Field Crops Research 302 (2023) 109044

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

The influence of nitrogen and variety on rice grain moisture content dry-down

James Brinkhoff[a,*], Brian W. Dunn[b], Tina Dunn[b]

[a] Applied Agricultural Remote Sensing Centre, University of New England, Armidale 2351, NSW, Australia
[b] New South Wales Department of Primary Industries, Yanco 2703, NSW, Australia

ARTICLE INFO

Keywords:
Rice
Nitrogen
Maturity
Grain moisture content
Harvest timing

ABSTRACT

Rice field management around maturity and harvest are some of the most difficult decisions growers face. Field drainage and harvest timing affect quality, yield, and post-harvest drying costs. These decisions are informed by grain moisture content (MC). Over three years, three sites and three varieties, we studied the field dry-down rate and time to optimal harvest MC. We showed that field-specific parameters significantly affected these characteristics, including rice variety, Nitrogen applied (NA), mid-season N uptake (NU) and dry matter (DM). Increased N and DM is associated with increased MC and thus delays time to harvest. We developed models based on linear regression and nonlinear machine learning (ML) algorithms, including parameters describing these field-specific conditions. Cross validation across the three years provided a realistic expectation of model prediction errors. A linear model with the addition of nonlinear predictors achieved competitive performance compared with more complex and less interpretable ML models. When MC was modeled as a function of days since heading, similar or better accuracy was achieved to using accumulated weather parameters. Moisture content was predicted with mean absolute error of 2.1 %. The predicted time from heading to harvest MC was improved by the inclusion of field-specific parameters (N and variety) from mean absolute error of 6.8 days to 5.7 days. The final linear regression model explained 80 % of the moisture variability in the dataset, and provided estimates of dry-down rates, moisture as a function of time, and time to reach harvest moisture. This study shows the importance of including field-specific parameters when estimating of rice harvest timing, and provides methods to model these effects.

1. Introduction

Rice is one of the most important staple food crops, providing nutrition for much of the world's population (Muthayya et al., 2014). Rice growers are motivated to optimize their crops, with processors often providing premiums and imposing penalties depending on quality (McCauley and Way, 2002). A key factor determining rice quality is harvest grain moisture content (MC) (Wang et al., 2021), which is determined largely by harvest timing. However, paddy drainage and harvest timing decisions are some of the most difficult growers face. Therefore, data and insights leading to optimized decision-making are desired (Sarkar et al., 2018; Dunn and Dunn, 2021). It is well known that nitrogen fertilization and biomass are key drivers of yields (Dunn et al., 2016), but the effects of these parameters on maturity and harvest timing is less understood.

In temperate growing regions, rice is mostly grown in irrigated environments, where fields are ponded for the majority of the season (Humphreys et al., 2006; Brinkhoff et al., 2022). After establishment, accumulated temperature drives progression through the vegetative growth stages (Darbyshire et al., 2019; Sharifi et al., 2018). At panicle initiation (PI), the head begins to form inside the stem. Heading is defined as the time when 50 % of the stems have flowered. At physiological maturity, the grains have accumulated maximum dry matter (Rajanna and Andrews, 1970), which typically occurs around 26–28 % MC (Ward et al., 2021). Ponded fields are drained around this time as the plants no longer need water, and to ensure field trafficability for harvest. After grain moisture content has reduced to the recommended level (around 22 % as discussed below), crops are harvested. Harvested grains are dried and stored at a MC around 12.5 % (Calderwood et al., 1980).

Rice yield is highly dependent on nitrogen (N) availability (Dunn et al., 2016). Studies have shown N application before continuous ponding results in the highest efficiency of N recovery by the plants

* Corresponding author.
E-mail address: james.brinkhoff@une.edu.au (J. Brinkhoff).

https://doi.org/10.1016/j.fcr.2023.109044
Received 17 November 2022; Received in revised form 6 June 2023; Accepted 10 July 2023
Available online 18 July 2023

(Wilson et al., 1998) and is most important for determining yield (Dunn et al., 2016). However, mid-season N applications are often used to address deficiencies. A mid-season application at the PI growth stage is effective (Wilson et al., 1998), and results in more efficient recovery of the applied N than if applied at earlier vegetative growth stages (Peng and Cassman, 1998). It is not effective to address N deficiencies later in the season, for example at heading (Fageria and Baligar, 1999). In light of this, many studies seek to predict N uptake at PI using proximal and remote sensing methodologies (Lee et al., 2009; Inoue et al., 2012; Brinkhoff et al., 2021), in order to provide spatial N top-dressing recommendations. While anticipating operational remote-sensing-based systems, growers often sample crops for biomass and N uptake at PI, in order to determine optimal top-dressing rates (Dunn et al., 2016). Therefore, N and biomass status at the PI growth stage are well-known parameters in many rice growing systems, and are used in this study as a source of information on field-specific conditions.

Harvest timing is a critical decision growers face. If rice is harvested too early, head rice yield may be reduced (Siebenmorgen et al., 2007). In addition, if growers deliver grains with high MC, they may be penalized by processors for the increased energy costs incurred to dry the product to the target storage MC (Calderwood et al., 1980). On the other hand, if rice is harvested too late, the low MC can lead to grain fissure, which can be further exacerbated by mechanical stress during the harvest process or if rain and re-drying occurs (Calderwood et al., 1980). This results in lower quality rice and lower whole grain yield. Siebenmorgen et al. (2007) found optimal harvest timing to maximize head rice yield was at MC of 19–22 % for long-grain cultivars, and 22–24 % for medium-grain cultivars. In Australia, harvest is recommended at MC of 22 % (Ward et al., 2021). Paddy drainage and harvest timing decisions are complicated by the need to ensure field trafficability and possible adverse weather before harvest (Dingkuhn and Le Gal, 1996). Therefore, there is great interest in providing information to help optimize drainage and harvest timing (McCauley and Way, 2002).

Current practices to determine harvest timing in many regions are inaccurate, non-specific or laborious (Yang et al., 2021; Wang et al., 2021). There are rules-of-thumb related to grain color and grain texture (milky, doughy, hard) (Ward et al., 2021), which are subjective. Other guidelines approximate days between heading and harvest readiness (Wang et al., 2021). Some processors provide MC measurement services for field samples taken by growers. However, the results may not represent the harvest readiness of the whole field due to in-field variability (Dunn and Dunn, 2021), as samples are likely to be taken from locations that do not represent average field MC. Additionally, time and resource constraints on rice growers (who often manage many fields over large areas) motivate development of tools that can improve efficiency (Xu et al., 2019). A model to predict MC is highly desired, allowing targeted samples to be taken only at critical times, or perhaps even removing the need for MC sampling altogether.

Recent work has predicted rice MC at a point in time using spectroscopy (Lin et al., 2019), photos taken by a smartphone (Yang et al., 2021), and UAV imagery (Sarkar et al., 2018; Dunn and Dunn, 2021). Other work has sought to establish grain dry-down models as a function of weather parameters for soy and maize crops (Martinez-Feria et al., 2019; McCormick et al., 2021; Chazarreta et al., 2023) and rice (Lu and Siebenmorgen, 1994). However, in addition to weather and time, field-specific variables can also influence dry-down dynamics. Xu et al. (2022a) found that maize leaf area was correlated with grain moisture content at maturity, implying higher leaf area corresponds with slower dry-down. Corn grain moisture is impacted by hybrid (Ward et al., 2016), by in-field variability (Miao et al., 2006) and by nitrogen rate (Zhang et al., 2021). There is less information available on the impact of these parameters on rice dry-down.

