AgriEngineering

MDPI

Article

Hierarchical Spectral Modelling of Pasture Nutrition: From Laboratory to Sentinel-2 via UAV Hyperspectral

Jason Barnetson [1,*], Hemant Raj Pandeya [2] and Grant Fraser [1]

[1] Queensland Department of Environment, Tourism, Science and Innovation, Eco-Sciences Precinct, 41 Boggo Rd., Dutton Park, Brisbane, QLD 4102, Australia; grant.fraser@detsi.qld.gov.au
[2] Queensland Department of Primary Industries, Eco-Sciences Precinct, 41 Boggo Rd., Dutton Park, Brisbane, QLD 4102, Australia; rajhemant.pandeya@dpi.qld.gov.au
* Correspondence: jason.barnetson@detsi.qld.gov.au

Abstract

This study demonstrates a hierarchical spectral modelling approach for predicting pasture nutrition metrics using TabPFN (Tabular Prior-Data Fitted Network), a transformer-based machine learning architecture. In the face of climate variability, aligning stocking rates with pasture resources is crucial for sustainable livestock grazing, requiring accurate assessments of both pasture biomass and nutrient composition. Our research, conducted across diverse growth stages at five tropical and subtropical savanna rangeland properties in Queensland, Australia, with native and introduced C4 grasses, employed a hierarchical sampling and modelling strategy that scales from laboratory spectroscopy to Sentinel-2 satellite predictions via uncrewed aerial vehicle (UAV) hyperspectral imaging. Spectral data were collected from leaf (laboratory spectroscopy) through field (point measurements), UAV hyperspectral imaging, and Sentinel-2 satellite imagery. Traditional laboratory wet chemistry methods determined plant leaf and stem nutrient content, from which crude protein (CP = total nitrogen (TN) $\times$ 6.25) and dry matter digestibility (DMD = 88.9–0.779 $\times$ acid detergent fibre (ADF)) were derived. TabPFN models were trained at each spatial scale, achieving validation R^2 of 0.76 for crude protein at the leaf scale, 0.95 at the UAV scale, and 0.92 at the Sentinel-2 satellite scale. For dry matter digestibility, validation R^2 was 0.88 at the UAV scale and 0.73 at the Sentinel-2 scale. A pasture classification masking approach using a deep neural network with 98.6% accuracy (7 classes) was implemented to focus predictions on productive pasture areas, excluding bare soil and woody vegetation. The Sentinel-2 models were trained on 462 samples from 19 site–date combinations across 11 field sites. The TabPFN architecture provided notable advantages over traditional neural networks: no hyperparameter tuning required, faster training, and superior generalisation from limited training samples. These results demonstrate the potential for accurate and efficient prediction and mapping of pasture quality across large areas (100 s–1000 s km^2) using freely available satellite imagery and open-source machine learning frameworks.

Keywords: TabPFN; transformer; machine learning; hyperspectral; multi-spectral; Sentinel-2; pasture nutrition; crude protein; dry matter digestibility; remote sensing; rangeland monitoring; hierarchical modelling

Academic Editors: Wei Xue and Ruyin Cao

Received: 18 February 2026
Revised: 14 March 2026
Accepted: 1 April 2026
Published: 7 April 2026

1. Introduction

Global livestock production for meat and milk faces persistent challenges [1]. Effectively aligning stocking rates with available pasture resources, while ensuring sustainability

and profitability, demands innovative approaches [2]. This need is driven by increasing global demand for animal protein, heightened climatic variability, and growing consumer and governmental emphasis on environmental stewardship. Informed herd management hinges on a comprehensive understanding of pasture resources, particularly their nutritional status. Stocking rates must be precisely matched to long-term pasture availability to maintain livestock health, growth, and enterprise profitability [3]. Strategic adjustments to pasture resources, such as fertilization and soil management, rely on accurate assessments of soil and plant nutrient composition [1]. Traditionally, plant leaf nutrient status has been determined through costly and spatially limited chemical analyses.

