PREDICTION OF RICE BIOMASS USING MACHINE LEARNING ALGORITHMS

By

DERRAZ RADHWANE

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

December 2022

FP 2022 70

DEDICATION

To my lovely parents, who always kept praying for me day and night to achieve my goal. To my family members and all my friends who supported me all these years.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

PREDICTION OF RICE BIOMASS USING MACHINE LEARNING ALGORITHMS

By

DERRAZ RADHWANE

December 2022

Chair : Farrah Melissa Muharam, PhD
Faculty : Agriculture

Conventional rice sampling methods are effective. However, they are destructive, laborious, time-consuming, impractical for large fields, and subject to human error. Unmanned aerial vehicles (UAVs) may address these issues. Machine learning algorithms (MLs) can predict rice biomass from UAV-based vegetation indices (VIs). Nevertheless, VIs are highly collinear, noisy, and their large dataset collection is expensive. These issues affect the MLs' model performance, stability (under/overfitting), variance, and confidence. This study aims to: (i) compare the base and ensemble MLs' model performance, variance, stability, and confidence for predicting rice biomass using collinear (multicollinearity context (MCC)) and non-collinear (non-multicollinearity context (NMCC)) VIs; (ii) compare the rice above ground biomass (TAGB) predictability from noised and Kalman filter' denoised VIs using histogram gradient boosting regressor (HGBR); (iii) develop a trigonometric-Euclidean-smoother interpolator (TESI), including linear (LN-TESI), cubic (C-TESI), quadratic (Q-TESI), and logarithmic (L-TESI) interpolators, for continuous time-series and non-time-series VIs data augmentation, and compare them to the tabular variational autoencoder (TVAE) and the conditional tabular generative adversarial network (CTGAN) for preventing DNN's under/overfitting. A split-plot randomised complete block design (RCBD) experiment was conducted in a rice granary at Terengganu, Malaysia, with 120 quadrants. Each quadrant provides five rice biomass traits during the tillering, booting, and milking stages. A MicaSense Red-Edge multispectral camera mounted on a DJI quadcopter drone was used to acquire the blue, green, red, red-edge, and NIR bands to extract the VIs values corresponding to each quadrant. Besides the biomass dataset, the non-time-series fertiliser dataset and the time-series oil palm and rice datasets were also collected to validate the TESI, TVAE, and CTGAN results. For the first objective, the MLs model performance and stability were better in MCC than in NMCC for predicting all rice biomass traits. The ensemble MLs outperformed the base MLs for predicting all rice biomass traits in MCC and NMCC. All base and ensemble MLs achieved inconsistent patterns of coefficient of determination (R^2) and root

mean squared error (RMSE) variances in MCC and NMCC. Multicollinearity and the base-ensemble MLs concept did not affect the model confidence; rather, the latter was subject to the cross-effects of the ML and dataset characteristics. For the second objective, the denoised VIs (R^2 = 0.74-0.95, RMSE = 2.43–13.94 g q^{-1}) outperformed the noised VIs (R^2 = 0.63-0.90, RMSE = 3.28–17.91 g q^{-1}) for the TAGB prediction. The denoised VIs achieved the highest R^2 and lowest RMSE values at the booting stage (R^2 = 0.93-0.95, RMSE = 8.22-9.30 g q^{-1}), then tillering (R^2 = 0.75-0.84, RMSE = 2.43-2.96 g q^{-1}), and then milking stages (R^2 = 0.74-0.80, RMSE = 13.34-13.94 g q^{-1}). The HGBR achieved the lowest overfitting on the denoised VIs at the booting stage with a training-testing R^2's change (ΔR^2) of 0.02-0.09 and a training-testing RMSE's change (ΔRMSE) of 1.93-6.54 q^{-1}, tillering (ΔR^2 = 0.08-0.21, ΔRMSE = 1.23-2.36 g q^{-1}), and then milking stages (ΔR^2 = 0.14-0.25, ΔRMSE = 5.57-10.02 g q^{-1}). For the third objective, the TESI, TVAE, and CTGAN were applied to increase the four datasets' sizes. The TESI retained the features' original probability distribution in the four datasets. The C-TESI achieved the lowest mean squared error mean percentage (MAE_P) on the oil palm (0.60–2.85%), rice (0.77–1.72%), and fertiliser datasets (2.04–2.21%). The TESI retained the variance inflation factor (VIF) ranges less than 10 on the four datasets; the TESI retained a VIF range of 1.99–10.06 or reduced the VIF range to 1.55–6.66. Furthermore, the TESI retained the Spearman's r (r_s) range of 0.79–0.97 or increased it to 0.81-0.99 on the four datasets. The DNN achieved the highest R^2 (0.77–0.99) and lowest RMSE ranges (2.8E+01–8.1E+05) on the four datasets augmented with the TESI. The Q-TESI, C-TESI, and L-TESI overcame the LN-TESI in retaining the features' original probability distribution, minimising the augmentation loss, reducing the VIF, increasing the r_s, and decreasing the DNN under- and overfitting. Overall, as most of the agronomic research is conducted based on a few sensors' bands, vegetation indices are highly collinear. Therefore, exploring the multilevel sensitivity of different MLs to multicollinearity may address the methodological choices of several future agronomic studies. Additionally, stable VI-biomass models accurately reflect rice yield potential, which may be significantly improved by VIs' denoising. Further, the Q-TESI, C-TESI, and L-TESI minimise the proportionality of interpolation error to the square of the distance between the data points compared to the LN-TESI. Consequently, the Q-TESI, C-TESI, and L-TESI may approximate the nonlinear changes of crop phenology in time-spaced sampling, thereby reducing the cost of sampling for scientists. Furthermore, they intensify non-time series zonal, synthetic sampling, which reduces sampling labour.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

RAMALAN BIOJISIM PADI MENGGUNAKAN ALGORITMA PEMBELAJARAN MESIN

Oleh

DERRAZ RADHWANE

Disember 2022

Pengerusi : Farrah Melissa Muharam, PhD
Fakulti : Pertanian

Kaedah persampelan padi secara konvensional adalah berkesan. Namun, kaedah ini merosakkan, melelahkan, mengambil masa, tidak praktikal untuk kawasan yang besar, serta terdedah kepada ralat manusia. Kenderaan udara tanpa pemandu (UAV) mungkin dapat mengatasi masalah-masalah ini. Algoritma pembelajaran mesin (ML) boleh memprediksi biojisim padi melalui indeks-indeks vegetasi (VIs) yang berdasarkan UAV. Walau bagaimanapun, VIs bersifat sangat kolinear, hingar, dan pengumpulan data yang besar adalah mahal. Masalah-masalah ini mempengaruhi prestasi, kestabilan (underfitting/overfitting), varians, dan keyakinan model ML. Tujuan kajian ini adalah untuk: (i) membandingkan prestasi, varians, kestabilan, dan keyakinan model ML asas dan model ML gabungan dalam memprediksi biojisim padi menggunakan VIs yang kolinear (konteks multikolinearan (MCC)) dan tidak kolinear (konteks bukan multikolinearan (NMCC)); (ii) membandingkan kebolehprediksian biojisim padi di atas tanah (TAGB) daripada VIs yang hingar dan dinyahhingar dengan penapis Kalman melalui *histogram gradient boosting regressor* (HGBR); (iii) membangunkan sebuah interpolator yang bersifat trigonometric-Euclidean-lancar (TESI), termasuk interpolator linear (LN-TESI), interpolator kubik (C-TESI), interpolator kuadratik (Q-TESI), dan interpolator logaritmik (L-TESI), bagi kelangsungan augmentasi data VIs secara siri masa dan bukan siri masa, serta membandingkannya dengan pengekod automatik bervariasi jadual (TVAE) dan jaringan pengadversarial generatif jadual bersyarat (CTGAN) bagi mencegah DNN daripada underfitting/overfitting. Eksperimen dengan rekabentuk rawak lengkap (RCBD) split-plot dijalankan dalam kawasan jelapang padi di Terengganu, Malaysia, dengan 120 kuadran. Setiap kuadran memberikan lima ciri-ciri biojisim padi pada peringkat vegetatif, bunting, dan pengisian susu. Kamera multispektral MicaSense Red-Edge yang dipasang pada dron DJI quadcopter digunakan untuk mendapatkan jalur biru, hijau, merah, pinggir merah, dan NIR bagi mengekstrak nilai VIs setiap kuadran. Selain set data biojisim, data baja bukan siri masa dan data siri masa kelapa sawit dan padi juga diambil untuk verifikasi keputusan TESI, TVAE, dan CTGAN. Bagi

objektif pertama, prestasi dan kestabilan model ML adalah lebih baik dalam konteks MCC berbanding daripada NMCC dalam meramal semua ciri-ciri biojisim padi. Model ML gabungan mengungguli model ML asas dalam memprediksi semua ciri-ciri biojisim padi di MCC dan NMCC. Semua ML asas dan ML gabungan menunjukkan trend pekali penentuan (R^2) yang tidak konsisten dan ralat punca min kuasa dua (RMSE) yang bervariasi pada MCC dan NMCC. Multikolinearan dan konsep ML asas-gabungan tidak mempengaruhi keyakinan model; sebaliknya, yang terakhir terikat dengan kesan silang dari ciri-ciri ML dan set data. Manakala, untuk objektif kedua, VIs yang dinyahhingar (R^2 = 0.74-0.95, RMSE = 2.43-13.94 g q^{-1}) mendominasi VIs yang hingar (R^2 = 0.63-0.90, RMSE = 3.28-17.91 g q^{-1}) dalam ramalan TAGB. VIs yang dinyahhingar memperoleh nilai R^2 tertinggi dan RMSE yang terendah yang pada peringkat bunting (R^2 = 0.93-0.95, RMSE = 8.22-9.30 g q^{-1}), kemudian pada peringkat vegetatif (R^2 = 0.75-0.84, RMSE = 2.43-2.96 g q^{-1}), dan kemudian tahap pengisian susu (R^2 = 0.74-0.80, RMSE = 13.34-13.94 g q^{-1}). Model HGBR mencapai overfitting yang paling rendah pada VIs yang dinyahhingar pada peringkat bunting dengan perubahan R^2 data latihan-ujian (ΔR^2) sebesar 0.02-0.09 dan perubahan RMSE latihan-ujian (ΔRMSE) sebanyak 1.93-6.54 q^{-1}, peringkat vegetatif (ΔR^2 = 0.08-0.21, ΔRMSE = 1.23-2.36 g q^{-1}), dan seterusnya peringkat pengisisan susu (ΔR^2 = 0.14-0.25, ΔRMSE = 5.57-10.02 g q^{-1}). Pada objektif ketiga, TESI, TVAE, dan CTGAN diaplikasikan untuk meningkatkan saiz empat set data. TESI mengekalkan ciri taburan kebarangkalian asal dalam empat set data. C-TESI mencapai purata ralat min kuasa dua (MAEP) yang terendah pada data kelapa sawit (0.60-2.85%), padi (0.77-1.72%), dan baja (2.04-2.21%). TESI mengekalkan faktor inflasi varians (VIF) kurang dari 10 bagi empat set data; TESI juga memelihara VIF dengan julat 1.99-10.06 atau mengurangkan julat VIF kepada 1.55-6.66. Selain itu, TESI memelihara r Spearman (r_s) dengan julat 0.79-0.97 atau meningkatkannya kepada 0.81-0.99 dalam semua set data. DNN mencapai R^2 yang paling tinggi (0.77-0.99) dan RMSE yang paling rendah (2.8E+01-8.1E+05) dalam empat set data yang diaugmentasikan dengan TESI. Model Q-TESI, C-TESI, dan L-TESI mengatasi LN-TESI dalam mengekalkan ciri taburan kebarangkalian asal, meminimalkan kehilangan augmentasi, mengurangkan VIF, meningkatkan r_s, dan mengurangkan underfitting dan overfitting DNN. Secara keseluruhan, kerana kebanyakan penyelidikan agronomi dilakukan berdasarkan beberapa jalur spektral sensor, indeks vegetasi adalah sangat kolinear. Justeru, penerokaan kepekaan ML di pelbagai peringkat terhadap multikolineran berupaya membantu dalam pemilihan metodologi kajian agronomi di masa hadapan. Selain itu, model VI-biojisim yang stabil mempresentasikan potensi hasil padi dengan tepat, dan boleh ditambah baik secara signifikan dengan penyahhingaran VIs. Seterusnya, Q-TESI, C-TESI, dan L-TESI meminimumkan kadar ralat interpolasi dengan kuasa dua jarak di antara titik-titik data berbanding LN-TESI. Oleh itu, Q-TESI, C-TESI, dan L-TESI berupaya menganggarkan perubahan tidak linear fenologi tanaman dalam persampelan berselang, di mana hal ini dapat mengurangkan kos persampelan bagi para saintis. Maka, mereka dapat memperbanyakkan zon bukan siri masa dan persampelan sintetik, yang dapat mengurangkan tenaga kerja bagi persampelan.