In this work, we aimed to characterize field dry-down of rice. In particular, we studied how grain MC is influenced by nitrogen, biomass and variety. We tested a variety of nitrogen rates (0–180 kg/ha), and short, medium and long grain varieties over three years. MC models with predictors including a variety of accumulated weather and time variables were assessed. The accuracy improvement resulting from adding field-specific variables (variety, nitrogen and biomass) to the models was quantified. The results provided insights that will help growers take into account field-specific conditions when choosing harvest timing, in order to maximize productivity and quality.

2. Methods

2.1. Sites, years and experiments

Three experiment sites with variable soil types were included in the study. Site 1 (RRAPL 35.34°S, 145.52°E) has a grey clay soil type. Site 2 (LFS 34.61°S, 146.36°E) has a grey self-mulching clay soil. Site 3 (YAI 34.61°S, 146.42°E) has a red-brown earth soil.

There were variable weather conditions over the three years, with 2019 being hotter and drier than 2020 and 2021 (Fig. 1a).

Three semi-dwarf rice varieties with different grain types were included (Troldahl et al., 2014). Reiziq is a bold medium grain variety with high yield potential and average grain weight of 29 mg. Langi is a long grain soft cooking (low amylose) variety with 23 mg grain weight. Opus is a short grain sushi variety, also with 23 mg grain weight.

Table 1 provides a description of the eight experiments. Rice seed was drill sown (Dunn and Ford, 2018) with a row spacing of 20 cm and target plant population of 100–200 plants/m^2. Plot sizes ranged from 55 to 135 m^2. The plots were intermittently irrigated until ponding, following common practice for drill sown rice in Australia (Ward et al., 2021). Each site included an experiment with standard sowing and ponding dates, and an experiment with delayed permanent water (DPW, ponding in late December). The DPW experiments were sown earlier to account for slower growth during the longer intermittent irrigation period, as recommended in Ward et al. (2021).

Plots at each experiment included multiple varieties and N rates from 0 to 180 kg/ha as listed in Table 1. There were 2 replicates of each variety and N rate combination in 2019, and 3 replicates in 2020 and 2021. The 2020 and 2021 experiments included all 3 varieties, while the 2019 experiments had 2 varieties each (Table 1). The 0 and 180 kg/ha plots of the DPW experiments in 2019 were not sampled because bird damage and lodging rendered them unrepresentative.

2.2. Plot sampling methodology

Each plot was first sampled around 7 January (close to the timing of the panicle initiation (PI) growth phase). These samples were used to determine the N uptake (NU kg/ha) and above-ground dry matter (DM kg/ha). These samples involved gathering above-ground biomass from 1 m × 4 rows. The samples were dried at 60 °C and weighed. Subsamples of whole plants were ground and mixed prior to analysis for N concentration by Dumas combustion. The N concentration was multiplied by the measured DM weight to calculate NU in kg/ha.

Secondly, the heading date for each plot was recorded. This was defined as the date on which 50 % of stems had flowered.

Thirdly, each plot was sampled at multiple dates to obtain a time-series of grain moisture. The median number of samples per plot was 4, and the maximum was 8. Samples were collected after morning dew had evaporated. Each sample consisted of three hand-grabs, which were threshed, resulting in approximately 500 g of grain.

The moisture content of the samples were tested using a Cropscan 2000B near-infrared transmission instrument. Prior to this work, this instrument was calibrated to determine rice moisture content using 197 field samples covering the moisture range of 12–30 %. The reference moisture content of these samples was determined using the two-stage air-oven reference method (AACC, 1999). The developed calibration was tested independently in the following season, producing an R^2 of 0.97.

In all, 911 moisture samples were obtained over the three years

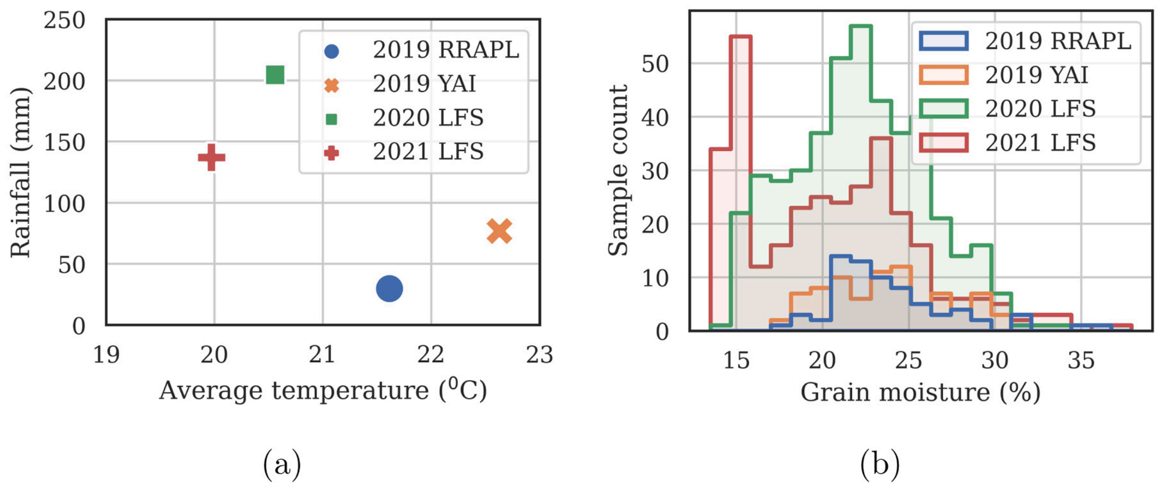

Fig. 1. (a) February–March average temperature and cumulative rainfall over each of the site-years. (b) Histogram of sampled moisture contents.

Table 1
List of experiments, with the number of plots and moisture samples collected for each. In the varieties column, S = short, M = medium and L = long grain.

Year	Site	Sowing	Ponding	N rates (kg/ha)	Varieties	Plots	Samples
2019	RRAPL	29 Oct	8 Dec	0,60,120,180	S,M	16	49
2019	RRAPL	10 Oct	21 Dec	60,120	S,M	8	21
2019	YAI	25 Oct	3 Dec	0,60,120,180	M,L	16	57
2019	YAI	17 Oct	24 Dec	60,120	M,L	8	25
2020	LFS	23 Oct	21 Nov	0,60,120,180	S,M,L	36	210
2020	LFS	11 Oct	20 Dec	0,60,120,180	S,M,L	36	288
2021	LFS	24 Oct	25 Nov	0,60,120,180	S,M,L	33	77
2021	LFS	14 Oct	23 Dec	0,60,120,180	S,M,L	36	246

(Table 1). The MC was distributed around the target harvest MC of 22 % (see MC histogram in Fig. 1b). The MC data for each plot was linearly interpolated on a daily basis, the date where the interpolated moisture was closest to 22 % was recorded. Then the number of days from heading to MC = 22 % was calculated. Of the 189 sampled plots, 177 had a moisture sample within 1 % of 22 %.