This research aims to advance efficient and spatially explicit alternatives to traditional plant leaf chemical analysis, enhancing the confidence in pasture nutrient assessments. As highlighted by [1], spectroscopy and machine learning are promising tools for predicting pasture nutrient status. Spectroscopy, the study of light energy capture across the electromagnetic spectrum (ER) [1,4], particularly the visible, near-infrared (NIR), and short-wave infrared (SWIR) regions, is crucial for vegetation analysis. Spectrophotometry further quantifies light absorption by chemical substances [5], enabling the determination of specific leaf chemical compositions based on known absorption features [6]. Absorption features in the red-edge and NIR regions correlate strongly with plant biochemical compounds and pigments, such as leaf nitrogen and crude protein, vital for photosynthesis and plant health [7]. The SWIR region, less influenced by pigment absorption, is more sensitive to plant fibre compounds (cellulose, hemicellulose, lignin) and water [8]. Understanding forage digestibility, which is directly linked to plant chemical components, is essential for sustainable livestock management.

Recent studies have demonstrated the viability of multi-scale spectral approaches for pasture and vegetation quality assessment. At the UAV scale, [9] achieved successful prediction of crude protein and acid detergent fibre in temperate grasslands using imaging spectroscopy, while [10] predicted crude protein, digestibility, lignin and cellulose in temperate pastures using airborne hyperspectral imagery. At the satellite scale, Reference [11] demonstrated Sentinel-2-based estimation of pasture crude protein and neutral detergent fibre content, and [12] mapped leaf nitrogen distribution in semi-arid environments using field spectroscopy models applied to Sentinel-2 imagery. Reference [13] established a hierarchical model-based inference framework combining UAV and Sentinel-2 data for forest biomass estimation, providing methodological precedent for multi-scale approaches. More recently, Reference [14] developed a physically informed multi-scale deep neural network for estimating foliar nitrogen concentration, demonstrating the potential of deep learning architectures for cross-scale vegetation chemistry prediction. Reference [15] further demonstrated effective prediction of Mediterranean pasture quality using field spectroscopy with random forest and partial least squares regression. These studies collectively establish the technical foundation for spectroscopy-based pasture nutrition monitoring, though few have attempted the full hierarchical integration from laboratory to satellite scales pursued in the present work.

Machine learning (ML) models are increasingly recognized for their potential in analyzing spectral data and plant leaf chemical features [16]. As highlighted by [1], ML models are crucial for developing, evaluating, and refining methods to predict plant leaf chemical compounds from reflectance spectra. While traditional statistical models like partial least squares regression and random forest have been widely used, deep learning has emerged as a dominant approach. Recent advances in transformer-based architectures, particularly TabPFN (Tabular Prior-Data Fitted Network) [17], have demonstrated exceptional performance on tabular regression tasks with small-to-medium sized datasets typical of remote sensing calibration studies. TabPFN leverages in-context learning through pre-

training on synthetic tabular data, eliminating the need for hyperparameter tuning while achieving state-of-the-art performance. This non-linear pattern recognition is particularly valuable for investigating the intricate chemical and biophysical relationships associated with pasture nutrition.

The specific research objectives of this study are:

1. **Multi-scale spectral characterisation:** To characterise and compare spectral responses of pasture swards across multiple spatial scales (leaf, plant, UAV, and satellite) and identify optimal spectral features for predicting pasture crude protein and dry matter digestibility.
2. **TabPFN model development:** To develop and evaluate TabPFN-based predictive models for crude protein (CP) and dry matter digestibility (DMD) across all spatial scales, demonstrating the advantages of transformer-based architectures for remote sensing applications.
3. **Pasture classification for target purity:** To develop and integrate a pasture classification masking approach as an essential preprocessing step that ensures spectral target purity by isolating pasture pixels from woody vegetation and bare soil, a critical requirement for reliable chemistry retrieval in heterogeneous rangeland landscapes.
4. **Temporal dynamics assessment:** To assess the temporal dynamics of predicted pasture chemistry metrics across multiple field sites and seasonal cycles, validating the physiological responsiveness of model predictions to environmental conditions.
5. **Operational scalability demonstration:** To demonstrate the operational scalability of satellite-based pasture nutrition monitoring using freely available Sentinel-2 imagery and open-source machine learning frameworks.

