome scale over the North American grasslands: A case study," *Remote Sens. Environ.*, vol. 79, no. 2/3, pp. 266–278, Feb. 2002.

[60] B. Xu, X. Yang, W. Tao, Z. Qin, H. Liu, and J. Miao, "Remote sensing monitoring upon the grass production in China," *Acta Ecologica Sin.*, vol. 27, no. 2, pp. 405–413, Feb. 2007.

[61] B. A. M. Bouman, A. H. C. M. Schapendonk, and W. Stol, *Description of the Growth Model LINGRA as Implemented in CGMS*. Wageningen, The Netherlands: AB-DLO, 1996.

[62] A. H. C.M. Schapendonk, W. Stol, D. W.G. van Kraalingen, and B. A.M. Bouman, "LINGRA, a sink/source model to simulate grassland productivity in Europe," *Eur. J. Agronomy*, vol. 9, no. 2/3, pp. 87–100, Nov. 1998.

[63] F. Maselli, G. Argenti, M. Chiesi, L. Angeli, and D. Papale, "Simulation of grassland productivity by the combination of ground and satellite data," *Agriculture Ecosyst. Environ.*, vol. 165, pp. 163–172, Jan. 2013.

[64] A. de Wit, G. Duveiller, and P. Defourny, "Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations," *Agricultural Forest Meteorology*, vol. 164, pp. 39–52, Oct. 2012.

[65] A. V. M. Ines, N. N. Das, J. W. Hansen, and E. G. Njoku, "Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction," *Remote Sens. Environ.*, vol. 138, pp. 149–164, Nov. 2013.

[66] R. E. E. Jongschaap, *Integrating Crop Growth Simulation and Remote Sensing to Improve Resource Use Efficiency in Farming Systems*. Wageningen, The Netherlands: Wageningen Universiteit, 2006.

[67] B. Ji, Y. Sun, S. Yang, and J. Wan, "Artificial neural networks for rice yield prediction in mountainous regions," *J. Agricultural Sci.*, vol. 145, no. 03, pp. 249–261, 2007.

[68] S. S. Panda, D. P. Ames, and S. Panigrahi, "Application of vegetation indices for agricultural crop yield prediction using neural network techniques," *Remote Sens.*, vol. 2, no. 3, pp. 673–696, Mar. 2010.

[69] C. Z. Serele, Q. H. J. Gwyn, J. B. Boisvert, E. Pattey, N. McLaughlin, and G. Daoust, "Corn yield prediction with artificial neural network trained using airborne remote sensing and topographic data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2000, vol.1, pp. 384–386.

[70] Y. Uno *et al.*, "Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data," *Comput. Electron. Agriculture*, vol. 47, no. 2, pp. 149–161, May 2005.

[71] M. E. J. Cutler, D. S. Boyd, G. M. Foody, and A. Vetrivel, "Estimating tropical forest biomass with a combination of SAR image texture and Landsat TM data: An assessment of predictions between regions," *ISPRS J. Photogrammetry Remote Sens.*, vol. 70, pp. 66–77, Jun. 2012.

[72] I. Ali, F. Cawkwell, S. Green, and N. Dwyer, "Application of statistical and machine learning models for grassland yield estimation based on a hypertemporal satellite remote sensing time series," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2014, pp. 5060–5063.

[73] Y. Xie, Z. Sha, M. Yu, Y. Bai, and L. Zhang, "A comparison of two models with Landsat data for estimating above ground grassland biomass in Inner Mongolia, China," *Ecological Modelling*, vol. 220, no. 15, pp. 1810–1818, Aug. 2009.

[74] X. Yang, B. Xu, J. Yunxiang, L. Jinya, and X. Zhu, "On grass yield remote sensing estimation models of China's northern farming-pastoral ecotone," in *Advances in Computational Environment Science*, G. Lee, Ed., Berlin, Germany: Springer, 2012, pp. 281–291.

[75] J. Clevers, G. Van Der Heijden, S. Verzakov, and M. Schaepman, "Estimating grassland biomass using SVM band shaving of hyperspectral data," *Photogrammetric Eng. Remote Sens.*, vol. 73, no. 10, p. 1141–1148, 2007.

[76] J.-S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans. Syst. Man Cybern.*, vol. 23, no. 3, pp. 665–685, May 1993.

[77] M. K. Goyal, B. Bharti, J. Quilty, J. Adamowski, and A. Pandey, "Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, fuzzy logic, and ANFIS," *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5267–5276, Sep. 2014.

[78] M. Hosoz, H. M. Ertunc, M. Karabektas, and G. Ergen, "ANFIS modelling of the performance and emissions of a diesel engine using diesel fuel and biodiesel blends," *Appl. Thermal Eng.*, vol. 60, no. 1/2, pp. 24–32, Oct. 2013.

[79] H. M. Jiang, C. K. Kwong, W. H. Ip, and T. C. Wong, "Modeling customer satisfaction for new product development using a PSO-based ANFIS approach," *Appl. Soft Comput.*, vol. 12, no. 2, pp. 726–734, Feb. 2012.