2.3. Predictors of moisture dry-down

We first assessed how the field-specific parameters (such as variety, nitrogen and dry matter, Table 2) impact the number of days from heading to MC = 22 % using Tukey's honest significant difference (HSD) test.

Daily weather information for each site was downloaded from the SILO dataset (Jeffrey et al., 2001). This data included vapor pressure deficit (VPD), reference evapotranspiration (ETo) and solar radiation. Degree days were calculated with base temperature of 10°C, which is the value used in many rice PI phenology models (Darbyshire et al., 2019; Sharifi et al., 2017). Specific to each plot, the 5 time and weather variables were accumulated from the heading date (listed in Table 2). We chose the heading date as the starting point because it is easily observed in each rice field and represents the beginning of the maturing growth stages. We determined the power of the accumulated weather variables to predict grain moisture dry-down by calculating the correlation coefficients between each of them and the 911 moisture samples.

Table 2
Predictive variables for MC models. The first five variables were accumulated from heading to the date MC is being predicted for. $\sum$Days is the number of days since heading. GDD is growth degree days, ETo is reference evapotranspiration, and VPD is vapor pressure deficit. For the static variables, NA is nitrogen applied, NU is nitrogen uptake and DM is the above-ground dry matter.

Group	Variable	Units
Cumulative from heading	$\sum$Days	Days
	$\sum$GDD	°C
	$\sum$ETo	mm
	$\sum$SolarRadiation	MJ/m^2
	$\sum$VPD	hPa
Field-specific	NA	kg/ha
	NU	kg/ha
	DM	g/m^2
	Variety	Short, medium, long

2.4. Dry-down model training and testing

Many models were trained and tested using combinations of predictor variables (Table 2) and algorithms. As an example, a model f with 3 predictors [$\sum$GDD, NU, Variety], is trained to predict MC at time step t, and is described by:

$$\mathrm{MC}(t) = f\left(\sum_{\mathrm{Heading}}^{t} \mathrm{GDD}, \mathrm{NU}, \mathrm{Variety}\right) \tag{1}$$

where f is a function that is determined using the algorithms described below.

We used a variety of linear and nonlinear algorithms to model the relationship between combinations of the variables in Table 2 and grain moisture. The scikit-learn, statsmodels and LightGBM packages (Pedregosa et al., 2011; Seabold and Perktold, 2010; Ke et al., 2017) were used to implement the algorithms. These included the tree-based models random forest (RF) and LightGBM (LGBM); support vector regression (SVR) with radial basis function kernel; and linear regression (LR). The grain dry-down vs time characteristic is nonlinear (Martinez-Feria et al., 2019), which the RF, LGBM and SVR algorithms can model. For the linear models, we used polynomials of the cumulative variables (Table 2) to allow the nonlinear characteristic to be described.

Randomly splitting train and test data can result in over-optimistic error estimates for agricultural studies (Brinkhoff et al., 2019). Therefore, we adopted a leave-one-year-out cross-validation strategy to ensure we obtained realistic estimates of prediction errors when applying models to new seasons. This involved running 3 experiments for each algorithm/variable combination, (i) train model on 2018 + 2019, test on 2020, (ii) train model on 2018 + 2020, test on 2019, (iii) train model on 2019 + 2020, test on 2018. The model predictions for grain moisture (%) were assessed using standard metrics including mean absolute error (MAE), and Lin's concordance correlation coefficient (LCCC) (Lin, 1989). LCCC has a range of 0–1 and measures how close predicted values are to the actual values (i.e. how close points are to the 1:1 line on an actual vs predicted graph).

After predicting the moisture for each date using the developed models, we determined the date MC was predicted to reach 22 % (see the example in Fig. 2). This was then compared to the actual MC = 22 % dates from the interpolated field samples (described above). The errors between actual and predicted 22 % dates over all plots were assessed using the MAE and LCCC.

After assessing the various predictors and algorithms using the cross validation strategy described above, we selected the most suitable combination, and developed a model trained on all data. This model is discussed as a basis for understanding the impact of accumulated time/weather and field-specific variables on rice grain moisture dry-down dynamics, and to provide a basis for predicting grain moisture in new growing seasons.

3. Results

3.1. Factors affecting dry-down rates and time to reach MC = 22 %

Heading occurred between 22 January and 19 February (189 plots). MC = 22 % was reached between 10 March and 24 April (177 plots). On average, there were 51 days between heading and MC = 22 %.

Days from heading to MC = 22 % for the three varieties was compared (Fig. 3b). The medium grain variety took longer to reach MC = 22 % compared to the long grain variety (by 6.5 days, $p = 0.001$), and compared to the short grain variety (by 2.6 days, $p < 0.01$). The short grain variety was slower to reach harvest moisture than the long grain variety, but the difference was not significant ($p = 0.1$).

The number of days from heading to MC = 22 % was similar in 2019 and 2020, but was less in 2021 by approximately 5 days ($p = 0.001$, Fig. 3c).

Fig. 3a shows the average moisture at each day since heading, separated by applied N rates. There were significant differences in days from heading to MC = 22 % between all combinations of N rates (0,60,120,180 kg/ha, $p < 0.01$), with higher rates taking longer to reach MC = 22 % (Fig. 3d). The time from heading to MC = 22 % was significantly and positively correlated with both PI dry matter (Fig. 3e, $R^2 = 0.18$) and PI N uptake (Fig. 3f, $R^2 = 0.28$).

We calculated the correlation between MC and the five variables accumulated from heading (Table 2), using all 911 samples. ∑SolarRadiation was best correlated with MC, explaining 74 % of the variation, and ∑ETo and ∑Days both explained 69 %. ∑GDD and ∑VPD were less correlated (66 % and 47 % respectively).

3.2. Leave-one-year-out cross-validated model comparisons

The leave-one-year-out cross-validation results of models trained using a range of input features are shown in Table 3.

Experiment set 1 in Table 3 compares models built using the LGBM algorithm and a single cumulative predictor. ∑SolarRadiation is the best predictor, giving moisture prediction MAE = 2.3 %, and days to MC = 22 % MAE of 7 days. Second best was ∑Days, with very similar moisture prediction error metrics to radiation. We noted that the ∑Days predictor also has the advantage of being very simple to understand and apply, whereas solar radiation is less familiar and not available in many weather observation and forecast products. Because of this, we selected ∑Days for further investigations. Other predictors (VPD, GDD and ETo) yielded poorer errors.

Experiment set 2 (Table 3) adds the field-specific variables. Adding variety (short, medium and long grain) to ∑Days reduced MC MAE from 2.3 % to 2.2 %, and the MC = 22 % MAE by more than 1 day (MAE 7.3. days). These and following results indicate a small reduction in moisture prediction errors leads to a significant reduction in the days to MC = 22 % prediction error, which is due to the dry-down curve having a low slope (on the order of 0.5 %/day, Fig. 2).

Of the nitrogen and biomass variables, adding applied N (NA) led to models with the lowest errors, which was slightly better than NU. Both of these N variables were better features than DM. We selected NU for further investigations, as this parameter is often measured by growers to guide mid-season N application (Dunn, 2008), and is expected to describe some of the variability due to residual soil N, which could lead to uncertainty in model predictions based on NA.