By addressing these objectives, this research seeks to improve the efficiency, cost-effectiveness, and sustainability of rangeland grazing through the provision of more accurate, spatially and temporally explicit estimations of pasture nutrition. This study examined five representative locations in tropical and subtropical savanna and woodland environments over two growing seasons. Plant pasture chemical composition and spectral reflectance data were collected using destructive sampling and field, UAV, and satellite measurements. TabPFN machine learning models were employed within a hierarchical spectral modelling framework—from laboratory to Sentinel-2 via UAV hyperspectral—for key pasture nutrition metrics: crude protein (CP) and dry matter digestibility (DMD).

2. Materials and Methods

2.1. Study Area

Five cattle grazing properties were chosen for this case study, located within the Desert Uplands, Brigalow Belt, and Southeast Queensland bioregions. All five properties were located within the Great Barrier Reef (GBR) catchment of Queensland, Australia. Grazing (77%) is the dominant agricultural land use of the GBR [18] and contributes to increased levels of sediments and nutrients. Improved grazing land management practices contribute to reduced sedimentation and nutrient deposition.

The climate of the study area is a predominantly hot, humid summer zone for Property A (Baralaba region) and warm, humid summer zone for the more northerly and westerly properties (B, C, D) as illustrated in Figure 1. Property E is located in a subtropical climate zone in southeast Queensland. Each field site was chosen to represent the main productive vegetation groups of each property. Field sites were sampled up to four times during 2023–2024, capturing late wet (April–May 2023), dry (August–November 2023), wet (January–April 2024), and dry (August–October 2024) seasons to assess temporal variability in pasture nutrition.

Table 1 details the dominant vegetation groups sampled across each property and field site.