ACKNOWLEDGEMENTS

First, Alhamdulillah, I am so grateful to Allah S.W.T for inspiring me with patience and giving me the strength to complete this thesis in the prescribed period. I am too thankful to my supervisor, Assoc. Prof. Dr. Farrah Melissa Muharam, for her supervision, knowledge, guidance, patience, and support during the PhD period. Her expertise and help play a significant part in this thesis. Allah will bless her and reward her efforts with his boundless gifts.

Besides, I would like to thank all my supervisory committee members for their valuable efforts: Dr Nurulhuda Binti Khairudin, Dr Noraini Binti Md Ahmad Jaafar, and Dr Ng Keng Yap. They guided me with patience and friendship. They significantly contributed to my thesis. I would also like to acknowledge the support of all my parents, brothers, and sister. Moreover, I appreciate the support of all my friends, Harireche Tarek Amar, Fatah Bousbia, and others. Special thanks to United Malacca Berhad (UMB) for providing the research fund. Furthermore, I thank Universiti Putra Malaysia and Malaysia's Ministry of Higher Education for allowing me to realise my ambition.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Farrah Melissa binti Muharam, PhD
Associate Professor
Faculty of Agriculture
Universiti Putra Malaysia
(Chairman)

Nurulhuda binti Khairudin, PhD
Senior Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Noraini binti Md Ahmad Jaafar, PhD
Senior Lecturer
Faculty of Agriculture
Universiti Putra Malaysia
(Member)

Ng Keng Yap, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 13 April 2023

Declaration by the Graduate Student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and the copyright of the thesis are fully-owned by Universiti Putra Malaysia, as stipulated in the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from the supervisor and the office of the Deputy Vice-Chancellor (Research and innovation) before the thesis is published in any written, printed or electronic form (including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials) as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld in accordance with the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2015-2016) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: ________________________ Date: __________________

Name and Matric No.: Radhwane Derraz

Declaration by Members of the Supervisory Committee

This is to confirm that:

- the research and the writing of this thesis were done under our supervision;
- supervisory responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2015-2016) are adhered to.

Signature:	______________________
Name of Chairman of Supervisory Committee:	Assoc. Prof. Dr. Farrah Melissa Muharam
Signature:	______________________
Name of Member of Supervisory Committee:	Dr. Nurulhuda Binti Khairudin
Signature:	______________________
Name of Member of Supervisory Committee:	Dr. Noraini Binti Md Ahmad Jaafar
Signature:	______________________
Name of Member of Supervisory Committee:	Dr. Ng Keng Yap

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

°C	degree celsius
GL	Green leaf biomass
DL	Dead leaf biomass
STEM	Stem biomass
SO	Storage organ biomass
TAGB	Total above ground biomass
RCBD	Randomized complete block design
TESI	Trigonometric-Euclidean-smoother interpolator
TVAE	Tabular variational autoencoder
CTGAN	Conditional tabular generative adversarial network
DNN	Deep neural network
HGBR	Histogram gradient boosting regressor
SNR	Signal-to-Noise ratio
KF	Kalman filter
RF	Random forest
PLSR	Partial least squared regressor
K-NN	K nearest neighbours
SVR	Support vector regressor
DT	Decision tree
ESR	Ensemble stacking regressor
VI	Vegetation index
UAVs	Unmanned aerial vehicles
R^2	Coefficient of determination
RMSE	Root mean squared error
EBDTs	Ensemble bagging decision trees
MCC	Multicollinearity context
NMCC	Non-multicollinearity context
BRF	Bagged Random Forest
BPLSR	Bagged Partial least squared regressor
BK-NN	Bagged K nearest neighbours
BSVR	Bagged Support vector regressor
BDT	Bagged Decision tree
N	Nitrogen
P	Phosphorus
K	Potassium
GLCM	Gray level co-occurrence matrix

CHAPTER 1

INTRODUCTION

1.1 Study Background

Rice is a staple food for half of the world's population, including Malaysia (Prasad et al., 2017). Whereas, Malaysia's rice production reached 1.8 million mega tonnes (MMT) in 2016 with an average annual growth of 1.62% in 2000–2016, while Malaysia's rice consumption reached 2.7 MT in the same year with an average annual growth of 1.75% in 2000–2016 (Omar et al., 2019). Furthermore, the gap between rice production and consumption is expected to widen around 2026; Malaysia's rice production is expected to reach 2.2 MMT, while Malaysia's rice consumption is expected to reach 3.5 MMT, and Malaysia's rice import volume is expected to reach 1.3 MMT. Additionally, Malaysia's rice self-sufficiency level (SSL) may reach 70% in 2026, compared to 80% in 1990 (Omar et al., 2019). Furthermore, Malaysia's population is expected to exceed 32.8 million in 2020, increasing to 41.1 million in mid-2050 (PRB, 2021). Thus, the growing population rate is expected to aggravate Malaysia's rice dilemma. Given Malaysia's growing rice production-consumption gap, increasing rice yield production is urgently needed to meet rice food security demands. Therefore, according to the 2011–2020 National Agrofood Policy (NAFP), Malaysia's domestic rice production became the priority for the country's staple food security (Nodin et al., 2022).

Crop yield is a function of the dry matter (biomass) partitioning amount in the source-sink system (Marcelis, 1996; Smith et al., 2018). Consequently, rice crop biomass reflects crop yield potential (Li et al., 2020; Mansaray, Zhang, et al., 2020). Thus, estimating rice crop biomass allows the prediction of rice yield (Gouranga and Kumar, 2014). The sampling of rice crop biomass traits, such as green leaf (GL) biomass, dead leaf (DL) biomass, stem (STEM) biomass, storage organ (SO) biomass, and the total above ground (TAGB), allows for the accurate estimation of rice biomass.

1.2 Problem Statement

Conventional rice sampling methods are effective. However, they are destructive, laborious, time-consuming, expensive, impractical for large fields, and subject to human error (Catchpole and Wheelert, 1992; Daughtry, 1990). Instead, remote sensing tools, such as satellites, ground-platform sensors, and unmanned aerial vehicles (UAVs), are novel methods to address these issues (Wójtowicz et al., 2016; Zheng & Moskal, 2009). However, on the one hand, satellites are limited by the low-resolution of acquired imagery, cloud cover, and revisit times for on-nadir data. Ground-platform sensors, on the other hand, are time-consuming, tedious, and have a limited spectrum (Rudd et al., 2017). UAVs may address the issues of satellites and ground-platform sensors (Barbedo, 2019). UAVs provide accurate rice biomass estimation (Devia et al., 2019); they provide rice biomass estimative spectral and hyperspectral data, such as vegetation indices (VIs) (Ndikumana et al., 2018b; Tilly et al., 2013; Tubaña et al., 2011).

However, VIs are sensitive to multicollinearity (Grüner et al., 2020); VIs are calculated from the same spectral bands (Fiorillo et al., 2020). Multicollinearity causes several statistical and numerical issues for machine learning algorithms (MLs). The former consists of the difficulties in testing individual regression coefficients due to inflated standard errors, resulting in a poor determination of significant features. The latter consists of difficulties in the computer's calculations due to numerical instability (Siegel, 2016). Further, VIs are noisy. VIs' noise increases the MLs' model complexity, which may cause model under- or overfitting (Atzberger et al., 2011; Broge and Leblanc, 2001; Quan et al., 2015). Furthermore, large VI data collection is expensive, while the MLs' expected generalisation error is inversely proportional to their training dataset size (Demir et al., 2021). For instance, connectionism-based MLs, such as deep neural networks (DNNs), are data-greedy MLs, and their training requires large datasets (Adadi, 2021; Antoniou et al., 2018).

1.3 Hypotheses Justification

Currently, accurate rice biomass estimation practises are leaning toward novel alternative methods. Methods that are inexpensive, time-effective, less laborious, less subject to human error, and efficient for large fields, such as UAVs (Lu et al., 2021). The latter provides descriptive multispectral data on rice biomass, such as VIs. The VI maximises sensitivity to the rice crop's vegetation-based physiological and phenological traits, including the green leaf, dead leaf, stem, storage organ, and total above-ground biomass (Friedl, 2018). Therefore, using machine learning algorithms (MLs), VIs are reliable predictors of rice biomass. The latter reflects rice yield potential and allows for the selection of high-yielding rice cultivars (Mansaray et al., 2020b). Nevertheless, the VIs are collinear and noisy (Atzberger et al., 2011; Grüner et al., 2020), and their large data collection is expensive (Taylor and Nitschke, 2018). The VIs' multicollinearity causes the MLs’ computational and statistical errors, the model's performance, stability, variance, and confidence (Siegel, 2016). The noise in the VIs may cause the MLs to develop a complex but unstable model. Additionally, training MLs, such as deep neural networks (DNNs), requires large datasets. Small and noisy VIs datasets can cause the model to underfit or overfit (Saseendran et al, 2019). Multicollinearity handling, denoising, and size augmentation of VIs' datasets are the expected solutions to improve the MLs model's performance, minimise its under- or overfitting, reduce its variance, and increase its confidence.

MLs, such as the guided regularised random forest (GRRF) (Fiorillo et al., 2020), the normal random forest (RF) (Triscowati et al., 2019), and the Bayesian algorithms (Jaya et al., 2020), were successfully applied to handle multicollinearity in other research domains. Likewise, denoising MLs filters, such as the fast Fourier transform (Choudhary and Gautam, 2022), wavelet transform (Fernandas et al., 2003), and the Kalman filter (Peesapati et al., 2013), were also successfully applied for spectral and spectral-like data denoising. Additionally, MLs, such as variational autoencoders and generative adversarial neural networks, were effective for spectral and spectral-like data augmentation (Miao et al., 2016; Xu et al., 2019).

1.4 Objectives

The current study focuses on the multicollinearity handling, denoising, and augmentation of UAV-based VIs using machine learning algorithms to address the three VIs issues, filling the revealed gap for accurately predicting rice biomass. In this regard, this study aims to:

1. Compare the base and ensemble MLs' model performance, variance, stability, and confidence for predicting rice biomass using collinear (multicollinearity context (MCC)) and non-collinear (non-multicollinearity context (NMCC)) VIs.

2. Compare the rice above ground biomass (TAGB) predictability from noised and Kalman filter' denoised VIs using histogram gradient boosting regressor (HGBR).

3. Develop a trigonometric-Euclidean-smoother interpolator (TESI), including linear (LN-TESI), cubic (C-TESI), quadratic (Q-TESI), and logarithmic (L-TESI) interpolators, for continuous time-series and non-time-series VIs data augmentation, and compare them to the tabular variational autoencoder (TVAE) and the conditional tabular generative adversarial network (CTGAN) for preventing DNN's under- or overfitting.

1.5 The study's scope and limitations

The current study considers rice biomass as a direct indicator of rice yield potential. Therefore, it explores new methods for rice biomass sampling. Along this line, the current study examines conventional rice sampling methods, revealing their drawbacks. Further, it promotes the use of remote sensing technologies as an alternative for rice sampling. Furthermore, it reveals the limitations of satellites and ground-based sensors and suggests UAVs as more potential tools for rice biomass sampling. Considering this, the study explores the use of UAV-based VIs and revolves around their issues (i.e., multicollinearity, noise, and data size) in predicting rice biomass. To fill the VI gap, the study proposes three possible solutions to the VI issues. Therefore, the current study aims to handle the multicollinearity, noise, and data size of UAV-based VIs using machine learning algorithms for accurately predicting rice biomass. To that end, the study proposed the objectives stated in Section 1.4. To achieve these objectives, the study conducted a split-plot randomised complete block design

(RCBD) experiment for rice biomass collection. Additionally, it used a UAV for spectral data acquisition. The spectral data was processed to extract the VIs. The study applied different concept based MLs to handle the three VI issues and then used them to predict rice biomass. At this point, handling the three VI issues was expected to help accurately predict rice biomass.

However, the study scope focuses on: (i) sampling biomass in a single field in a single granary (i.e., IADA KETARA) in Malaysia; (ii) sampling biomass in an experimental field of only 3.3 km2 of area; (iiii) sampling biomass only in the off-season; (iv) sampling biomass only during three growth stages (ie., tillering, booting, and milking); (v) limiting rice biomass sampling to 120 samples per growth stage; (vi) experimenting only on five *Oryza sativa* cultivars (i.e., including MR269, MR297, UPUTRA, MR219, and MR220CL2); (vii) limiting the imagery's spectral resolution to five bands (i.e., blue, green, red, RedEdge, and NIR) within the 450–880 spectrum; (viii) limiting the imagery's spatial resolution to 3 cm/pixel; (ix) and studying only VIs to predict rice biomass.