Experiment set 3 investigated other nonlinear algorithms, to

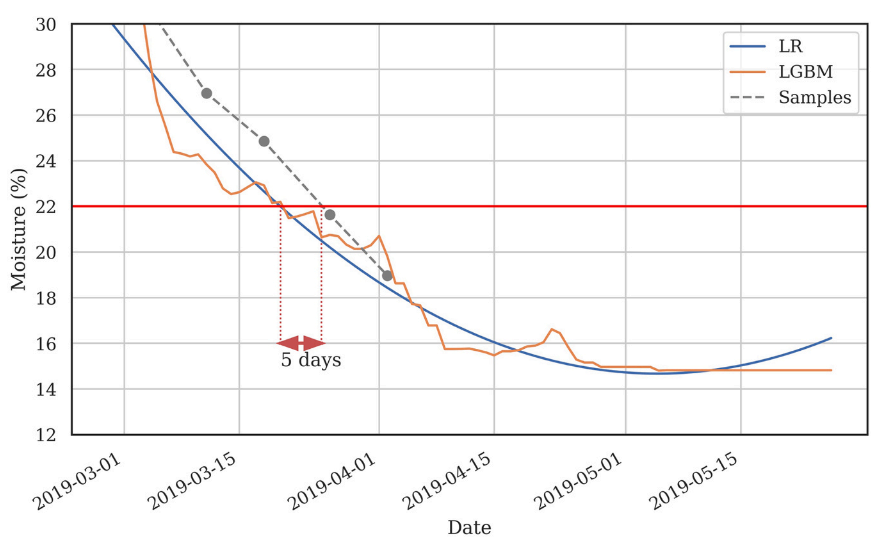

Fig. 2. Predicted and sampled moisture for a plot from the 2019 YAI experiment, illustrating how the error (5 days) between actual and predicted MC = 22 % date is calculated. Predictions from 2 models are shown: a linear regression (LR) model with features [∑Days, $\sum Days^2$, PI N, Variety], and a Light Gradient Boosting Machine (LGBM) model with features [∑Days, PI N, Variety], both trained on independent 2020–2021 data.

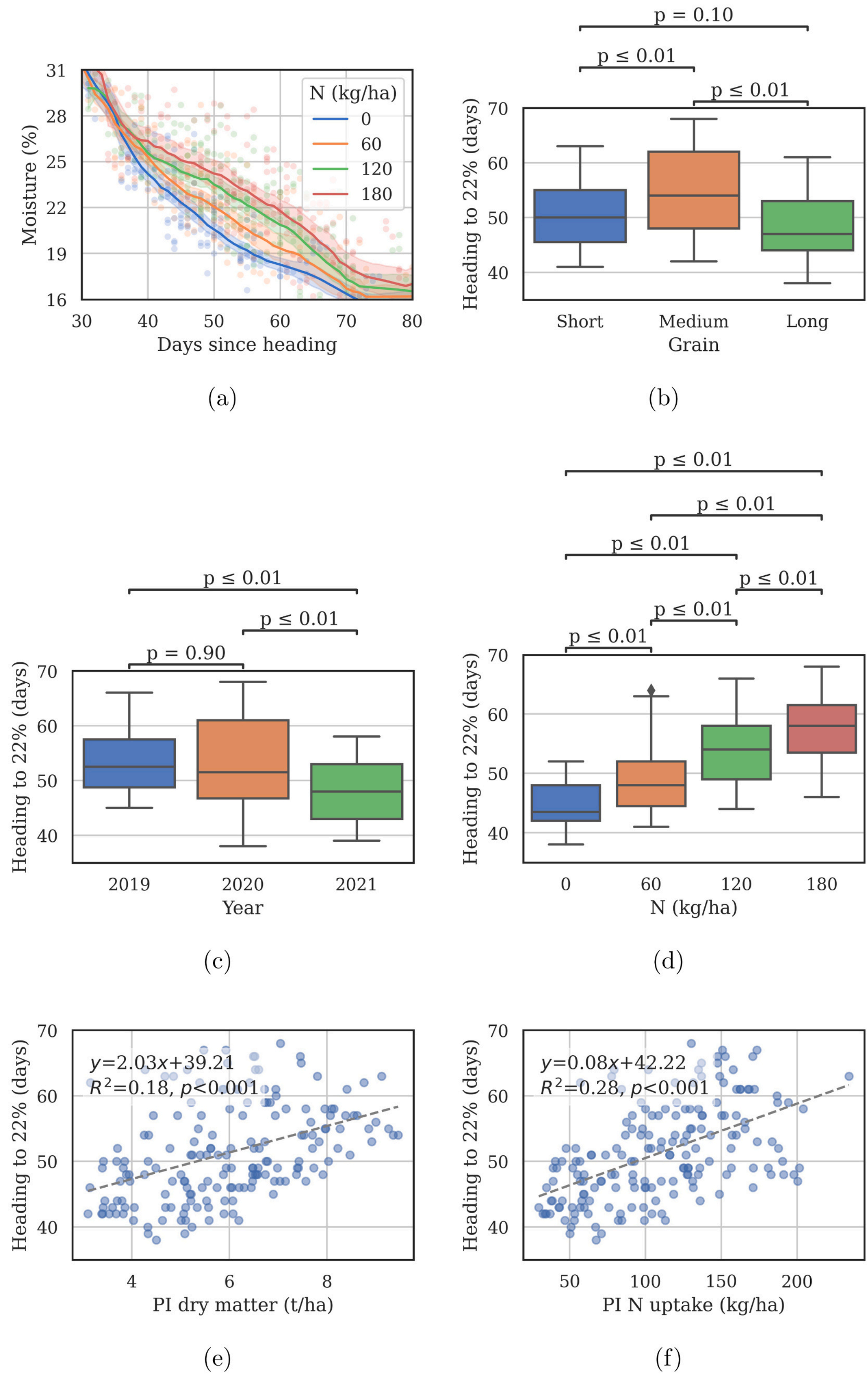

Fig. 3. (a) Dry down curves, averaged per day, with shaded areas showing the 95 % confidence interval. The time between heading and MC = 22 % as a function of (b) applied nitrogen rate, (c) variety, (d) year, (e) dry matter at PI and (f) nitrogen uptake at PI. Box plots show p-values from the Tukey HSD tests.

Table 3
Table of cross-validated model results, LR = Linear Regression, LGBM = Light Gradient Boosting Machine, RF = Random Forest, SVR = Support Vector Regression. Bold numbers and darker shading indicate higher accuracy.

Experiment set	Features	Algorithm	Moisture prediction MAE (%)	Moisture prediction LCCC	Days to 22% prediction MAE (Days)	Days to 22% prediction LCCC
1	ΣETo	LGBM	2.59	0.71	9.08	0.41
	ΣGDD		2.61	0.7	8.56	0.47
	ΣSolarRadiation		2.25	0.79	7.02	0.5
	ΣVPD		3.50	0.47	23.45	0.13
	ΣDays		2.28	0.78	8.56	0.27
2	ΣDays, Variety	LGBM	2.20	0.8	7.25	0.45
	ΣDays, NA		2.26	0.79	**5.72**	**0.71**
	ΣDays, DM		2.33	0.77	7.53	0.47
	ΣDays, NU		2.27	0.79	6.1	0.65
3	ΣDays, NU, Variety	LGBM	2.12	0.81	5.73	**0.71**
		RF	2.22	0.81	5.95	0.7
		SVR	2.37	0.77	8.87	0.32
4	ΣDays, $\Sigma Days^2$	LR	2.29	0.79	6.79	0.43
	ΣDays, NU, Variety		2.37	0.76	6.42	0.64
	ΣDays, $\Sigma Days^2$, NU, Variety		**2.06**	**0.83**	**5.72**	0.68
	ΣDays, $\Sigma Days^2$, $\Sigma Days^3$, NU, Variety		2.11	0.82	5.76	0.67

compare with LGBM, with the [∑Days, NU, Variety] variable set. LGBM gave the lowest errors compared with RF and SVR.