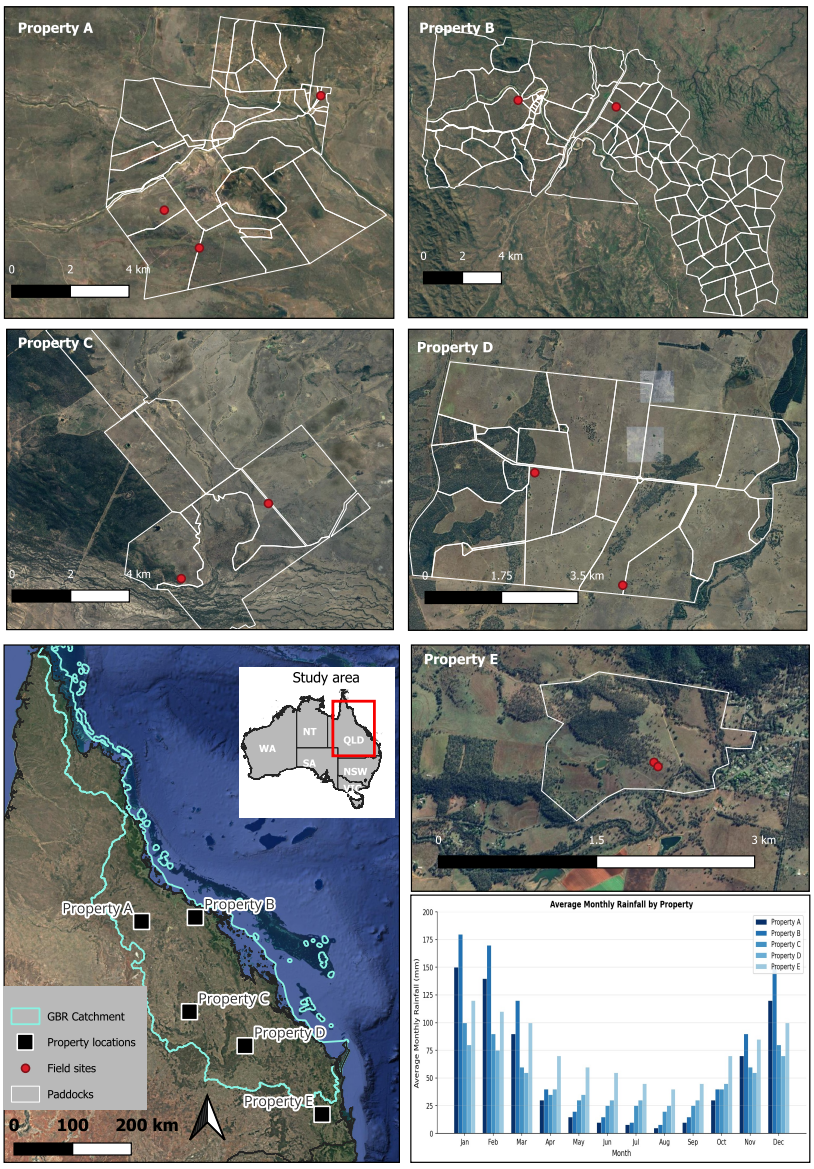

Figure 1. Locations across GBR catchment and bioregions. Property A–E locations illustrated. Red box indicates inset boundary. Background imagery: Google Earth satellite imagery (January 2026; © Google, Maxar Technologies, Westminster, CO, USA). Hillshade relief derived from Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global digital elevation data [19].

Table 2 details the sampling schedule for each field site, including coordinates and visit dates.

Table 1. Vegetation groups and dominant species sampled at each property.

Property	Vegetation Group	Dominant Species
Property A (Sites 1–3)	Grassland	Native/Naturalised grasses
Property B (Sites 1–2)	Grassland (Site 1), Woodland (Site 2)	Native grasses, shrubs
Property C (Sites 1–2)	Sparse woodland	Native/Naturalised grasses
Property D (Sites 1–2)	Woodland	Native/Naturalised grasses
Property E (Sites 1–2)	Grassland	Native/Naturalised grasses

Table 2. Field site locations and sampling schedule.

Property	Field Site	Longitude	Latitude (South)	Visit Dates	Season	Sampling Type [1]
A	Site 1	145.8957566	20.5097843	29 April 2023	Late wet	FS
				30 September 2023	Dry	FS, UAV
				18 February 2024	Wet	FS, UAV
				20 September 2024	Dry	FS, UAV
				18 August 2025	Dry	UAV
A	Site 2	145.9227423	20.5389117	2 May 2023	Late wet	FS
				2 October 2023	Dry	FS, UAV
A	Site 3	146.0163052	20.4216030	16 February 2024	Wet	FS, UAV
				22 September 2023	Dry	FS, UAV
				21 August 2025	Dry	UAV
B	Site 1	147.9482863	20.2945290	7 May 2023	Late wet	FS
				7 October 2023	Dry	FS, UAV
				9 February 2024	Wet	FS, UAV
				26 September 2024	Dry	FS, UAV
				13 August 2025	Dry	UAV
B	Site 2	147.9000386	20.2912746	9 May 2023	Late wet	FS
				10 October 2023	Dry	FS, UAV
				11 February 2024	Wet	FS, UAV
				29 September 2024	Dry	FS, UAV
C	Site 1	147.7126628	23.6817412	22 May 2023	Late wet	FS
				29 October 2023	Dry	FS, UAV
				11 April 2024	Wet	FS, UAV
				25 October 2024	Dry	FS, UAV
				10 September 2025	Dry	UAV
C	Site 2	147.7418179	23.6566114	24 May 2023	Late wet	FS
				31 October 2023	Dry	FS, UAV
				13 April 2024	Wet	FS, UAV
				26 October 2024	Dry	FS, UAV
D	Site 1	149.7133254	24.8590296	29 May 2023	Late wet	FS
				5 November 2023	Dry	FS, UAV
				17 April 2024	Wet	FS, UAV
				28 October 2024	Dry	FS, UAV
				15 September 2025	Dry	UAV
D	Site 2	149.6933999	24.8335359	31 May 2023	Late wet	FS
				8 November 2023	Dry	FS, UAV
				19 April 2024	Wet	FS, UAV
				30 October 2024	Dry	FS, UAV
E	Site 1	152.3962129	27.2509679	15 March 2023	Wet	FS
				30 August 2023	Dry	FS, UAV
				30 January 2024	Wet	FS, UAV
				27 August 2024	Dry	FS, UAV
				6 May 2025	Dry	UAV
E	Site 2	152.3966005	27.2513793	15 March 2023	Wet	FS
				30 August 2023	Dry	FS, UAV