The above-mentioned scope boundaries may limit the generalisations in the conclusions. The spatial invariability of the field may mask the effects of soil variability on the biomass production of each cultivar. Additionally, rice cultivation in the off-season requires suitable irrigation conditions, which may affect rice biomass production; rising temperatures and changes in rainfall patterns have direct effects on rice biomass and yield, as well as indirect effects through changes in irrigation water availability in most regions. Moreover, the number of rice cultivars is limited, and different gene cultivars may produce different amounts of biomass per quadrant. Additionally, they may perform differently under different soil conditions because biomass production is influenced by (i) soil humidity, (ii) soil and air temperature, (iii) air humidity, (iv) photoperiod, (v) light intensity, (vi) soil fertility, i.e., soil nutrient availability, and (vii) genotype. Sampling biomass only in three growth stages lowers the quality of spectral imagery because different days may examine different climatic conditions. Consequently, densifying the UAV flights may cover such an issue and give insightful results regarding the rice crop phenology. Additionally, increasing the spectral and spatial resolutions of bands may better describe the vegetative dynamism of rice biomass. Finally, studying only VIs to predict rice biomass is not the only option; texture features and colour spaces may also be appropriate for this purpose and suffer from the same VI issues. Overall, insufficient funding justifies the limitations of the current study.

1.6 Thesis Outline

The present study highlights rice biomass as a direct indicator of rice yield potential. The latter's determination enables the selection of the most productive rice cultivars. To that end, this study uses an UAV as a non-destructive, time-efficient, inexpensive, less laborious, and large-field practical tool for sampling rice biomass. Further, it harnesses different concept MLs for UAV-based VIs analysis and rice biomass prediction. In accordance with that, the present study comprises five chapters. In Chapter 1, it revealed the three VI issues—the VI sensitivity to multicollinearity, the VI noise, and the VI small data size—that cause the MLs several computational and statistical issues, weakening the prediction of rice biomass. Further, it hypothesises three expected solutions to resolve the three VI issues, filling the VI gap: using MLs for VI multicollinearity handling, an ML filter for VI denoising, and MLs for VI data augmentation. To that end, the study established three objectives to achieve: (i) compare the base and ensemble MLs' model performance, variance, stability, and confidence for predicting rice biomass using collinear (multicollinearity context (MCC)) and non-collinear (non-multicollinearity context (NMCC)) VIs; (ii) compare the rice above ground biomass (TAGB) predictability from noised and Kalman filter' denoised VIs using histogram gradient boosting regressor (HGBR); (iii) develop a trigonometric-Euclidean-smoother interpolator (TESI), including linear (LN-TESI), cubic (C-TESI), quadratic (Q-TESI), and logarithmic (L-TESI) interpolators, for continuous time-series and non-time-series VIs data augmentation, and compare them to the tabular variational autoencoder (TVAE) and the conditional tabular generative adversarial network (CTGAN) for preventing DNN's under- or overfitting. In Chapter 2, it classified rice sampling methods as destructive and non-destructive, highlighted UAV-VIs as alternative non-destructive tools for rice biomass sampling and reviewed the use of MLs for VI multicollinearity handling, denoising, and data augmentation. In Chapter 3, it described the study field, the split-plot randomised complete block design (RCBD) experiment for rice biomass data collection, the MicaSense Red-Edge multispectral camera and DJI quadcopter UAV flight, UAV spectral data acquisition, UAV spectral data processing, UAV-VIs extraction, the use of MLS for VI multicollinearity handling, denoising, and data augmentation, the use of MLs and VIs for rice biomass prediction, and the MLs model's performance, stability, variance, and confidence evaluation. In Chapter 4, the study analyses the multicollinearity, denoising, and augmentation results, extracts and interprets their key findings, concludes the general hypotheses, and links them to the literature (findings generalisation). In Chapter 5, the study draws the general conclusion, highlights each objective significance and limitations, and set the future recommendations.

REFERENCES

Abbas, A., Ansumali, S., 2010. Global Potential of Rice Husk as a Renewable Feedstock for Ethanol Biofuel Production. Bioenergy Res. 3, 328–334. https://doi.org/10.1007/s12155-010-9088-0

Abdi, H., 2010. Partial least squares regression and projection on latent structure regression (PLS Regression). Wiley Interdiscip. Rev. Comput. Stat. 2, 97–106. https://doi.org/10.1002/wics.51

Abdul Halim, N.S. adah, Abdullah, R., Karsani, S.A., Osman, N., Panhwar, Q.A., Ishak, C.F., 2018. Influence of soil amendments on the growth and yield of rice in acidic soil. Agronomy 8, 1–11. https://doi.org/10.3390/agronomy8090165

Abid, M., Ahmad, N., Jahangir, M., Ahmad, I., 1986. Effect of Zinc , Iron and Manganese on Growth and Yield of Rice (Oryza Sativa L .). Pak J. Agri. Sc. 39, 3–6.

Aboneh, T., Rorissa, A., Srinivasagan, R., 2022. Stacking-Based Ensemble Learning Method for Multi-Spectral Image Classification. Technologies 10, 17. https://doi.org/10.3390/technologies10010017

Adadi, A., 2021. A survey on data-efficient algorithms in big data era. J. Big Data 8, 1–54. https://doi.org/10.1186/s40537-021-00419-9

Adeluyi, O., Harris, A., Foster, T., Clay, G.D., 2022. Exploiting centimetre resolution of drone-mounted sensors for estimating mid-late season above ground biomass in rice. Eur. J. Agron. 132, 126411. https://doi.org/10.1016/j.eja.2021.126411

Agapiou, A., Alexakis, D.D., Sarris, A., Hadjimitsis, D.G., 2013. Orthogonal equations of multi-spectral satellite imagery for the identification of un-excavated archaeological sites. Remote Sens. 5, 6560–6586. https://doi.org/10.3390/rs5126560

Aggarwal, C.C., 2018. Neural Networks and Deep Learning, An Introduction to Neural Networks, in: Springer Cham. pp. 1–52. https://doi.org/10.1007/978-3-319-94463-0

Aghbari, Z. Al, Saeed, M.M., 2021. Leveraging Association Rules in Feature Selection to Classify Text, Springer, Singapore. Springer, Singapore. https://doi.org/https://doi.org/10.1007/978-981-16-3728-5_53

Ahamed, T., Tian, L., Zhang, Y., Ting, K.C., 2011. A review of remote sensing methods for biomass feedstock production. Biomass and Bioenergy 35, 2455–2469. https://doi.org/10.1016/j.biombioe.2011.02.028

Al-Hajj, R., Assi, A., Mohamad, M.F., 2019. Stacking-Based Ensemble of

Support Vector Regressors for One-Day Ahead Solar Irradiance Prediction. IEEE; 2019 8th Int. Conf. Renew. Energy Res. Appl. 428–433.

Alin, A., 2010. Multicollinearity. Wiley Interdiscip. Rev. Comput. Stat. 2, 370–374. https://doi.org/10.1002/wics.84

Anbumozhi, V., Yamaji, E., Tabuchi, T., 1998. Rice crop growth and yield as influenced by changes in ponding water depth, water regime and fertigation level. Agric. Water Manag. 37, 241–253. https://doi.org/10.1016/S0378-3774(98)00041-9

Anderson, H.E., 1982. Aids to Determining Fuel Models For Estimating Fire Behavior. USDA For, Ser, Gen. Tech. Rep. INT-122, Ogden, UT 44, 42–53.

Andrew, M., Noble, I., Lange, R., 1979. A non-destructive method for estimating the weight of forage on shrubs. Rangel. J. 1, 225. https://doi.org/10.1071/rj9790225

Antoniou, A., Storkey, A., Edwards, H., 2018. Data augmentation using generative adversarial networks. Stat 1050, 8. https://doi.org/10.48550/arXiv.1711.04340

Antoulas, A.C., 1988. Rational interpolation and the Euclidean algorithm. Linear Algebra Appl. 108, 157–171. https://doi.org/10.1016/0024-3795(88)90185-1

Arifin, F., Robbani, H., Annisa, T., Ma'Arof, N.N.M.I., 2019. Variations in the Number of Layers and the Number of Neurons in Artificial Neural Networks: Case Study of Pattern Recognition. J. Phys. Conf. Ser. 1413, 1–7. https://doi.org/10.1088/1742-6596/1413/1/012016

Arora, S., Zhang, Y., 2017. Do GANs actually learn the distribution? An empirical study. arXiv 1–11. https://doi.org/10.48550/arXiv.1706.08224

Arslan, M., Guzel, M., Demirci, M., Ozdemir, S., 2019. SMOTE and Gaussian Noise Based Sensor Data Augmentation. 2019 4th Int. Conf. Comput. Sci. Eng. 1–5. https://doi.org/10.1109/UBMK.2019.8907003

Atanbori, J., Montoya-P, M.E., Selvaraj, M.G., French, A.P., Pridmore, T.P., 2019. Convolutional Neural Net-Based Cassava Storage Root Counting Using Real and Synthetic Images. Front. Plant Sci. 10, 1–14. https://doi.org/10.3389/fpls.2019.01516

Atzberger, C., Richter, K., Vuolo, F., Darvishzadeh, R., Schlerf, M., 2011. Why confining to vegetation indices? Exploiting the potential of improved spectral observations using radiative transfer models. Remote Sens. Agric. Ecosyst. Hydrol. XIII 8174, 263–278. https://doi.org/10.1117/12.898479

Auffarth, B., López, M., Cerquides, J., 2010. Comparison of Redundancy and

Relevance Measures for Feature Selection in Tissue Classification of CT Images. Adv. Data Min. Appl. Theor. Asp. 6171, 248–262. https://doi.org/10.1007/978-3-642-14400-4_20

Aziira, A.H., Setiawan, N.A., Soesanti, I., 2020. Generation of Synthetic Continuous Numerical Data Using Generative Adversarial Networks. J. Phys. Conf. Ser. 1577, 1–11. https://doi.org/10.1088/1742-6596/1577/1/012027

Baig, M.H.A., Zhang, L., Shuai, T., Tong, Q., 2014. Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. Remote Sens. Lett. 5, 423–431. https://doi.org/10.1080/2150704X.2014.915434

Baldi, P., Sadowski, P.J., 2013. Understanding dropout. Proc. 26th Int. Conf. Neural Inf. Process. Syst. 2, 2814–2822.

Barbedo, J.G.A., 2019. A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses. Drones 3, 1–27. https://doi.org/10.3390/drones3020040

Bauer, F., Lukas, M.A., 2011. Comparingparameter choice methods for regularization of ill-posed problems. Math. Comput. Simul. 81, 1795–1841. https://doi.org/10.1016/j.matcom.2011.01.016

Bel, A.J.E. va., Offler, C.E., Patrick, J.W., 2003. PHOTOSYNTHESIS AND PARTITIONING | Sources and Sinks. Encycl. Appl. Plant Sci. https://doi.org/10.1016/B0-12-227050-9/00089-2

Benoit, K., 2011. Linear regression models with logarithmic transformations. London Sch. Econ. 22, 23–36.

Bernier, P.Y., Robitaille, G., 2004. A plane intersect method for estimating fine root productivity of trees from minirhizotron images. Plant Soil 265, 165–173. https://doi.org/10.1007/s11104-005-0056-y

Bojanowski, J., Kowalik, W., Bochenek, Z., 2009. Noise reduction of NDVI time-series: a robust method based on Savitzky-Golay filter. Rocz. Geomatyki 7, 13–21.

Boon Teck, T., Pei Shan, F., Radin Firdaus, R.B., Mou Leong, T., Mahinda Senevi, G., 2021. Impact of climate change on rice yield in malaysia: A panel data analysis. Agric. 11. https://doi.org/10.3390/agriculture11060569

Breiman, L., 1996. Stacked Regressions. Mach. Learn. 49–64.

Broge, N.H., Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sens. Environ. 76, 156–172. https://doi.org/10.1016/S0034-4257(00)00197-8

Brown, J.K., Oberheu, R.D., Johnston, C.M., 1982. Handbook for inventorying surface fuels and biomass in the interior West. USDA For. Ser. Gen. Tech. Rep. INT 129.