Noting that experiment set 1 showed cumulative radiation was a better predictor than days when only a single predictor is used, we also trained LGBM models using predictors [∑SolarRadiation, NU, Variety] to compare with experiment set 3, which used ∑Days. In this case, the predictions with ∑SolarRadiation were slightly worse (MAE = 2.2 % and 5.8 days) than using the simpler ∑Days, indicating the predictive advantage of ∑SolarRadiation is not clear when field-specific variables are included.

Finally, experiment set 4 investigated using linear regression (LR), which has the advantage of providing easily interpretable predictions in contrast to the other algorithms. It also produces smooth predictions vs time, which is not the case for tree-based models such as LGBM (Fig. 2), because the latter model predictions are based on ensembling discrete prediction values. The LR model with [∑Days, NU, Variety] (2nd row, Experiment 4, Table 3) produced higher errors than the nonlinear models. However, when $\sum Days^2$ was added, the predictions were competitive or better than the nonlinear model predictions. There was no advantage to adding a third-order polynomial term ($\sum Days^3$).

The prediction accuracy of the LR models with and without the field specific variables (NU and Variety) are compared in Fig. 4, using the cross-validation strategy. Adding the field-specific variables to the models improves moisture prediction MAE from 2.3 % to 2.1 %, and improves MC = 22 % predictions by more than a day, from 6.8 to 5.7 days.

3.3. Final model

We determined the following from the above experiments:

- Linear regression gave good accuracy relative to more complex and less interpretable algorithms, provided a second-order polynomial of the cumulative predictor was used to account for nonlinear dry-down vs time.
- The simple ∑Days predictor was similar or better than weather-based cumulative predictors.
- Variety did improve model predictions.
- Predictions using NU were not quite as good as NA, but are likely to generalize better to new sites and seasons, due to uncertainty about pre-existing soil N.

Therefore, we trained a LR model on all data (n = 911) using variable set [∑Days, $\sum Days^2$, NU, Variety]. The model performance and parameters are given in Table 4.

For example, for the medium grain variety, the equation is:

$$\mathrm{MC}(\%) = 45.8 + 0.016 \times \mathrm{NU} - 0.65 \times \sum \mathrm{Days} + 0.0034 \times \sum \mathrm{Days}^2 \quad (2)$$

This quadratic equation can be solved to find the number of days from heading to any moisture content. For the medium grain variety with a typical panicle initiation N uptake of 100 kg/ha (Dunn et al., 2016), the model predicts 55 days from heading to MC = 22 % (Fig. 5).

Eq. (2) can be differentiated to find the dry down rate:

$$\frac{d\mathrm{MC}}{d\sum \mathrm{Days}} = 0.0068 \times \sum \mathrm{Days} - 0.65 \quad (3)$$

This describes the expected characteristic, that dry-down rate slows as moisture reduces towards equilibrium (Martinez-Feria et al., 2019). At 30 days after heading, the dry-down rate is − 0.45 %/day. At 55 days it is − 0.28%/day.

Predicted dry-down for the different varieties and a range of NU are shown in Fig. 5. Eq. (2) indicates that for every 50 kg/ha increase in N uptake, there is a 0.8% increase in MC. This increases time to harvest moisture, as the experimental data also showed (Fig. 3d). The medium grain variety was predicted to have more than 1 % higher grain moisture than the other varieties (Table 4), which is reflected in the slower time to harvest for this variety in the experimental data (Fig. 3b).

4. Discussion

We developed predictive and descriptive models for rice MC dry-down in the field for long, medium and short grain rice varieties. The models were able to describe much of the variability in the moisture samples (R^2 = 0.8). Incorporating factors related to field variability improves model predictions and provides understanding of how these factors affect time to harvest. Variety, applied nitrogen, mid-season nitrogen uptake and dry matter were found to be important drivers of dry-down variability. Adding field-specific variables such as nitrogen and variety improved the average prediction error of days from heading to harvest moisture by around a day (from MAE = 6.8–5.7 days for the linear models).