[80] R. K. Kharb, S. L. Shimi, S. Chatterji, and M. F. Ansari, "Modeling of solar PV module and maximum power point tracking using ANFIS," *Renewable Sustain. Energy Rev.*, vol. 33, pp. 602–612, May 2014.

[81] A. Mellit and S. A. Kalogirou, "ANFIS-based modelling for photovoltaic power supply system: A case study," *Renewable Energy*, vol. 36, no. 1, pp. 250–258, Jan. 2011.

[82] A. Nazari, A. A. Milani, and G. Khalaj, "Modeling ductile to brittle transition temperature of functionally graded steels by ANFIS," *Appl. Math. Modelling*, vol. 36, no. 8, pp. 3903–3915, Aug. 2012.

[83] S. Sabzi, P. Javadikia, H. Rabani, and A. Adelkhani, "Mass modeling of Bam orange with ANFIS and SPSS methods for using in machine vision," *Measurement*, vol. 46, no. 9, pp. 3333–3341, Nov. 2013.

[84] A. Sadrmomtazi, J. Sobhani, and M. A. Mirgozar, "Modeling compressive strength of EPS lightweight concrete using regression, neural network and ANFIS," *Construction Building Mater.*, vol. 42, pp. 205–216, May 2013.

[85] T. R. Kiran and S. P. S. Rajput, "An effectiveness model for an indirect evaporative cooling (IEC) system: Comparison of artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS) and

fuzzy inference system (FIS) approach," *Appl. Soft Comput.*, vol. 11, no. 4, pp. 3525–3533, Jun. 2011.
[86] D. S. Ali Haghighat Mesbahi, "Performance prediction of a specific wear rate in epoxy nanocomposites with various composition content of polytetrafluoroethylen (PTFE), graphite,short carbon fibers (CF) and nano-TiO2 using adaptive neuro-fuzzy inference," *Composites Part B*, 2011.
[87] H. Esen and M. Inalli, "ANN and ANFIS models for performance evaluation of a vertical ground source heat pump system," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8134–8147, Dec. 2010.
[88] M. Iphar, "ANN and ANFIS performance prediction models for hydraulic impact hammers," *Tunnelling Underground Space Technol.*, vol. 27, no. 1, pp. 23–29, Jan. 2012.
[89] A. Karami and S. Afiuni-Zadeh, "Sizing of rock fragmentation modeling due to bench blasting using adaptive neuro-fuzzy inference system (ANFIS)," *Int. J. Mining Sci. Technol.*, vol. 23, no. 6, pp. 809–813, Nov. 2013.
[90] G. Karimi, S. B. Sedaghat, and R. Banitalebi, "Designing and modeling of ultra low voltage and ultra low power LNA using ANN and ANFIS for Bluetooth applications," *Neurocomputing*, vol. 120, pp. 504–508, Nov. 2013.
[91] B. Rahmanian, M. Pakizeh, S. A. A. Mansoori, M. Esfandyari, D. Jafari, H. Maddah, and A. Maskooki, "Prediction of MEUF process performance using artificial neural networks and ANFIS approaches," *J. Taiwan Inst. Chem. Eng.*, vol. 43, no. 4, pp. 558–565, Jul. 2012.
[92] M. Talebizadeh and A. Moridnejad, "Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ANN and ANFIS models," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4126–4135, Apr. 2011.
[93] Y. Varol, E. Avci, A. Koca, and H. F. Oztop, "Prediction of flow fields and temperature distributions due to natural convection in a triangular enclosure using adaptive-network-based fuzzy inference system (ANFIS) and artificial neural network (ANN)," *Int. Commun. Heat Mass Transfer*, vol. 34, no. 7, pp. 887–896, Aug. 2007.
[94] Z. Yuan, L.-N. Wang, and X. Ji, "Prediction of concrete compressive strength: Research on hybrid models genetic based algorithms and ANFIS," *Adv. Eng. Softw.*, vol. 67, pp. 156–163, Jan. 2014.
[95] J. Sobhani, M. Najimi, A. R. Pourkhorshidi, and T. Parhizkar, "Prediction of the compressive strength of no-slump concrete: A comparative study of regression, neural network and ANFIS models," *Construction Building Mater.*, vol. 24, no. 5, pp. 709–718, May 2010.
[96] I. Yilmaz and O. Kaynar, "Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5958–5966, May 2011.
[97] N. Guindin-Garcia, A. A. Gitelson, T. J. Arkebauer, J. Shanahan, and A. Weiss, "An evaluation of MODIS 8- and 16-day composite products for monitoring maize green leaf area index," *Agricultural Forest Meteorology*, vol. 161, pp. 15–25, Aug. 2012.
[98] J. Rouse, R. Haas, J. Schell, and D. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," in *Proc. 3rd Earth Resources Technol. Satellite Symp.*, 1973, vol. 1, pp. 309–317.
[99] Z. Jiang, A. R. Huete, K. Didan, and T. Miura, "Development of a two-band enhanced vegetation index without a blue band," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3833–3845, Oct. 2008.
[100] A. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, no. 3, pp. 295–309, Aug. 1988.
[101] J. Qi, A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian, "A modified soil adjusted vegetation index," *Remote Sens. Environ.*, vol. 48, no. 2, pp. 119–126, May 1994.
[102] G. Rondeaux, M. Steven, and F. Baret, "Optimization of soil-adjusted vegetation indices," *Remote Sens. Environ.*, vol. 55, no. 2, pp. 95–107, Feb. 1996.
[103] G. Civelekoglu, N. O. Yigit, E. Diamadopoulos, and M. Kitis, "Prediction of bromate formation using multi-linear regression and artificial neural networks," *Ozone Sci. Eng.*, vol. 29, no. 5, pp. 353–362, 2007.
[104] B. Yegnanarayana, *Artificial Neural Networks*. New Delhi, India: PHI Learning Pvt. Ltd., 2009.
[105] I. A. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, Dec. 2000.
[106] M. H. Hassoun, *Fundamentals of Artificial Neural Networks*. Cambridge, MA, USA: MIT Press, 1995.
[107] T. Takagi and M. Sugeno, "Deviation of Fuzzy Control Rules from Human Operator's Control Actions," in *Proc. IFAC Symp. Fuzzy Inform. Knowl. Representation Decision Anal.*, 1983, pp. 55–60.
[108] J. Benesty, M. M. Sondhi, and Y. Huang, *Springer Handbook of Speech Processing*. Berlin, Germany: Springer, 2008.
[109] S. Mojeddifar, H. Ranjbar, and H. Nezamabadipour, "Adaptive neuro-fuzzy inference system application for hydrothermal alteration mapping using ASTER data," *J. Mining Environ.*, vol. 4, no. 2, pp. 83–96, Oct. 2013.
[110] S. Rajesh, S. Arivazhagan, K. P. Moses, and R. Abisekaraj, "ANFIS based land cover/land use mapping of LISS IV imagery using optimized wavelet packet features," *J. Indian Soc. Remote Sens.*, vol. 42, no. 2, pp. 267–277, Jun. 2014.
[111] E. Giovanis, "Study of discrete choice models and adaptive neuro-fuzzy inference system in the prediction of economic crisis periods in USA," *Social Science Research Network*, Rochester, NY: SSRN Scholarly, Mar. 2012, Art no. 2014737.
[112] B. Issac and N. Israr, *Case Studies in Intelligent Computing: Achievements and Trends*. Boca Raton, FL, USA: CRC Press, 2014.