Burruss, G.W., Bray, T.M., 2005. Confidence intervals. Encycl. Soc. Meas. 455–462. https://doi.org/10.1016/B0-12-369398-5/00060-8

Buscema, M., 1998. Back propagation neural networks. Subst. Use Misuse 33, 233–270. https://doi.org/10.3109/10826089809115863

Cao, L., 2010. Domain-driven data mining: Challenges and prospects. IEEE Trans. Knowl. Data Eng. 22, 755–769. https://doi.org/10.1109/TKDE.2010.32

Casanova, D., Epema, G.F., Goudriaan, J., 1998. Monitoring rice reflectance at field level for estimating biomass and LAI. F. Crop. Res. 55, 83–92. https://doi.org/10.1016/S0378-4290(97)00064-6

Catchpole, W.R., Wheelert, C.J., 1992. Review Estimating plant biomass : A review of techniques. Aust. J. Ecol. 17, 121–131. https://doi.org/10.1111/j.1442-9993.1992.tb00790.x

Cen, H., Wan, L., Zhu, J., Li, Y., Li, X., Zhu, Y., Weng, H., Wu, W., Yin, W., Xu, C., Bao, Y., Feng, L., Shou, J., He, Y., 2019. Dynamic monitoring of biomass of rice under different nitrogen treatments using a lightweight UAV with dual image-frame snapshot cameras. Plant Methods 15, 32. https://doi.org/10.1186/s13007-019-0418-8

Chang, K.W., Shen, Y., Lo, J.C., 2005. Predicting rice yield using canopy reflectance measured at booting stage. Agron. J. 97, 872–878. https://doi.org/10.2134/agronj2004.0162

Chao, Z., Liu, N., Zhang, P., Ying, T., Song, K., 2019. Estimation methods developing with remote sensing information for energy crop biomass: A comparative review. Biomass and Bioenergy 122, 414–425. https://doi.org/10.1016/j.biombioe.2019.02.002

Chawla, N. V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: Synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357. https://doi.org/10.1613/jair.953

Chen, J., Mowlaei, M.E., Shi, X., 2020. Population-scale Genomic Data Augmentation Based on Conditional Generative Adversarial Networks. Based Cond. Gener. Advers. Networks. Proc. 11th ACM Int. Conf. Bioinformatics, Comput. Biol. Heal. Informatics 26, 1–6. https://doi.org/10.1145/3388440.3412475

Chen, P., Liu, S., Zhao, H., Jia, J., 2020. Gridmask data augmentation. arXiv Prepr. arXiv 1–9. https://doi.org/https://doi.org/10.48550/arXiv.2001.04086

Chen, X., Kingma, D.P., Salimans, T., Duan, Y., Dhariwal, P., Schulman, J., Sutskever, I., Abbeel, P., 2017. Variational lossy autoencoder. arXiv Prepr. arXiv 2, 1–17. https://doi.org/10.48550/arXiv.1611.02731

Chen, Z., Miao, Y., Lu, J., Zhou, L., Li, Y., Zhang, H., Lou, W., Zhang, Z., Kusnierek, K., Liu, C., 2019. In-season diagnosis of winter wheat nitrogen status in smallholder farmer fields across a village using unmanned aerial vehicle-based remote sensing. Agronomy 9, 619. https://doi.org/10.3390/agronomy9100619

Cheng, Q., Xu, H., Fei, S., Li, Z., Chen, Z., 2022. Estimation of Maize LAI Using Ensemble Learning and UAV Multispectral Imagery under Different Water and Fertilizer Treatments. Agriculture 12, 1267. https://doi.org/10.3390/agriculture12081267

Cheng, T., Song, R., Li, D., Zhou, K., Zheng, H., Yao, X., Tian, Y., Cao, W., Zhu, Y., 2017. Spectroscopic estimation of biomass in canopy components of paddy rice using dry matter and chlorophyll indices. Remote Sens. 9, 319. https://doi.org/10.3390/rs9040319

Cheung, T.-H., Yeung, D.-Y., 2021. MODALS: Modality-agnostic automated data augmentation in the latent space. Int. Conf. Learn. Represent. 1–18. https://doi.org/https://hdl.handle.net/1783.1/109714

Choudhary, C., Gautam, A.K., 2022. Comparison of FFT and continuous wavelet transform to analysis of ECG signal. Int. J. Res. Publ. Rev. 3, 1000–1003.

Christmann, A., Steinwart, I., 2008. Support vector machines, 1st ed, Springer, New York, NY. https://doi.org/https://doi.org/10.1007/978-0-387-77242-4

Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I., 1990. STL: A seasonal-trend decomposition. J. Off. Stat. 6, 3–73.

Colorado, J.-D., Calderon, F., Mendez, D., Petro, E., Rojas, J.P., Correa, E.S., Mondragon, I.F., Rebolledo, M.C., Jaramillo-Botero, A., 2020a. A novel NIR-image segmentation method for the precise estimation of above-ground biomass in rice crops. PLoS One 15, e0239591. https://doi.org/10.1371/journal.pone.0239591

Colorado, J.-D., Cera-Bornacelli, N., Caldas, J.S., Petro, E., Rebolledo, M.C., Cuellar, D., Calderon, F., Mondragon, I.F., Jaramillo-Botero, A., 2020b. Estimation of nitrogen in rice crops from UAV-captured images. Remote Sens. 12, 3396. https://doi.org/10.3390/rs12203396

Crawford, S.L., 2006. Correlation and regression. Circulation 114, 2083–2088. https://doi.org/https://doi.org/10.1161/CIRCULATIONAHA.105.586495

Crist, E.P., 1985. A TM Tasseled Cap equivalent transformation for reflectance factor data. Remote Sens. Environ. 17, 301–306. https://doi.org/10.1016/0034-4257(85)90102-6

Cseresnyés, I., Szitár, K., Rajkai, K., Füzy, A., Mikó, P., Kovács, R., Takács, T., 2018. Application of electrical capacitance method for prediction of plant root mass and activity in field-grown crops. Front. Plant Sci. 9, 1–11. https://doi.org/10.3389/fpls.2018.00093

Cui, Z., Chen, W., Chen, Y., 2017. Convolutional neural networks for time series classification. J. Syst. Eng. Electron. 28, 162-169. https://doi.org/10.21629/JSEE.2017.01.18

Dagum, E.B., 2010. Time series modeling and decomposition. Statistica 70, 433–457. https://doi.org/10.6092/issn.1973-2201/3597

Daoud, J.I., 2018. Multicollinearity and regression analysis. J. Phys. Conf. Ser. 949, 012009. https://doi.org/10.1088/1742-6596/949/1/012009

Daughtry, C.S.T., 1990. Direct measurements of canopy structure. Remote Sens. Rev. 5, 45–60. https://doi.org/10.1080/02757259009532121

de Haan-Rietdijk, S., Voelkle, M.C., Keijsers, L., Hamaker, E.L., 2017. Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. Front. Psychol. 8, 1849. https://doi.org/10.3389/fpsyg.2017.01849

De, S., Bhakta, I., Phadikar, S., Majumder, K., 2022. Agricultural Image Augmentation with Generative Adversarial Networks GANs, in: Das, A.K., Nayak, J., Naik, B., Vimal, S., Pelusi, D. (Eds.), Computational Intelligence in Pattern Recognition. Springer Nature Singapore, Singapore, pp. 335–344.

De Swaef, T., Maes, W.H., Aper, J., Baert, J., Cougnon, M., Reheul, D., Steppe, K., Roldán-Ruiz, I., Lootens, P., 2021. Applying rgb-and thermal-based vegetation indices from UAVs for high-throughput field phenotyping of drought tolerance in forage grasses. Remote Sens. 13, 1–24. https://doi.org/10.3390/rs13010147

Demir, S., Mincev, K., Kok, K., Paterakis, N.G., 2021. Data augmentation for time series regression: Applying transformations, autoencoders and adversarial networks to electricity price forecasting. Appl. Energy 304, 117695. https://doi.org/10.1016/j.apenergy.2021.117695

Devia, C.A., Rojas, J.P., Petro, E., Martinez, C., Mondragon, I.F., Patino, D., Rebolledo, M.C., Colorado, J., 2019. High-Throughput Biomass Estimation in Rice Crops Using UAV Multispectral Imagery. J. Intell. Robot. Syst. 96, 573–589. https://doi.org/10.1007/s10846-019-01001-5

Dhanabal, S., Chandramathi, S., 2011. A Review of various k-Nearest Neighbor Query Processing Techniques. Int. J. Comput. Appl. 31, 14–22.

Dietterich, T.G., 2002. Ensemble learning. Dep. Comput. Sci. Oregon state Univ.

Doersch, C., 2016. Tutorial on Variational Autoencoders. arXiv Prepr. arXiv 1606, 1–23. https://doi.org/10.48550/arXiv.1606.05908

Dong, X., Yu, Z., Cao, W., Shi, Y., Ma, Q., 2020. A survey on ensemble learning. Front. Comput. Sci. 14, 241–258. https://doi.org/10.1007/s11704-019-8208-z

Du, K.-L., Swamy, M.N.S., 2019. Neural networks and statistical learning, Springer London. Springer-Verlag London. https://doi.org/10.1007/978-1-4471-7452-3

Fahad, S., Adnan, M., Noor, M., Arif, M., Alam, M., Khan, I.A., Ullah, H., Wahid, F., Mian, I.A., Jamal, Y., Basir, A., Hassan, S., Saud, S., Amanullah, Riaz, M., Wu, C., Khan, M.A., Wang, D., 2018. Major constraints for global rice production, in: Advances in Rice Research for Abiotic Stress Tolerance. Elsevier Inc., pp. 1–22. https://doi.org/10.1016/B978-0-12-814332-2.00001-0

Fan, C., Chen, M., Tang, R., Wang, J., 2022. A novel deep generative modeling-based data augmentation strategy for improving short-term building energy predictions. Build. Simul. 15, 197–211. https://doi.org/10.1007/s12273-021-0807-6

Fang, H., Liang, S., 2014. Leaf Area Index Models, in: Reference Module in Earth Systems and Environmental Sciences. Elsevier Inc., pp. 1–11. https://doi.org/10.1016/b978-0-12-409548-9.09076-x

Fang, H., Liang, S., 2008. Leaf Area Index Models. Encycl. Ecol. https://doi.org/10.1016/B978-008045405-4.00190-7

Fawakherji, M., Potena, C., Prevedello, I., Pretto, A., Bloisi, D.D., Nardi, D., 2020. Data Augmentation Using GANs for Crop/Weed Segmentation in Precision Farming. CCTA 2020 - 4th IEEE Conf. Control Technol. Appl. 279–284. https://doi.org/10.1109/CCTA41146.2020.9206297

Felderhof, L., Gillieson, D., 2014. Near-infrared imagery from unmanned aerial systems and satellites can be used to specify fertilizer application rates in tree crops. Can. J. Remote Sens. 37, 376–386. https://doi.org/10.5589/m11-046

Feng, H., Jiang, N., Huang, C., Fang, W., Yang, W., Chen, G., Xiong, L., Liu, Q., 2013. A hyperspectral imaging system for an accurate prediction of the above-ground biomass of individual rice plants. Rev. Sci. Instrum. 84, 095107. https://doi.org/10.1063/1.4818918

Feng, L., Zhang, Z., Ma, Y., Du, Q., Williams, P., Drewry, J., Luck, B., 2020. Alfalfa yield prediction using UAV-based hyperspectral imagery and ensemble learning. Remote Sens. 12. https://doi.org/10.3390/rs12122028

Feng, W., Wu, Y., He, L., Ren, X., Wang, Yangyang, Hou, G., Wang, Yonghua, Liu, W., Guo, T., 2019. An optimized non-linear vegetation index for estimating leaf area index in winter wheat. Precis. Agric. 20, 1157–1176. https://doi.org/10.1007/s11119-019-09648-8

Fernandas, F.C.A., Van Spaendonck, R.L.C., Burrus, C.S., 2003. A new framework for complex wavelet transforms. IEEE Trans. Signal Process. 51, 1825–1837. https://doi.org/10.1109/TSP.2003.812841

Fields, T., Hsieh, G., Chenou, J., 2019. Mitigating drift in time series data with noise augmentation. 2019 Int. Conf. Comput. Sci. Comput. Intell. 227–230. https://doi.org/10.1109/CSCI49370.2019.00046

Fiorillo, E., Di Giuseppe, E., Fontanelli, G., Maselli, F., 2020. Lowland rice mapping in Sédhiou region (Senegal) using sentinel 1 and sentinel 2 data and random forest. Remote Sens. 12, 1–23. https://doi.org/10.3390/rs12203403

Fletcher, J.E., Robinson, M.E., 1996. A Capacitance Meter for Estimating Forage Weight. J. Range Manag. 49, 520–529. https://doi.org/10.2111/RANGELANDS-D-10-00090.1

Friedl, M.A., 2018. Remote sensing of croplands, in: Comprehensive Remote Sensing. Elsevier, pp. 78–95. https://doi.org/10.1016/B978-0-12-409548-9.10379-3