Cumulative solar radiation was a good predictor of moisture dry-

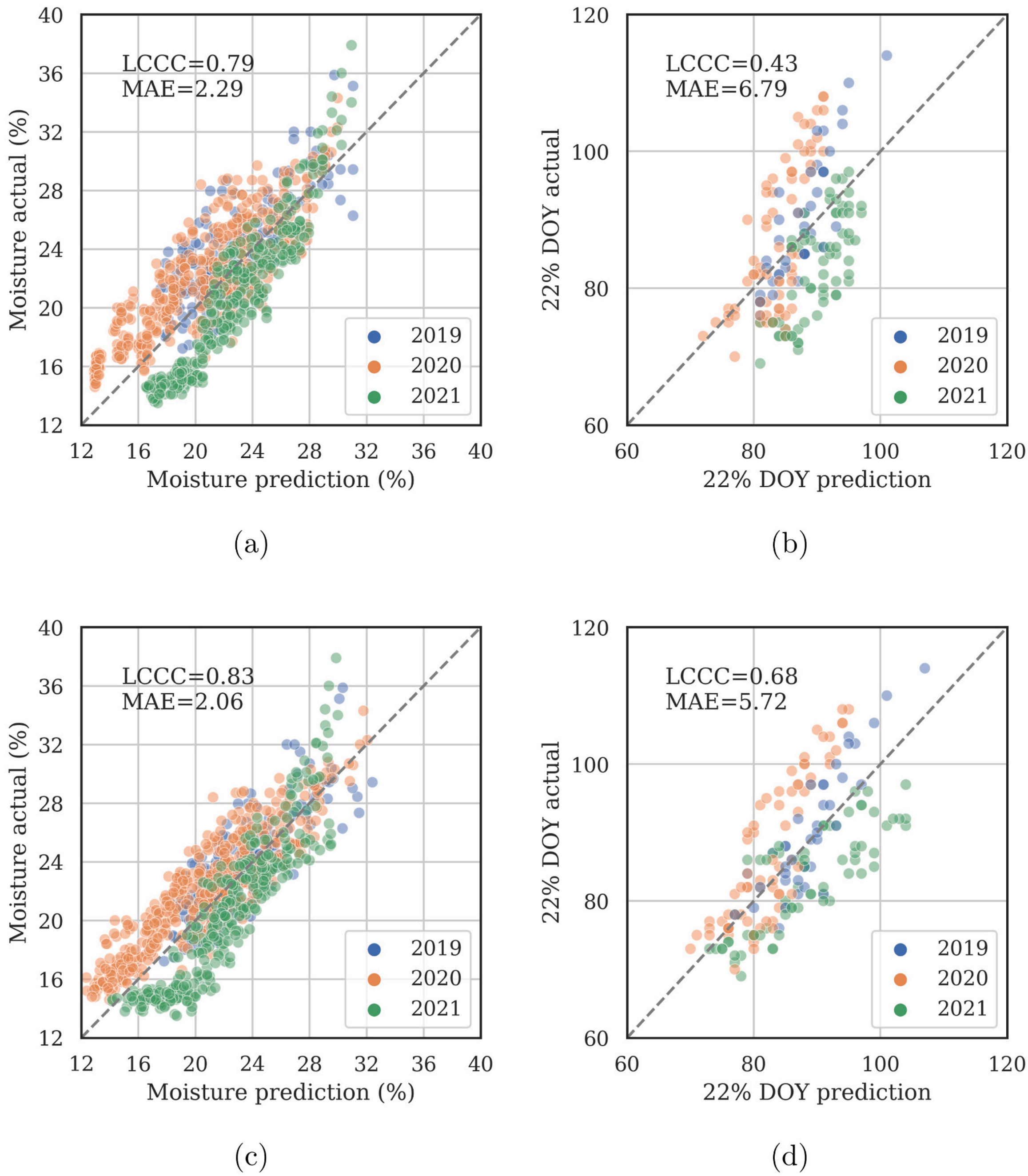

Fig. 4. Cross-validation actual vs predicted moisture for the 911 samples (left) and day of year (DOY) of MC = 22 % (right). Predictions are from linear regression models with input features being (a–b) [$\sum$Days, $\sum \text{Days}^2$], and (c–d) [$\sum$Days, $\sum \text{Days}^2$, NU, Variety].

Table 4
Linear regression model trained on all 911 samples. R^2 was 0.80, and $p < 0.001$. The p-value of all model coefficients were < 0.001. Short, Medium and Long are the intercept, specific to each rice variety. NU is N uptake around panicle initiation (kg/ha). $\sum$Days is the number of days between heading and the date MC is being predicted for.

Predictor	Coefficient	95 % confidence interval
Short	44.6	43.1–46.1
Medium	45.8	44.3–47.3
Long	43.8	42.3–45.2
NU	0.016	0.013–0.019
$\sum$Days	− 0.65	− 0.70 to − 0.60
$\sum \text{Days}^2$	0.0034	0.0030–0.0038

down, and degree days was not. However, we found that cumulative days since heading was almost as good a predictor of grain MC as any cumulative weather variables. Similarly, McCauley and Way (2002) found that weather did not explain the variability in rice dry-down between years, and that the effect of rainfall on MC was not consistent. Also, while Martinez-Feria et al. (2019) included temperature and humidity in the equilibrium moisture component of their model, they similarly found little advantage to using weather variables over cumulative days in their dry-down model component. However, more recent work on maize dry-down (Chazarreta et al., 2023) showed that when a wider range of sowing dates are included, weather variables do have a significant effect on dry-down parameters. Similarly, it is possible that if more years of data were available for this work, the cumulative weather variables could have been more important, as the dataset would

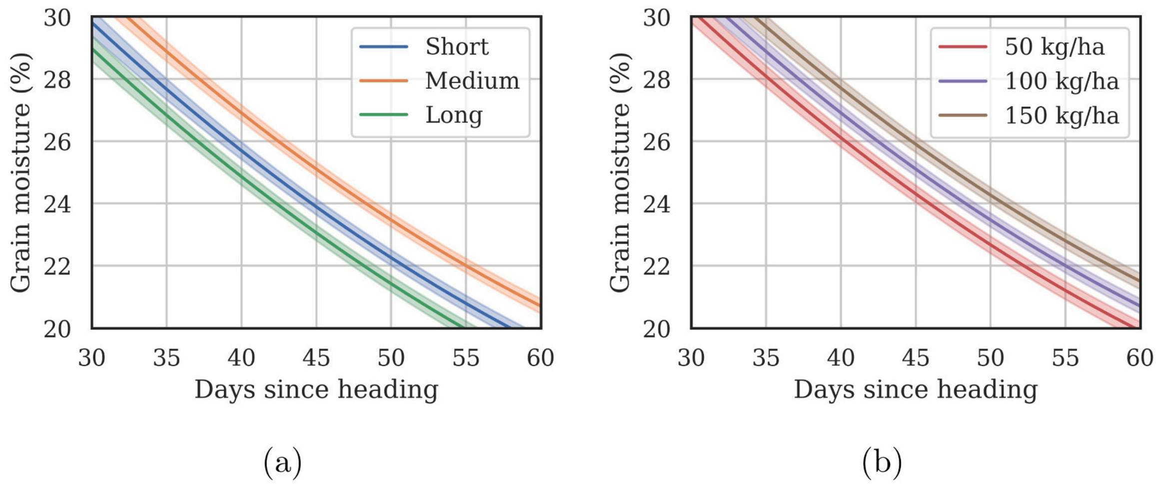

Fig. 5. Predicted dry-down curves using the linear regression model with coefficients in Table 4. The shaded areas are the 95 % confidence intervals. (a) For three varieties, with N uptake set to 100 kg/ha. (b) For three N uptake values, and the medium grain variety.

encompass a wider range of weather variability. Along similar lines, Lobell and Burke (2010) found a dataset with a sufficiently wide range of weather variability from spatial and seasonal heterogeneity was crucial to accurately quantify how weather affects yield. Until sufficient data is available to prove the advantage of using weather predictor variables, the decision to use cumulative days as the main predictor of dry-down has the advantage of producing models that are simple to apply and interpret, while maintaining accuracy.

Higher nitrogen rates and biomass were associated with higher moisture content, and longer time to reach harvest moisture. This is similar to the trend found for corn (Zhang et al., 2021). Similarly, Xu et al. (2022a) found larger leaf area (which may be caused by higher nitrogen) was associated with higher moisture content in maize.

We found the medium grain variety (Reiziq) had higher moisture content and longer time to reach harvest maturity. It is possible this could be explained by the larger grain volume and higher grain weight of this variety compared with the short and long varieties (29 vs 23 mg, see the Section 2). However experiments with more varieties and a range of grain sizes would be needed to confirm this.

Machine learning algorithms were able to predict the nonlinear dry-down of MC with respect to time (Fig. 2). The LightGBM algorithm has produced state-of-the-art performance on tabular and forecasting problems, such as the one in this work (Ke et al., 2017; Makridakis et al., 2022), and we found it produced the best results among the machine learning algorithms we tried. However, when we added a nonlinear predictor (the square of the days since heading), linear regression produced similar accuracy. Linear regression had the benefit of producing smooth predictions vs time, whereas the tree-based models tended to produce non-smooth predictions because of their discrete nature (Fig. 2). Linear regression also gave the advantage of providing a descriptive model, that can be interrogated to determine factors such as the time to reach a specific moisture content, expected dry-down rates, and how these change with variety and nitrogen.

We used indicators of N and biomass obtained relatively early in the season, at the panicle initiation growth phase. Potentially, samples taken later in the season, closer to maturity, may improve quantification of the effects of these parameters on dry-down. However, samples are often taken at PI rather than later in the season due to the need to topdress at that time to address nitrogen deficiencies (Wilson et al., 1998; Fageria and Baligar, 1999; Dunn et al., 2016), making our choice a practical one due to data availability.