Iftikhar Ali received the B.S. degree in mathematics from Bahauddin Zakariya University Multan Pakistan, in 2008, and the M.Sc. degree in geodesy and geoinformation science (major in remote sensing) from the Technical University of Berlin, Berlin, Germany, in 2011. He received the Ph.D. degree in remote sensing from the University College Cork, Cork, Ireland, in March 2016.

From June 2011 to June 2012 he worked as a Remote Sensing Researcher at the Institute for Applied Remote Sensing, EURAC, Italy and the Technical University of Berlin, Germany. He also worked as a JPL Visiting Student Researcher at JPL, NASA, CA, USA. He is currently a Postdoc Researcher at the Remote Sensing Research Group, Vienna University of Technology, Wien, Austria. His research interests include remote sensing based biophysical parameters retrieval and application of machine learning algorithms for remote sensing based parameters retrieval.

Fiona Cawkwell received the B.Sc. degree in geography from the University of Edinburgh, Edinburgh, U.K., in 1997, and the M.Sc. degree in remote sensing from the University of London, London, U.K., in 1998 and the Ph.D. degree in remote sensing from the University of Bristol, Bristol, U.K, in 2003.

Since 2006, she has been working at the Earth Observation Research Group, Geography Department, University College Cork, Cork, Ireland, which specializes in land cover mapping and monitoring using optical and microwave satellite images.

Edward Dwyer received the Graduate degree in electronic engineering from the Trinity College Dublin, Dublin, Ireland. He received the M.Sc. and the Ph.D. degrees in satellite remote sensing applications.

He is currently working as the Executive-Director of the EurOcean Foundation, European Centre for Information on Marine Science and Technology, Lisbon, Portugal. From 2002 to 2014, he worked as a Team Leader of the Applied Remote Sensing and GIS group at the Coastal and Marine Research Centre, University College Cork, Cork, Ireland and from 2010 as the Deputy Director of the Coastal and Marine Research Centre. His thematic experience covers the use of both optical and SAR data for terrestrial and ocean applications.

Stuart Green received the degree in geophysics (hons.) from the University of Liverpool, Liverpool, U.K., in 1992 and the M.Sc. degree in remote sensing from Dundee University, Dundee, U.K, in 1993. He is currently working toward the Ph.D. degree in remote sensing of grass and herd management from the University College Cork, Cork, Ireland, in 2016.

After working in University College Dublin, Dublin, Ireland, on a project on airborne measurements of eutrophic status in Irish lakes, he joined as a Remote Sensing Specialist in Teagasc, the Irish Agricultural Research Body, Oak Park, Carlow, Ireland. His research interests include the remote sensing of pasture biomass, farm habitats and farm drainage.