Gao, B.-C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gastal, F., Lemaire, G., Durand, J., 2015. Quantifying crop responses to nitrogen and avenues to improve nitrogen-use efficiency, in: Crop Physiology. Elsevier Inc., pp. 161–206. https://doi.org/10.1016/B978-0-12-417104-6.00008-X

Gellman, M.D. (Ed.), 2020. Multiple Regression, in: Encyclopedia of Behavioral Medicine. Springer International Publishing, Cham, p. 1433. https://doi.org/10.1007/978-3-030-39903-0_301228

Gitelson, A.A., 2004. Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. J. Plant Physiol. 161, 165–173. https://doi.org/10.1078/0176-1617-01176

Gnyp, M.L., Miao, Y., Yuan, F., Ustin, S.L., Yu, K., Yao, Y., Huang, S., Bareth, G., 2014. Hyperspectral canopy sensing of paddy rice aboveground biomass at different growth stages. F. Crop. Res. 155, 42–55. https://doi.org/10.1016/j.fcr.2013.09.023

Goel, N.S., Qin, W., 1994. Influences of canopy architecture on relationships between various vegetation indices and LAI and Fpar: A computer simulation. Remote Sens. Rev. 10, 309–347.

https://doi.org/10.1080/02757259409532252

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2020. Generative adversarial networks. Commun. ACM 63, 139–144. https://doi.org/10.1145/3422622

Gouranga, K., Kumar, A., 2014. Forecasting rainfed rice yield with biomass of early phenophases, peak intercepted PAR and ground based remotely sensed vegetation indices. J. Agrometeorol. 16, 94–103. https://doi.org/10.54386/jam.v16i1.1492

Grami, A., 2016. Introduction to Digital Communications, Probability, Random Variables, and Random Processes, in: Academic Press. pp. 151–216. https://doi.org/10.1016/B978-0-12-407682-2.00004-1

Grandvalet, Y., 2004. Bagging equalizes influence. Mach. Learn. 55, 251–270. https://doi.org/10.1023/B:MACH.0000027783.34431.42

Grattan, S.R., Zeng, L., C.Shannon, M., Roberts, S.R., 2002. Rice is more sensitive to salinity than previously thought. Calif. Agric. 56, 189–198. https://doi.org/10.3733/ca.v056n06p189

Gregorich, M., Strohmaier, S., Dunkler, D., Heinze, G., 2021. Regression with highly correlated predictors: Variable omission is not the solution. Int. J. Environ. Res. Public Health 18. https://doi.org/10.3390/ijerph18084259

Grüner, E., Wachendorf, M., Astor, T., 2020. The potential of UAV-borne spectral and textural information for predicting aboveground biomass and N fixation in legume-grass mixtures. PLoS One 15, 1–21. https://doi.org/10.1371/journal.pone.0234703

Gurney, K., 2018. An Introduction to Neural Networks, Associative memories: the Hopfield net, in: CRC Press. pp. 134–163. https://doi.org/10.1201/9781315273570

Gurumurthy, S., Sarvadevabhatla, R.K., Babu, R.V., 2017. DeLiGAN: Generative Adversarial Networks for Diverse and Limited Data. 2017 IEEE Conf. Comput. Vis. Pattern Recognit. 4941–4949. https://doi.org/10.1109/CVPR.2017.525

Han, Y., Yu, L., 2012. A variance reduction framework for stable feature selection. Stat. Anal. Data Min. 5, 428–445. https://doi.org/10.1002/sam.11152

Hanafi, M.M., Hartinie, A., Shukor, J., Mahmud, T.M.M., 2009. Upland rice varieties in Malaysia: agronomic and soil physico-chemical characteristics. Pertanika J. Trop. Agric. Sci. 32, 225–246.

Hang, H., Huang, T., Cai, Y., Yang, H., Lin, Z., 2021. Gradient Boosted Binary Histogram Ensemble for Large-scale Regression. arXiv Prepr. arXiv arXiv-

2106. https://doi.org/10.48550/arXiv.2106.01986

Haralabopoulos, G., Torres, M.T., Anagnostopoulos, I., McAuley, D., 2021. Text data augmentations: permutation, antonyms and negation. Expert Syst. Appl. 177, 114769. https://doi.org/10.1016/j.eswa.2021.114769

Haruna, Y., Qin, S., Mbyamm Kiki, M.J., 2022. An Improved Approach to Detection of Rice Leaf Disease with GAN-Based Data Augmentation Pipeline. SSRN Electron. J. https://doi.org/10.2139/ssrn.4135061

Hashim, M.F.C., Nurulhuda, K., Haidar, A.N., Muharam, F.M., Nurulhuda, K., Berahim, Z., Ismail, M.R., Zad, S.N.M., Zulkafli, Z., 2022. Physiological and Yield Responses of Five Rice Varieties to Nitrogen Fertilizer Under Farmer's Field in IADA Ketara, Terengganu, Malaysia. Sains Malaysiana 51, 359–368. https://doi.org/10.17576/jsm-2022-5102-03

Hassan, A., Shahzad, A.N., Qureshi, M.K., 2022. Rice Production and Crop Improvement Through Breeding and Biotechnology, in: Sarwar, N., Atique-ur-Rehman, Ahmad, S., Hasanuzzaman, M. (Eds.), Modern Techniques of Rice Crop Production. Springer Singapore, Singapore, pp. 605–627. https://doi.org/10.1007/978-981-16-4955-4_30

Haydock, K.. P., Shaw, N.H.H., 1975. The Comparative yield method for estimating dry matter yield of pasture.\rMétodo de rendimiento comparativo. Aust. J. Exp. Agric. Anim. Husb. 15, 663–670.

Helland, I.S., 1988. On the structure of partial least squares regression. Commun. Stat. - Simul. Comput. 17, 581–607. https://doi.org/10.1080/03610918808812681

Hernández-Lobato, D., Martínez-Muñoz, G., Suárez, A., 2006. Pruning in ordered regression bagging ensembles. 2006 IEEE Int. Jt. Conf. Neural Netw. Proc. 1266–1273. https://doi.org/10.1109/ijcnn.2006.246837

Hird, J.N., McDermid, G.J., 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. Remote Sens. Environ. 113, 248–258. https://doi.org/10.1016/j.rse.2008.09.003

Ho, J., Chen, X., Srinivas, A., Duan, Y., Abbeel, P., 2019. Flow++: Improving flow-based generative models with variational dequantization and architecture design. Proc. 36th Int. Conf. Mach. Learn. PMLR 97, 2722–2730. https://doi.org/10.48550/arXiv.1902.00275

Hoerl, A.E., Kennard, R.W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 12, 55–67. https://doi.org/10.1080/00401706.1970.10488634

Hopkins, Wi.G., Huner, N.P.A., 1995. Photomorphogenesis: Responding To Light, in: An Introduction to Plant Physiology. Library ofCongress Cataloging-in-Publication Data, p. 528.

Hu, Y., Shen, J., Qi, Y., 2021. Estimation of rice biomass at different growth stages by using fractal dimension in image processing. Appl. Sci. 11, 7151. https://doi.org/10.3390/app11157151

Huang, F., Xia, X., Huang, Y., Lv, S., Chen, Q., Pan, Y., Zhu, X., 2022. Comparison of Winter Wheat Extraction Methods Based on Different Time Series of Vegetation Indices in the Northeastern Margin of the Qinghai–Tibet Plateau: A Case Study of Minhe, China. Remote Sens. 14. https://doi.org/10.3390/rs14020343

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci. 454, 903–995. https://doi.org/10.1098/rspa.1998.0193

Huang, S., Miao, Y., Yuan, F., Gnyp, M.L., Yao, Y., Cao, Q., Wang, H., Lenz-Wiedemann, V.I.S., Bareth, G., 2017. Potential of RapidEye and WorldView-2 satellite data for improving rice nitrogen status monitoring at different growth stages. Remote Sens. 9, 227. https://doi.org/10.3390/rs9030227

Huete, A.., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Huete, A., Didan, K., A, T.M., A, E.P.R., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Hutchings, S.S., Schmautz, J.E., 1969. A Field Test of the Relative-Weight-Estimate Method for Determining Herbage Production. J. Range Manag. 22, 408. https://doi.org/10.2307/3895852

IRRI, Africa Rice, C., 2010. Global Rice Science Partnership (GRiSP). Counc. Partnersh. Rice Res. Asia Metro Manila, Philipp. 105.

Islam, M.M., Tian, Y., Cheng, Y., Wang, Y., Hu, P., 2018. A deep neural network based regression model for triglyceride concentrations prediction using epigenome-wide DNA methylation profiles. BMC Proc. 12, 55–60. https://doi.org/10.1186/s12919-018-0121-1

Iwana, B.K., Uchida, S., 2021a. Time Series Data Augmentation for Neural Networks by Time Warping with a Discriminative Teacher. 2020 25th Int. Conf. Pattern Recognit. 3558–3565. https://doi.org/10.1109/ICPR48806.2021.9412812

Iwana, B.K., Uchida, S., 2021b. An empirical survey of data augmentation for time series classification with neural networks. PLoS One 16, e0254841. https://doi.org/10.1371/journal.pone.0254841

Iwana, B.K., Uchida, S., 2021c. An empirical survey of data augmentation for time series classification with neural networks. PLoS One 16, e0254841. https://doi.org/10.1371/journal.pone.0254841

Jagtenberg, W.D., 1970. Predicting the Best Time To Apply Nitrogen To Grassland in Spring. Grass Forage Sci. 25, 266–271. https://doi.org/10.1111/j.1365-2494.1970.tb01202.x

Jaya, I.G.N.M., Ruchjana, B.N., Abdulah, A.S., 2020. Comparison of Different Bayesian and Machine Learning Methods in Handling Multicollinearity Problem: a Monte Carlo Simulation Study. ARPN J. Eng. Appl. Sci. 15, 1998–2011.

Jiang, Q., Fang, S., Peng, Y., Gong, Y., Zhu, R., Wu, X., Ma, Y., Duan, B., Liu, J., 2019. UAV-based biomass estimation for rice-combining spectral, TIN-based structural and meteorological features. Remote Sens. 11. https://doi.org/10.3390/RS11070890

Jimenez-Sierra, D.A., Correa, E.S., Benítez-Restrepo, H.D., Calderon, F.C., Mondragon, I.F., Colorado, J.D., 2021. Novel feature-extraction methods for the estimation of above-ground biomass in rice crops. Sensors 21, 1–14. https://doi.org/10.3390/s21134369

Jordan, C.F., 1969. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. Ecology 50, 663–666. https://doi.org/10.2307/1936256

Jørgensen, N., 2019. Carbohydrates. Encycl. Inl. Waters. https://doi.org/10.1016/B978-012370626-3.00258-1

Kadyan, V., Bawa, P., Hasija, T., 2021. In domain training data augmentation on noise robust Punjabi Children speech recognition. J. Ambient Intell. Humaniz. Comput. 13, 2705–2721. https://doi.org/10.1007/s12652-021-03468-3

Kanke, Y., Tubaña, B., Dalen, M., Harrell, D., 2016a. Evaluation of red and red-edge reflectance-based vegetation indices for rice biomass and grain yield prediction models in paddy fields. Precis. Agric. 17, 507–530. https://doi.org/10.1007/s11119-016-9433-1

Kanke, Y., Tubaña, B., Dalen, M., Harrell, D., 2016b. Evaluation of red and red-edge reflectance-based vegetation indices for rice biomass and grain yield prediction models in paddy field. Precis. Agric. 17, 507–530. https://doi.org/10.1007/s11119-016-9433-1

Karki, S., Rizal, G., Quick, W.P., 2013. Improvement of photosynthesis in rice (Oryza sativa L.) by inserting the C4 pathway. Rice 6, 2–8. https://doi.org/10.1186/1939-8433-6-28

Katrutsa, A., Strijov, V., 2017. Comprehensive study of feature selection methods to solve multicollinearity problem according to evaluation criteria.

Expert Syst. Appl. 76, 1–11. https://doi.org/10.1016/j.eswa.2017.01.048

Kaufman, Y., Tanre., D., 1992. Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS. IEEE Trans. Geosci. Remote Sens. 30, 260–270. https://doi.org/10.1109/36.134076

Kawamura, K., Ikeura, H., Phongchanmaixay, S., Khanthavong, P., 2018. Canopy Hyperspectral Sensing of Paddy Fields at the Booting Stage and PLS Regression can Assess Grain Yield. Remote Sens. 10, 1249. https://doi.org/10.3390/rs10081249

Kawamura, K., Tsujimoto, Y., Rabenarivo, M., Asai, H., Andriamananjara, A., Rakotoson, T., 2017. Vis-NIR spectroscopy and PLS regression with waveband selection for estimating the total C and N of paddy soils in Madagascar. Remote Sens. 9, 1081. https://doi.org/10.3390/rs9101081

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y., 2017. LightGBM: A highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 2017-Decem, 3147–3155.