Our model does not explain all of the variability in grain moisture, and variability in the time to reach 22 % MC. Some of the remaining variability is due to inherent errors in sample gathering, processing and uneven distribution of grain moisture contents (Kocher et al., 1990). Some may be caused by additional in- and between-field variability caused by other factors that have not been incorporated in the model (for example soil, other nutrient limitations and water management). Remote sensing is able to describe nitrogen status (Brinkhoff et al., 2019), biomass (Xu et al., 2022b) and vegetation water content (Yilmaz et al., 2008). Previous work has demonstrated estimation of phenological stage (Yang et al., 2020), optimal harvest timing (Meng et al., 2015) and prediction of grain moisture content at the time an image was acquired (Dunn and Dunn, 2021; Sarkar et al., 2018). Therefore, remote sensing potentially offers the opportunity to capture field variability that may be able to further reduce some of our model's uncertainty.

There are several factors that influence the decisions regarding when to drain rice fields and when to start harvest. These include the soil type, weather forecast and the requirement to ensure field trafficability for harvesters (Dingkuhn and Le Gal, 1996). The purpose of this study was to present data and models that contribute to understanding another crucial factor: predicting grain moisture content and thus optimal harvest date. Specifically, the results have shown how field-specific conditions such as variety, nitrogen status and biomass impact these decisions.

5. Conclusion

This work developed predictive and descriptive models of rice grain moisture content dry-down. It provides understanding of how field-specific parameters such as rice variety and nitrogen affect dry-down and time to optimal harvest moisture content. Higher biomass and nitrogen were associated with longer times to reach harvest moisture. Models were able to predict time to harvest moisture with an average of less than 6 days error (using a robust leave-one-year-out cross-validation train/test methodology). We anticipate combining remote sensing with such dry-down models will further improve accuracy through quantifying field-specific parameters, with the aim of providing grain moisture forecast tools for growers to aid field drainage and harvest timing decisions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data that has been used is confidential.

Acknowledgements

This work was funded by AgriFutures Australia, Grants PRO-013078 (Real-time remote-sensing based monitoring for the rice industry) and PRJ-009790 (Rice variety nitrogen and agronomic management).