Kim, H., Choi, B.S., Huh, M.Y., 2016. Booster in High Dimensional Data Classification. IEEE Trans. Knowl. Data Eng. 28, 29–40. https://doi.org/10.1109/TKDE.2015.2458867

Kim, S., Kim, N.H., Choi, J.H., 2020. Prediction of remaining useful life by data augmentation technique based on dynamic time warping. Mech. Syst. Signal Process. 136, 106486. https://doi.org/10.1016/j.ymssp.2019.106486

Kim, S.R., Prasad, A.K., El-Askary, H., Lee, W.K., Kwak, D.A., Lee, S.H., Kafatos, M., 2014. Application of the Savitzky-Golay filter to land cover classification using temporal MODIS vegetation indices. Photogramm. Eng. Remote Sensing 80, 675–685. https://doi.org/10.14358/PERS.80.7.675

Kingma, D.P., Welling, M., 2014. Auto-Encoding Variational Bayes. Stat 1050, 1. https://doi.org/10.48550/arXiv.1312.6114

Kira, K., Rendell, L.A., 1992. A Practical Approach to Feature Selection. Mach. Learn. Proc. 1992 249–256. https://doi.org/10.1016/B978-1-55860-247-2.50037-1

Koppe, W., Gnyp, M.L., Hütt, C., Yao, Y., Miao, Y., Chen, X., Bareth, G., 2012. Rice monitoring with multi-temporal and dual-polarimetric terrasar-X data. Int. J. Appl. Earth Obs. Geoinf. 21, 568–576. https://doi.org/10.1016/j.jag.2012.07.016

Kramer, O., 2013. Dimensionality Reduction with Unsupervised Nearest Neighbors, Intelligent Systems Reference Library. https://doi.org/10.1007/978-3-642-38652-7

Kushwah, J.-S., Kumar, A., Patel, S., Soni, R., Gawande, A., Gupta, S., 2021. Comparative study of regressor and classifier with decision tree using modern tools. Mater. Today Proc. 56, 3571–3576. https://doi.org/10.1016/j.matpr.2021.11.635

Kwak, G.H., Park, N.W., 2019. Impact of texture information on crop classification with machine learning and UAV images. Appl. Sci. 9. https://doi.org/10.3390/app9040643

Lafi, S.Q., Kaneene, J.B., 1992. An explanation of the use of principal-components analysis to detect and correct for multicollinearity. Prev. Vet. Med. 13, 261–275. https://doi.org/10.1016/0167-5877(92)90041-D

Lan, L., You, L., Zhang, Z., Fan, Z., Zhao, W., Zeng, N., Chen, Y., Zhou, X., 2020. Generative Adversarial Networks and Its Applications in Biomedical Informatics. Front. public Heal. 8, 164. https://doi.org/10.3389/fpubh.2020.00164

Lanorte, A., Manzi, T., Nolè, G., Lasaponara, R., 2015. On the use of the Principal Component Analysis (PCA) for evaluating vegetation anomalies from LANDSAT-TM NDVI Temporal Series in the Basilicata Region (Italy). Comput. Sci. Its Appl. -- ICCSA 2015 9158, 204–216. https://doi.org/10.1007/978-3-319-21410-8_16

Lavda, F., Gregorová, M., Kalousis, A., 2020. Improving VAE Generations of Multimodal Data Through Data-Dependent Conditional Priors. Front. Artif. Intell. Appl. 325, 1254–1261. https://doi.org/10.3233/FAIA200226

Lehmann, T.M., Gönner, C., Spitzer, K., 1999. Survey: Interpolation methods in medical image processing. IEEE Trans. Med. Imaging 18, 1049–1075. https://doi.org/10.1109/42.816070

Li, H., Luan, Y., 2005. Boosting proportional hazards models using smoothing splines, with applications to high-dimensional microarray data. Bioinformatics 21, 2403–2409. https://doi.org/10.1093/bioinformatics/bti324

Li, P., Zhang, X., Wang, W., Zheng, H., Yao, X., Tian, Y., Zhu, Y., Cao, W., Chen, Q., Cheng, T., 2020. Estimating aboveground and organ biomass of plant canopies across the entire season of rice growth with terrestrial laser scanning. Int. J. Appl. Earth Obs. Geoinf. 91, 102132. https://doi.org/10.1016/j.jag.2020.102132

Li, R., Li, X., Chen, G., Lin, C., 2020. Improving Variational Autoencoder for Text Modelling with Timestep-Wise Regularisation. Improv. Var. Autoencoder Text Model. with Timestep-Wise Regul. Proc. 28th Int. Conf. Comput. Linguist. 2381–2397. https://doi.org/10.18653/v1/2020.coling-main.216

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine learning in agriculture: A review. Sensors (Switzerland) 18, 1–29.

https://doi.org/10.3390/s18082674

Lim, J.S., Abdul Manan, Z., Wan Alwi, S.R., Hashim, H., 2012. A review on utilisation of biomass from rice industry as a source of renewable energy. Renew. Sustain. Energy Rev. 16, 3084–3094. https://doi.org/10.1016/j.rser.2012.02.051

Liu, B., Wang, X., Dixi, M., Kwitt, R., Vasconcelos, N., 2018. Feature space transfer for data augmentation. Proc. IEEE Conf. Comput. Vis. pattern Recognit. 9090-9098. 201. https://doi.org/10.48550/arXiv.1801.04356

Liu, K. Lou, Li, Y.Z., Hu, H.W., 2015. Predicting Ratoon rice growth rhythm based on NDVI at key growth stages of main rice. Chil. J. Agric. Res. 75, 410–417. https://doi.org/10.4067/S0718-58392015000500005

Liu, M., Stella, F., Hommersom, A., J.F.Lucas, P., Boer, L., Bischoff, E., 2019. A comparison between discrete and continuous time Bayesian networks in learning from clinical time series data with irregularity. Artif. Intell. Med. 95, 104–117. https://doi.org/10.1016/j.artmed.2018.10.002

Loh, W.Y., 2011. Classification and regression trees. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 1, 14–23. https://doi.org/10.1002/widm.8

Lu, J., Eitel, J.U.H., Engels, M., Zhu, J., Ma, Y., Liao, F., Zheng, H., Wang, X., Yao, X., Cheng, T., Zhu, Y., Cao, W., Tian, Y., 2021. Improving Unmanned Aerial Vehicle (UAV) remote sensing of rice plant potassium accumulation by fusing spectral and textural information. Int. J. Appl. Earth Obs. Geoinf. 104, 102592. https://doi.org/10.1016/j.jag.2021.102592

Maimon, O., Rokach, L., 2005. Data mining and knowledge discovery handbook, Choice Reviews Online. Springer Science+Business Media, Inc. https://doi.org/10.5860/choice.48-5729

Major, D.J., Baret, F., Guyot, G., 2007. A ratio vegetation index adjusted for soil brightness. Int. J. Remote Sens. 11, 727–740. https://doi.org/10.1080/01431169008955053

Mallick, J., Talukdar, S., Ahmed, M., 2022. Combining high resolution input and stacking ensemble machine learning algorithms for developing robust groundwater potentiality models in Bisha watershed, Saudi Arabia. Appl. Water Sci. 12, 1–19. https://doi.org/10.1007/s13201-022-01599-2

Mallick, N., Tripathy, S., Mishra, A.N., Deb, S., Sahoo, R., 2021. Estimation of impact parameter and transverse spherocity in heavy-ion collisions at the LHC energies using machine learning. Phys. Rev. D 103, 94031. https://doi.org/10.1103/PhysRevD.103.094031

Mansaray, L.R., Kanu, A.S., Yang, L., Huang, J., Wang, F., 2020a. Evaluation of machine learning models for rice dry biomass estimation and mapping using quad-source optical imagery. GIScience Remote Sens. 57, 785–796.

https://doi.org/10.1080/15481603.2020.1799546

Mansaray, L.R., Zhang, K., Kanu, A.S., 2020b. Dry biomass estimation of paddy rice with Sentinel-1A satellite data using machine learning regression algorithms. Comput. Electron. Agric. 176, 105674. https://doi.org/10.1016/j.compag.2020.105674

Manthey, F.A., 2015. Starch: Sources and Processing, 1st ed, Encyclopedia of Food and Health. Elsevier Ltd. https://doi.org/10.1016/B978-0-12-384947-2.00658-9

Marcelis, L.F.M., 1996. Sink strength as a determinant of dry matter partitioning in the whole plant. J. Exp. Bot. 47, 1281–1291. https://doi.org/10.1093/jxb/47.Special_Issue.1281

Meier, U., 2001. Growth Stages ofMono- and Dicotyledonous Plants: BBCH Monograph. Fed. Biol. Res. Cent. Agric. For. Braunschweig, Ger. 66–70.

Miao, Y., Yu, L., Blunsom, P., 2016. Neural variational inference for text processing. Proc. 33rd Int. Conf. Int. Conf. Mach. Learn. 48, 1727–1736. https://doi.org/10.48550/arXiv.1511.06038

Michishita, R., Jin, Z., Chen, J., Xu, B., 2014. Empirical comparison of noise reduction techniques for NDVI time-series based on a new measure. ISPRS J. Photogramm. Remote Sens. 91, 17–28. https://doi.org/10.1016/j.isprsjprs.2014.01.003

Mikołajczyk, A., Grochowski, M., 2018. Data augmentation for improving deep learning in image classification problem. 2018 Int. Interdiscip. PhD Work. 17–122. https://doi.org/10.1109/IIPHDW.2018.8388338

Mirzaie, M., Darvishzadeh, R., Shakiba, A., Matkan, A.A., Atzberger, C., Skidmore, A., 2014. Comparative analysis of different uni- and multi-variate methods for estimation of vegetation water content using hyper-spectral measurements. Int. J. Appl. Earth Obs. Geoinf. 26, 1–11. https://doi.org/10.1016/j.jag.2013.04.004

Morcillo-Pallarés, P., Rivera-Caicedo, J.P., Belda, S., De Grave, C., Burriel, H., Moreno, J., Verrelst, J., 2019. Quantifying the robustness of vegetation indices through global sensitivity analysis of homogeneous and forest leaf-canopy radiative transfer models. Remote Sens. 11, 2418. https://doi.org/10.3390/rs11202418

Muharam, F.M., Nurulhuda, K., Zulkafli, Z., Tarmizi, M.A., Abdullah, A.N.H., Hashim, M.F.C., Zad, S.N.M., Radhwane, D., Ismail, M.R., 2021. Uav-and random-forest-adaboost (Rfa)-based estimation of rice plant traits. Agronomy 11, 915. https://doi.org/10.3390/agronomy11050915

Mukaka, M.M., 2012. Statistics corner: a guide to appropriate use of correlation in medical research. Malawi Med. J. 24, 69–71.

Munnaf, M.A., Haesaert, G., Van Meirvenne, M., Mouazen, A.M., 2020. Site-specific seeding using multi-sensor and data fusion techniques: A review, in: Advances in Agronomy. Elsevier Inc., pp. 241–323. https://doi.org/10.1016/bs.agron.2019.08.001

Munson, M.A., Caruana, R., 2009. On Feature Selection, Bias-Variance, and Bagging. Springer, Berlin, Heidelb. 144–159. https://doi.org/https://doi.org/10.1007/978-3-642-04174-7_10

Murozuka, E., Massange-Sánchez, J.A., Nielsen, K., Gregersen, P.L., Braumann, I., 2018. Genome wide characterization of barley NAC transcription factors enables the identification of grain-specific transcription factors exclusive for the Poaceae family of monocotyledonous plants. PLoS One 13, 1–28. https://doi.org/10.1371/journal.pone.0209769

Mutanga, O., Skidmore, A.K., 2010. Narrow band vegetation indices overcome the saturation problem in biomass estimation. Int. J. Remote Sens. 25, 3999–4014. https://doi.org/10.1080/01431160310001654923

Muthukumar, V., Vodrahalli, K., Sahai, A., 2019. Harmless interpolation of noisy data in regression. IEEE Int. Symp. Inf. Theory - Proc. 2019-July, 2299–2303. https://doi.org/10.1109/ISIT.2019.8849614

Naklang, K., Fukai, S., Nathabut, K., 1996. Growth of rice cultivars by direct seeding and transplanting under upland and lowland conditions. F. Crop. Res. 48, 115–123. https://doi.org/10.1016/S0378-4290(96)01029-5

Nanni, L., Maguolo, G., Paci, M., 2020. Data augmentation approaches for improving animal audio classification. Ecol. Inform. 57, 101084. https://doi.org/10.1016/j.ecoinf.2020.101084

Ndikumana, E., Minh, D.H.T., Nguyen, H.T.D., Baghdadi, N., Courault, D., Hossard, L., Moussawi, I. El, 2018a. Estimation of rice height and biomass using multitemporal SAR Sentinel-1 for Camargue, Southern France. Remote Sens. 10, 1–18. https://doi.org/10.3390/rs10091394

Ndikumana, E., Minh, D.H.T., Thu, D.N.H., Baghdadi, N., Courault, D., Hossard, L., Moussawi, brahim El, 2018b. Estimation of Rice Height and Biomass Using Multitemporal SAR Sentinel-1 for Camargue, Southern France 31. https://doi.org/10.1117/12.2325174

Nguyen, N.T., Mohapatra, P.K., Fujita, K., Nakabayashi, K., Thompson, J., 2003. Effect of nitrogen deficiency on biomass production, photosynthesis, carbon partitioning, and nitrogen nutrition status of Melaleuca and Eucalyptus species. Soil Sci. Plant Nutr. 49, 99–109. https://doi.org/10.1080/00380768.2003.10409985

Ni, R., Zhu, X., Lei, Y., Li, X., Dong, W., Zhang, C., Chen, T., Mburu, D.M., Hu, C., 2022. Effectiveness of Common Preprocessing Methods of Time Series for Monitoring Crop Distribution in Kenya. Agric. 12.

https://doi.org/10.3390/agriculture12010079

Nodin, M.N., Mustafa, Z., Hussain, S.I., 2022. Assessing rice production efficiency for food security policy planning in Malaysia: A non-parametric bootstrap data envelopment analysis approach. Food Policy 107, 102208. https://doi.org/https://doi.org/10.1016/j.foodpol.2021.102208

O'Brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673–690. https://doi.org/10.1007/s11135-006-9018-6

Okafor, E., Schomaker, L., Wiering, M.A., 2018. An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals. J. Inf. Telecommun. 2, 465–491. https://doi.org/10.1080/24751839.2018.1479932

Omar, S.C., Shaharudin, A., Tumin, S.A., 2019. The Status of the Paddy and Rice Industry in Malaysia, Khazanah Research Institute.

Park, Y., Ho, J.C., 2021. Tackling Overfitting in Boosting for Noisy Healthcare Data. IEEE Trans. Knowl. Data Eng. 33, 2995–3006. https://doi.org/10.1109/TKDE.2019.2959988

Paul, M.J., Lawlor, D.W., 2003. GENETIC MODIFICATION OF PRIMARY METABOLISM | Photosynthesis. Encycl. Appl. Plant Sci. https://doi.org/10.1016/B0-12-227050-9/00175-7

Paul, S., Masih, V.A., 2022. Study on new markets potential of hybrid paddy seed in Cooch Behar district of West Bengal. Pharma Innov. J. 11, 726–728.

Pechanec, J.F., Pickford, G.D., 1937. A Weight Estimate Method for the Determination of Range or Pasture Production 1 . Agron. J. 29, 894–904. https://doi.org/10.2134/agronj1937.00021962002900110003x

Peesapati, R., Sabat, S.L., Karthik, K.P., Nayak, J., Giribabu, N., 2013. Efficient hybrid Kalman filter for denoising fiber optic gyroscope signal. Optik (Stuttg). 124, 4549–4556. https://doi.org/10.1016/j.ijleo.2013.02.013

Perez, L., Wang, J., 2017. The Effectiveness of Data Augmentation in Image Classification using Deep Learning.

Pham, Q.B., Kumar, M., Di Nunno, F., Elbeltagi, A., Granata, F., Islam, A.R.M.T., Talukdar, S., Nguyen, X.C., Ahmed, A.N., Anh, D.T., 2022. Groundwater level prediction using machine learning algorithms in a drought-prone area. Neural Comput. Appl. 7. https://doi.org/10.1007/s00521-022-07009-7

Phillips, D.S.M., Clarke, S.E., 1971. The calibration of a weighted disc against pasture dry matter yield. Proc. New Zeal. Grassl. Assoc. 68–75.

Prasad, R., Shivay, Y.S., Kumar, D., 2017. Current Status, Challenges, and Opportunities in Rice Production, in: Rice Production Worldwide. Springer Cham, pp. 1–32. https://doi.org/10.1007/978-3-319-47516-5

PRB, population reference bureau, 2021. World Population Data Sheet.

Price, J.C., 1992. Estimating vegetation amount from visible and near infrared reflectances. Remote Sens. Environ. 41, 29–34. https://doi.org/10.1016/0034-4257(92)90058-R

Putten, P. Van Der, Someren, M. Van, 2004. A bias-variance analysis of a real world learning problem: The CoIL challenge 2000. Mach. Learn. 57, 177–195. https://doi.org/10.1023/B:MACH.0000035476.95130.99

Qi, G., Zhao, G., Xi, X., 2020. Soil salinity inversion of winter wheat areas based on satellite-unmanned aerial vehicle-ground collaborative system in coastal of the Yellow River Delta. Sensors 20, 6521. https://doi.org/10.3390/s20226521

Qian, Y., Ye, M., Zhou, J., 2013. Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features. IEEE Trans. Geosci. Remote Sens. 51, 2276–2291. https://doi.org/10.1109/TGRS.2012.2209657

Qiu, Z., Xiang, H., Ma, F., Du, C., 2020. Qualifications of Rice Growth Indicators Optimized at Different Growth Stages Using Unmanned Aerial Vehicle Digital Imagery Zhengchao. Remote Sens. 12, 3228. https://doi.org/10.3390/rs12193228

Quan, X., He, B., Li, X., 2015. A Bayesian Network-Based Method to Alleviate the Ill-Posed Inverse Problem: A Case Study on Leaf Area Index and Canopy Water Content Retrieval. IEEE Trans. Geosci. Remote Sens. 53, 6507–6517. https://doi.org/10.1109/TGRS.2015.2442999

Quinlan, J.R., 1987. Simplifying decision trees. Int. J. Man. Mach. Stud. 27, 221–234. https://doi.org/10.1016/S0020-7373(87)80053-6

Rahayu, S., Sugiarto*, T., Madu, L., Holiawati, H., Subagyo, A., 2017. Application of Principal Component Analysis (PCA) to Reduce Multicollinearity Exchange Rate Currency of Some Countries in Asia Period 2004-2014. Int. J. Educ. Methodol. 3, 75–83. https://doi.org/10.12973/ijem.3.2.75

Rasel, S.M.M., Chang, H.C., Ralph, T.J., Saintilan, N., Diti, I.J., 2021. Application of feature selection methods and machine learning algorithms for saltmarsh biomass estimation using Worldview-2 imagery. Geocarto Int. 36, 1075–1099. https://doi.org/10.1080/10106049.2019.1624988

Rasheed, B.A., Adnan, R., Saffari, S.E., 2016. Robust PC with wild bootstrap estimation of linear model in the presence of outliers, multicollinearity and heteroscedasticity error variance. AIP Conf. Proc. 1750.

https://doi.org/10.1063/1.4954632

Rasti, S., Bleakley, C.J., Holden, N.M., Whetton, R., Langton, D., O'Hare, G., 2022. A survey of high resolution image processing techniques for cereal crop growth monitoring. Inf. Process. Agric. 9, 300–315. https://doi.org/10.1016/j.inpa.2021.02.005

Reddy, D.E., 2017. Emerging Trends in Seedbanking for Food and Agriculture: An International Perspective. J. Agric. Food Inf. 18, 145–160. https://doi.org/10.1080/10496505.2017.1289092

Redpath, T.W., 1998. Signal-to-noise ratio in MRI. Br. J. Radiol. 71, 704–707. https://doi.org/10.1259/bjr.71.847.9771379

Richardson, A.J., Wiegand, C.L., 1977. Distinguishing vegetation from soil background information. Photogramm. Eng. Remote Sensing 43, 1541–1552.

Rodriguez-Galiano, V., Mendes, M.P., Garcia-Soldado, M.J., Chica-Olmo, M., Ribeiro, L., 2014. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). Sci. Total Environ. 476–477, 189–206. https://doi.org/10.1016/j.scitotenv.2014.01.001

Rostamian, A., Heidaryan, E., Ostadhassan, M., 2022. Evaluation of different machine learning frameworks to predict CNL-FDC-PEF logs via hyperparameters optimization and feature selection. J. Pet. Sci. Eng. 208, 109463. https://doi.org/10.1016/j.petrol.2021.109463

Roujean, J.L., Breon, F.M., 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sens. Environ. 51, 375–384. https://doi.org/10.1016/0034-4257(94)00114-3

Rudd, J.D., Roberson, G.T., Classen, J.J., 2017. Application of satellite, unmanned aircraft system, and ground-based sensor data for precision agriculture: A review. 2017 ASABE Annu. Int. Meet. 1700272. https://doi.org/10.13031/aim.201700272

Saberioon, M., Amin, M.S.M., Gholizadeh, A., Ezri, M.H., 2014. A review of optical methods for assessing nitrogen contents during rice growth. Appl. Eng. Agric. 30, 657–669. https://doi.org/10.13031/aea.30.10478

Sapkota, B., Singh, V., Neely, C., Rajan, N., Bagavathiannan, M., 2020. Detection of Italian ryegrass in wheat and prediction of competitive interactions using remote-sensing and machine-learning techniques. Remote Sens. 12. https://doi.org/10.3390/RS12182977

Sapkota, B.B., Popescu, S., Rajan, N., Leon, R.G., Reberg-Horton, C., Mirsky, S., Bagavathiannan, M. V., 2022. Use of synthetic images for training a deep learning model for weed detection and biomass estimation in cotton. Sci. Rep. 12, 19580. https://doi.org/10.1038/s41598-022-23399-z

Saseendran, A.T., Setia, L., Chhabria, V., Chakraborty, D., Roy, A.B., 2019. Impact of Noise in Dataset on Machine Learning Algorithms 1–8. https://doi.org/10.13140/RG.2.2.25669.91369

Sawicki, A., Zieliński, S.K., 2020. Augmentation of Segmented Motion Capture Data for Improving Generalization of Deep Neural Networks. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 12133 LNCS, 278–290. https://doi.org/10.1007/978-3-030-47679-3_24

Schober, P., Boer, C., Schwarte, L.A., 2018. Correlation coefficients: Appropriate use and interpretation. Anesth. Analg. 126, 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Sharma, S., James, W.L., 1981. Latent Root Regression: An Alternate Procedure for Estimating Parameters in the Presence of Multicollinearity. J. Mark. Res. 18, 154. https://doi.org/10.2307/3150950

Shen, D., Qin, C., Wang, C., Zhu, H., Chen, E., Xiong, H., 2021. Regularizing Variational Autoencoder with Diversity and Uncertainty Awareness, in: Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21) Regularizing. pp. 2964–2970. https://doi.org/10.24963/ijcai.2021/408

Shibayama, M., Akiyama, T., 1989. Seasonal visible, near-infrared and mid-infrared spectra of rice canopies in relation to LAI and above-ground dry phytomass. Remote Sens. Environ. 27, 119–127. https://doi.org/10.1016/0034-4257(89)90011-4

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on Image Data Augmentation for Deep Learning. J. Big Data 6. https://doi.org/10.1186/s40537-019-0197-0

Shrestha, N., 2020. Detecting Multicollinearity in Regression Analysis. Am. J. Appl. Math. Stat. 8, 39–42. https://doi.org/10.12691/ajams-8-2-1

Siegel, A.F., 2016. Multiple Regression: Predicting One Variable From Several Others, in: Practical Business Statistics. Elsevier Inc., pp. 355–418. https://doi.org/10.1016/B978-0-12-804250-2.00012-2

Sipper, M., Moore, J.H., 2022. AddGBoost: A gradient boosting-style algorithm based on strong learners. Mach. Learn. with Appl. 7, 100243. https://doi.org/10.1016/j.mlwa.2021.100243

Slinker, B.K., Neilands, T.B., Glantz, S.A., 2016. Primer of Applied Regression & Analysis of Variance Authors. McGraw-Hill Educ. 1216.