References

AACC, 1999. AAAC International Method 44-15.02. Moisture–Air-Oven Methods. Reapproval November 3, 1999. 〈https://doi.org/10.1094/AACCIntMethod-44-15.02〉.
Brinkhoff, J., Dunn, B.W., Robson, A.J., Dunn, T.S., Dehaan, R.L., 2019. Modeling mid-season rice nitrogen uptake using multispectral satellite data. Remote Sens. 11, 1837. https://doi.org/10.3390/rs11151837.
Brinkhoff, J., Dunn, B.W., Robson, A.J., 2021. Rice nitrogen status detection using commercial-scale imagery. Int. J. Appl. Earth Obs. Geoinf. 105, 102627 https://doi.org/10.1016/j.jag.2021.102627.
Brinkhoff, J., Houborg, R., Dunn, B.W., 2022. Rice ponding date detection in Australia using Sentinel-2 and Planet Fusion imagery. Agric. Water Manag. 273, 107907 https://doi.org/10.1016/j.agwat.2022.107907.
Calderwood, D.L., Bollich, C.N., Scott, J.E., 1980. Field drying of rough rice: effect on grain yield, milling quality, and energy saved. Agron. J. 72, 649–653. https://doi.org/10.2134/agronj1980.00021962007200040019x.
Chazarreta, Y.D., Carcedo, A.J.P., AlvarezPrado, S., Massigoge, I., Amas, J.I., Fernandez, J.A., Ciampitti, I.A., Otegui, M.E., 2023. Enhancing maize grain dry-down predictive models. Agric. For. Meteorol. 334, 109427 https://doi.org/10.1016/j.agrformet.2023.109427.
Darbyshire, R., Crean, E., Dunn, T., Dunn, B., 2019. Predicting panicle initiation timing in rice grown using water efficient systems. Field Crops Res. 239, 159–164. https://doi.org/10.1016/j.fcr.2019.05.018.
Dingkuhn, M., Le Gal, P.Y., 1996. Effect of drainage date on yield and dry matter partitioning in irrigated rice. Field Crops Res. 46, 117–126. https://doi.org/10.1016/0378-4290(95)00094-1.
Dunn, B., 2008. Improving topdressing recommendations for rice. 〈https://agrifuturesrice.squarespace.com/s/Improving-topdressing-recommendations-for-rice.pdf〉.
Dunn, B., Dunn, T., 2021. Predicting rice crop maturity using remote sensing. In: NSW Department of Primary Industries Southern NSW Research Results, pp. 137–9. 〈https://www.dpi.nsw.gov.au/_data/assets/pdf_file/0009/1365192/SRR21-book-web-cm29Oct2021.pdf〉.
Dunn, B., Ford, R., 2018. Management of drill sown rice. Primefact 1253, third edition. NSW Department of Primary Industries.
Dunn, B.W., Dunn, T.S., Orchard, B.A., 2016. Nitrogen rate and timing effects on growth and yield of drill-sown rice. Crop Pasture Sci. 67, 1149–1157. https://doi.org/10.1071/CP16331.
Fageria, N.K., Baligar, V.C., 1999. Yield and yield components of lowland rice as influenced by timing of nitrogen fertilization. J. Plant Nutr. 22, 23–32. https://doi.org/10.1080/01904169909365603.
Humphreys, E., Lewin, L.G., Khan, S., Beecher, H.G., Lacy, J.M., Thompson, J.A., Batten, G.D., Brown, A., Russell, C.A., Christen, E.W., Dunn, B.W., 2006. Integration of approaches to increasing water use efficiency in rice-based systems in southeast Australia. Field Crops Res. 97, 19–33. https://doi.org/10.1016/j.fcr.2005.08.020.
Inoue, Y., Sakaiya, E., Zhu, Y., Takahashi, W., 2012. Diagnostic mapping of canopy nitrogen content in rice based on hyperspectral measurements. Remote Sens. Environ. 126, 210–221. https://doi.org/10.1016/j.rse.2012.08.026.
Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ. Model. Softw. 16, 309–330. https://doi.org/10.1016/S1364-8152(01)00008-1.
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. In: Advances in Neural Information Processing Systems. Curran Associates, Inc. 〈https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html〉.
Kocher, M.F., Siebenmorgen, T.J., Norman, R.J., Wells, B.R., 1990. Rice kernel moisture content variation at harvest. Trans. ASAE 33, 0541–0548. https://doi.org/10.13031/2013.31363.
Lee, J.H., Kang, C.S., Roh, A.S., Park, K.Y., Lee, H.J., 2009. Assessment of N topdressing rate at panicle initiation stage with chlorophyll meter-based diagnosis in rice. J. Crop Sci. Biotechnol. 12, 195–200. https://doi.org/10.1007/s12892-009-0133-5.
Lin, L., He, Y., Xiao, Z., Zhao, K., Dong, T., Nie, P., 2019. Rapid-detection sensor for rice grain moisture based on NIR spectroscopy. Appl. Sci. 9, 1654. https://doi.org/10.3390/app9081654.
Lin, L.I.K., 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics 45, 255–268. https://doi.org/10.2307/2532051.
Lobell, D.B., Burke, M.B., 2010. On the use of statistical models to predict crop yield responses to climate change. Agric. For. Meteorol. vol. 150, pp. 1443–52. 〈https://doi.org/10.1016/j.agrformet.2010.07.008〉.
Lu, R., Siebenmorgen, T.J., 1994. Modeling rice field moisture content during the harvest season – Part I model development. Trans. ASAE 37, 545–551. https://doi.org/10.13031/2013.28111.
Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2022. M5 accuracy competition: results, findings, and conclusions. Int. J. Forecast. https://doi.org/10.1016/j.ijforecast.2021.11.013.
Martinez-Feria, R.A., Licht, M.A., Ordóñez, R.A., Hatfield, J.L., Coulter, J.A., Archontoulis, S.V., 2019. Evaluating maize and soybean grain dry-down in the field with predictive algorithms and genotype-by-environment analysis. Sci. Rep. 9, 7167. https://doi.org/10.1038/s41598-019-43653-1.
McCauley, G.N., Way, M.O., 2002. Drain and harvest timing affects on rice grain drying and whole-milled grain. Field Crops Res. 74, 163–172. https://doi.org/10.1016/S0378-4290(01)00205-2.
McCormick, R.F., Truong, S.K., Rotundo, J., Gaspar, A.P., Kyle, D., van Eeuwijk, F., Messina, C.D., 2021. Intercontinental prediction of soybean phenology via hybrid ensemble of knowledge-based and data-driven models. Silico Plants 3, diab004. https://doi.org/10.1093/insilicoplants/diab004.
Meng, J., Xu, J., You, X., 2015. Optimizing soybean harvest date using HJ-1 satellite imagery. Precis. Agric. 16, 164–179. https://doi.org/10.1007/s11119-014-9368-3.
Miao, Y., Mulla, D.J., Robert, P.C., 2006. Spatial variability of soil properties, corn quality and yield in two Illinois, USA fields: implications for precision corn management. Precis. Agric. 7, 5–20. https://doi.org/10.1007/s11119-005-6786-2.
Muthayya, S., Sugimoto, J.D., Montgomery, S., Maberly, G.F., 2014. An overview of global rice production, supply, trade, and consumption. Ann. N. Y. Acad. Sci. 1324, 7–14. https://doi.org/10.1111/nyas.12540.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.
Peng, S., Cassman, K.G., 1998. Upper threshholds of nitrogen uptake rates and associated nitrogen fertilizer efficiencies in irrigated rice. Agron. J. 90, 178–185. https://doi.org/10.2134/agronj1998.00021962009000020010x.
Rajanna, B., Andrews, C.H., 1970. Trends in seed maturation of rice (Oryza Sativa L.). Proc. Assoc. Off. Seed Anal. 60, 188–196.
Sarkar, T.K., Ryu, C.S., Kang, J.G., Kang, Y.S., Jun, S.R., Jang, S.H., Park, J.W., Song, H.Y., 2018. Artificial neural network-based model for predicting moisture content in rice using UAV remote sensing data. Korean J. Remote Sens. 34, 611–624. https://doi.org/10.7780/kjrs.2018.34.4.4.
Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and Statistical Modeling with Python. 6.
Sharifi, H., Hijmans, R.J., Hill, J.E., Linquist, B.A., 2017. Using stage-dependent temperature parameters to improve phenological model prediction accuracy in rice models. Crop Sci. 57, 444–453. https://doi.org/10.2135/cropsci2016.01.0072.
Sharifi, H., Hijmans, R.J., Hill, J.E., Linquist, B.A., 2018. Water and air temperature impacts on rice (Oryza sativa) phenology. Paddy Water Environ. 16, 467–476. https://doi.org/10.1007/s10333-018-0640-4.
Siebenmorgen, T.J., Bautista, R.C., Counce, P.A., 2007. Optimal harvest moisture contents for maximizing milling quality of long- and medium-grain rice cultivars. Appl. Eng. Agric. 23, 517–527. https://doi.org/10.13031/2013.23476.
Troldahl, D., Snell, P., Pallas, L., Ovenden, B., 2014. Rice variety guide 2014–15. Primefact 1112, 4th edition. NSW Department of Primary Industries.
Wang, J., Sun, X., Xu, Y., Wang, Q., Tang, H., Zhou, W., 2021. The effect of harvest date on yield loss of long and short-grain rice cultivars (Oryza sativa L.) in Northeast China. Eur. J. Agron. 131, 126382 https://doi.org/10.1016/j.eja.2021.126382.
Ward, J.K., Henry, W.B., Hock, M.W., 2016. Variability in harvest moisture and dry-down in multi-hybrid planting systems. Trans. ASABE 59, 1111–1115. https://doi.org/10.13031/trans.59.11572.
Ward, R., Brickhill, H., Bull, N., Dunn, B., Dunn, T., Fowler, J., Hart, J., Mauger, T., 2021. Rice Growing Guide 2021. NSW Department of Primary Industries. 〈https://www.dpi.nsw.gov.au/_data/assets/pdf_file/0004/1361173/RGG-2021-web-final-26Oct2021.pdf〉.
Wilson Jr., C.E., Bollich, P.K., Norman, R.J., 1998. Nitrogen application timing effects on nitrogen efficiency of dry-seeded rice. Soil Sci. Soc. Am. J. 62, 959–964. https://doi.org/10.2136/sssaj1998.03615995006200040016x.
Xu, C.c., Zhang, P., Wang, Y.y., Luo, N., Tian, B.j., Liu, X.w., Wang, P., Huang, S.b., 2022a. Grain yield and grain moisture associations with leaf, stem and root characteristics in maize. J. Integr. Agric. 21, 1941–1951. https://doi.org/10.1016/S2095-3119(20)63598-5.
Xu, J., Meng, J., Quackenbush, L.J., 2019. Use of remote sensing to predict the optimal harvest date of corn. Field Crops Res. 236, 1–13. https://doi.org/10.1016/j.fcr.2019.03.003.
Xu, L., Zhou, L., Meng, R., Zhao, F., Lv, Z., Xu, B., Zeng, L., Yu, X., Peng, S., 2022b. An improved approach to estimate ratoon rice aboveground biomass by integrating UAV-based spectral, textural and structural features. Precis. Agric. https://doi.org/10.1007/s11119-022-09884-5.
Yang, M.D., Hsu, Y.C., Tseng, W.C., Lu, C.Y., Yang, C.Y., Lai, M.H., Wu, D.H., 2021. Assessment of grain harvest moisture content using machine learning on smartphone images for optimal harvest timing. Sensors 21, 5875. https://doi.org/10.3390/s21175875.
Yang, Q., Shi, L., Han, J., Yu, J., Huang, K., 2020. A near real-time deep learning approach for detecting rice phenology based on UAV images. Agric. For. Meteorol. 287, 107938 https://doi.org/10.1016/j.agrformet.2020.107938.
Yilmaz, M.T., Hunt, E.R., Jackson, T.J., 2008. Remote sensing of vegetation water content from equivalent water thickness using satellite imagery. Remote Sens. Environ. 112, 2514–2522. https://doi.org/10.1016/j.rse.2007.11.014.
Zhang, Y.m., Xue, J., Zhai, J., Zhang, G.q., Zhang, W.x., Wang, K.r., Ming, B., Hou, P., Xie, R.z., Liu, C.w., Li, S.k., 2021. Does nitrogen application rate affect the moisture content of corn grains? J. Integr. Agric. 20, 2627–2638. https://doi.org/10.1016/S2095-3119(20)63401-3.