Smith, M.R., Rao, I.M., Merchant, A., 2018. Source-sink relationships in crop plants and their influence on yield development and nutritional quality. Front. Plant Sci. 871, 1–10. https://doi.org/10.3389/fpls.2018.01889

Stavrakoudis, D., Katsantonis, D., Kadoglidou, K., Kalaitzidis, A., Gitas, I.Z., 2019. stimating rice agronomic traits using drone-collected multispectral imagery. Remote Sens. 11, 545. https://doi.org/10.3390/rs11050545

Taiz, L., Zeiger, E., 2003. Plant physiology. Annals of Botany. https://doi.org/10.1017/9781108486392

Talukdar, B., 2020. Handling of Class Imbalance for Plant Disease Classification with Variants of GANs. 2020 IEEE 15th Int. Conf. Ind. Inf. Syst. ICIIS 2020 - Proc. 466–471. https://doi.org/10.1109/ICIIS51140.2020.9342728

Tanré, D., Deroo, C., Duhaut, P., Herman, M., Morcrette, J.J., Perbos, J., Deschamps, P.Y., 1990. Technical note Description of a computer code to simulate the satellite signal in the solar spectrum: the 5S code. Int. J. Remote Sens. 11, 659–668. https://doi.org/10.1080/01431169008955048

Taylor, L., Nitschke, G., 2018. Improving Deep Learning with Generic Data Augmentation. 2018 IEEE Symp. Ser. Comput. Intell. 1542–1547. https://doi.org/10.1109/SSCI.2018.8628742

Tian, L., Xue, B., Wang, Z., Li, D., Yao, X., Cao, Q., Zhu, Y., Cao, W., Cheng, T., 2021. Spectroscopic detection of rice leaf blast infection from asymptomatic to mild stages with integrated machine learning and feature selection. Remote Sens. Environ. 257, 112350. https://doi.org/10.1016/j.rse.2021.112350

Tilly, N., Hoffmeister, D., Cao, Q., Huang, S., Lenz-Wiedemann, V., Miao, Y., Bareth, G., 2014. Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. J. Appl. Remote Sens. 8, 083671. https://doi.org/10.1117/1.jrs.8.083671

Tilly, N., Hoffmeister, D., Cao, Q., Lenz-Wiedemann, V., Miao, Y., Bareth, G., 2013. Precise plant height monitoring and biomass estimation with Terrestrial Laser Scanning in paddy rice. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. 2, 295–300. https://doi.org/10.5194/isprsannals-II-5-W2-295-2013

Triscowati, D.W., Sartono, B., Kurnia, A., Domiri, D.D., Wijayanto, A.W., 2019. Multitemporal remote sensing data for classification of food crops plant phase using supervised random forest 1131102, 10. https://doi.org/10.1117/12.2547216

Tsunoo, E., Shibata, K., Narisetty, C., Kashiwagi, Y., Watanabe, S., 2021. Data augmentation methods for end-to-end speech recognition on distant-talk scenarios. arXiv Prepr. arXiv. https://doi.org/10.48550/arXiv.2106.03419

Tu, Y.K., Kellett, M., Clerehugh, V., Gilthorpe, M.S., 2005. Problems of correlations between explanatory variables in multiple regression analyses in the dental literature. Br. Dent. J. 199, 457–461. https://doi.org/10.1038/sj.bdj.4812743

Tubaña, B., Harrell, D., Walker, T., Teboh, J., Lofton, J., Kanke, Y., Phillips, S., 2011. Relationships of spectral vegetation indices with rice biomass and grain yield at different sensor view angles. Agron. J. 103, 1405–1413. https://doi.org/10.2134/agronj2011.0061

Tucker, C. j, 1980. A critical review of remote sensing and other methods for non-destructive estimation of standing crop biomass. Grass Forage Sci. 35, 177–182. https://doi.org/10.1111/j.1365-2494.1980.tb01509.x

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Vaghefi, N., Shamsudin, M.N., Radam, A., Rahim, K.A., 2016. Impact of climate change on food security in Malaysia: economic and policy adjustments for rice industry. J. Integr. Environ. Sci. 13, 19–35. https://doi.org/10.1080/1943815X.2015.1112292

van Vliet, J.A., Giller, K.E., 2017. Mineral Nutrition of Cocoa: A Review, in: Advances in Agronomy. Elsevier Inc., pp. 185–270. https://doi.org/10.1016/bs.agron.2016.10.017

Vapnik, V.N., Lerner, A., 1963. Pattern recognition using generalized portrait method. Autom. Remote Control 24, 2 Vapnik, V. N. 3 Lerner, A. 4 1963 5 Autom. Remot.

Vlascici, D., Pica, E.M., Fagadar-Cosma, E., Cosma, V., Bizerea, O., 2008. Thiocyanate and fluoride electrochemical sensors based on nanostructurated metalloporphyrin systems. J. Optoelectron. Adv. Mater. 10, 2303–2306.

Walker, B.H., 1970. An Evaluation of Eight Methods of Botanical Analysis on Grasslands in Rhodesia. J. Appl. Ecol. 7, 403. https://doi.org/10.2307/2401967

Wang, J., Huang, J., Gao, P., Wei, C., Mansaray, L.R., 2016. Dynamic mapping of rice growth parameters using HJ-1 CCD time series data. Remote Sens. 8, 1–19. https://doi.org/10.3390/rs8110931

Wang, J., Lan, C., Liu, C., Ouyang, Y., Qin, T., 2021. Generalizing to Unseen Domains: A Survey on Domain Generalization. IEEE Trans. Knowl. Data Eng. 4627–4635. https://doi.org/10.1109/TKDE.2022.3178128

Wang, J., Neskovic, P., Cooper, L.N., 2006. Neighborhood size selection in the k-nearest-neighbor rule using statistical confidence. Pattern Recognit. 39, 417–423. https://doi.org/10.1016/j.patcog.2005.08.009

Wang, L., Zhou, X., Zhu, X., Dong, Z., Guo, W., 2016. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. Crop J. 4, 212–219. https://doi.org/10.1016/j.cj.2016.01.008

Wang, Y., 2009. A Mathematical Introduction to Generative Adversarial Nets (GAN). arXiv Prepr. arXiv 00169. https://doi.org/10.48550/arXiv.2009.00169

Wang, Y., Li, J., 2005. The plant architecture of rice (Oryza sativa). Plant Mol. Biol. 59, 75–84. https://doi.org/10.1007/s11103-004-4038-x

Wei, J., Zou, K., 2020. EDA: Easy data augmentation techniques for boosting performance on text classification tasks. EMNLP-IJCNLP 2019 - 2019 Conf. Empir. Methods Nat. Lang. Process. 9th Int. Jt. Conf. Nat. Lang. Process. Proc. Conf. 6382–6388. https://doi.org/10.18653/v1/d19-1670

Wen, Q., Gao, J., Song, X., Sun, L., Xu, H., Zhu, S., 2019. RobustSTL: A Robust Seasonal-Trend Decomposition Algorithm for Long Time Series. Proc. AAAI Conf. Artif. Intell. 33, 5409–5416. https://doi.org/10.1609/aaai.v33i01.33015409

Wen, Q., Sun, L., Yang, F., Song, X., Gao, J., Wang, X., Xu, H., 2021. Time Series Data Augmentation for Deep Learning: A Survey. arXiv Prepr. arXiv 12478, 2002. https://doi.org/10.48550/arXiv.2002.12478

Winkworth, R.E., Perry, R.A., Rossetti, C.O., 1962. A Comparison of Methods of Estimating Plant Cover in an Arid Grassland Community. J. Range Manag. 15, 194. https://doi.org/10.2307/3895247

Wójtowicz, M., Wójtowicz, A., Piekarczyk, J., 2016. Application of remote sensing methods in agriculture. Commun. Biometry Crop Sci. 11, 31–50.

Wold, S., Ruhe, A., Wold, H., Dunn, W.J., 1984. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. SIAM J. Sci. Stat. Comput. 5, 735–743. https://doi.org/https://doi.org/10.1137/0905052

Wu, Y., Xu, L., 2021. Image generation of tomato leaf disease identification based on adversarial-vae. Agric. 11. https://doi.org/10.3390/agriculture11100981

Xiao, Y., Hua, J., Dougherty, E.R., 2007. Quantification of the impact of feature selection on the variance of cross-validation error estimation. Eurasip J. Bioinforma. Syst. Biol. 2007. https://doi.org/10.1155/2007/16354

Xu, L., Skoularidou, M., Cuesta-Infante, A., Veeramachaneni, K., 2019. Modeling tabular data using conditional GAN. Adv. Neural Inf. Process. Syst. 32.

Xu, Y., Noy, A., Lin, M., Qian, Q., Li, H., Jin, R., 2020. WeMix: How to Better Utilize Data Augmentation. arXiv Prepr. arXiv 01267, 2010. https://doi.org/10.48550/arXiv.2010.01267

Xue, J., Su, B., 2017. Significant remote sensing vegetation indices: A review of developments and applications. J. Sensors 2017. https://doi.org/10.1155/2017/1353691

Yang, F., Zhu, Y.M., Luo, J.H., Robini, M., Liu, J., Croisille, P., 2014. A Comparative Study of Different Level Interpolations for Improving Spatial Resolution in Diffusion Tensor Imaging. IEEE J. Biomed. Heal. Informatics 18, 1317–1327. https://doi.org/10.1109/JBHI.2014.2306937

Yang, M.-D., Boubin, J.G., Tsai, H.P., Tseng, H.H., Hsu, Y.C., Stewart, C.C., 2020. Adaptive autonomous UAV scouting for rice lodging assessment using edge computing with deep learning EDANet. Comput. Electron. Agric. 179, 105817. https://doi.org/10.1016/j.compag.2020.105817

Ye, H.-J., Zhan, D.-C., Chao, W.-L., 2022. Procrustean Training for Imbalanced Deep Learning. IEEE/CVF Int. Conf. Comput. Vis. 92–102. https://doi.org/10.1109/iccv48922.2021.00016

Yoo, W., Mayberry, R., Bae, S., Singh, K., He, Q., Jr, J.W.L., 2014. A study of effects of multicollinearity in the multivariable analysis. Int. J. Appl. Sci. Technol. 4, 9–19. https://doi.org/10.1186/s12889-022-12882-x

Yuan, J., Wang, D.L., Li, R., 2014. Remote sensing image segmentation by combining spectral and texture features. IEEE Trans. Geosci. Remote Sens. 52, 16–24. https://doi.org/10.1109/TGRS.2012.2234755

Zahari, S.M., Ramli, N.M., Mokhtar, B., 2014. Bootstrapped Parameter Estimation in Ridge Regression with Multicollinearity and Multiple Outliers. J. Appl. Environ. Biol. Sci 4, 150–156.

Zaji, A., Liu, Z., Xiao, G., Bhowmik, P., Sangha, J.S., Ruan, Y., 2022. Wheat spikes counting using object-level data augmentation. 2022 IEEE Int. Instrum. Meas. Technol. Conf. 1–6. https://doi.org/10.1109/I2MTC48687.2022.9806479

Zhang, C., Kovacs, J.M., 2012. The application of small unmanned aerial systems for precision agriculture: A review. Precis. Agric. 13, 693–712. https://doi.org/10.1007/s11119-012-9274-5

Zhang, C., Ma, Y., 2012. Ensemble Machine Learning, Methods and Applications, 1st ed, Ensemble Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-9326-7

Zhang, W., Liu, H., Wu, W., Zhan, L., Wei, J., 2020. Mapping rice paddy based on machine learning with sentinel-2 multi-temporal data: Model comparison and transferability. Remote Sens. 12, 1620.

https://doi.org/10.3390/rs12101620

Zhang, X., Han, Liangxiu, Dong, Y., Shi, Y., Huang, W., Han, Lianghao, González-Moreno, P., Ma, H., Ye, H., Sobeih, T., 2019. A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images. Remote Sens. 11, 1–16. https://doi.org/10.3390/rs11131554

Zhao, Q., Lenz-Wiedemann, V.I.S., Yuan, F., Jiang, R., Miao, Y., Zhang, F., Bareth, G., 2015. Investigating within-field variability of rice from high resolution satellite imagery in Qixing farm county, northeast China. ISPRS Int. J. Geo-Information 4, 236–261. https://doi.org/10.3390/ijgi4010236

Zhao, S., Song, J., Ermon, S., 2017. Towards deeper understanding of variational autoencoding models. arXiv Prepr. arXiv 08658, 1702. https://doi.org/10.48550/arXiv.1702.08658

Zheng, G., Moskal, L.M., 2009. Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors. Sensors 9, 2719–2745. https://doi.org/10.3390/s90402719

Zheng, H., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y., Cao, W., Zhu, Y., 2019. Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. Precis. Agric. 20, 611–629. https://doi.org/10.1007/s11119-018-9600-7

Zhou, X., Zheng, H.B., Xu, X.Q., He, J.Y., Ge, X.K., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C., 2017. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. ISPRS J. Photogramm. Remote Sens. 130, 246–255. https://doi.org/10.1016/j.isprsjprs.2017.05.003