[1] FS = Field spectrometer, UAV = Imaging UAV spectrometer.

2.2. Pasture Sampling

Pasture leaf chemistry at each field site was sampled from an area of approximately 1000 m^2. A transect of 100 m in length was located in an S-shaped configuration at each field

site. Sampling quadrats were located along the transect at nine evenly spaced locations. Each quadrat consists of a square quarter meter (0.5 × 0.5 m) area sample.

At each quadrat along the transect, plant samples were sorted by species into grass, forb, and shrub categories, and by growth state into live and dead classes. Sorted pasture leaf samples were oven dried for a minimum of 70 h at 60 degrees Celsius. Each sample was then sent for traditional chemistry analysis. The percent of ash-free Acid Detergent Fibre (ADF) was determined using an Ankom fibre digestion unit following the methods of [20]. Total percent nitrogen (TN) was determined using the methods developed by [21].

2.3. Derived Pasture Nutrition Metrics

To provide pasture nutrition metrics directly relevant to livestock management decisions, the laboratory-measured chemistry values were converted to industry-standard nutrition indicators:

Crude Protein (CP): Crude protein percentage was calculated from total nitrogen using the standard conversion factor [22]:

$$\text{CP (\%)} = \text{TN (\%)} \times 6.25 \quad (1)$$

This conversion factor is based on the assumption that plant proteins contain approximately 16% nitrogen by mass. Crude protein is the primary indicator of pasture nutritive value for livestock, with values above 7% generally considered adequate for maintaining cattle body condition.

Dry Matter Digestibility (DMD): Dry matter digestibility was derived from acid detergent fibre using the empirical relationship established for tropical and subtropical pastures [23]:

$$\text{DMD (\%)} = 88.9 - (0.779 \times \text{ADF (\%)}) \quad (2)$$

This relationship accounts for the inverse correlation between fibre content and digestibility, where higher ADF values indicate increased lignification and reduced forage digestibility. DMD values above 55% are generally considered adequate for maintenance of grazing cattle.

All subsequent modelling, results, and discussion refer to CP and DMD as the target variables, representing the practical pasture nutrition metrics of interest to livestock managers.

2.4. Pasture Spectral Response

Prior to each destructive pasture sampling effort, the spectral response of the sward was measured across four spatial scales spanning laboratory to satellite: leaf (laboratory), plant point (in situ quadrat), UAV (hyperspectral imaging), and satellite (Sentinel-2). This hierarchical sampling design enabled calibration and validation of predictive models at each scale, with progressively larger spatial footprints from individual leaves (∼1 cm^2) through UAV pixels (∼5 cm) to Sentinel-2 pixels (10 and 20 m).

2.4.1. Leaf Scale

Immediately following the sorting of collected pasture samples, spectral leaf-scale measurements were acquired for each life form class. A custom light-box providing near-ambient, diffuse lighting conditions was used to ensure repeatable illumination conditions across all samples. A RS-8800 portable spectroradiometer (Spectral Evolution, Lawrence, MA, USA) with a sensa-probe collected spectra in the 350–2500 nm range. After calibrating against a white Spectralon reference panel, surface reflectance was calculated by

dividing the target's radiance by the reference target's radiance. An average of five spectral measurements were taken for each sample to minimize measurement noise.

2.4.2. Plant Scale

In situ spectral measurements of the pasture were obtained using two complementary methodologies:

Method 1—Field Spectroradiometer: A backpack-mounted RS-8800 field spectroradiometer (Spectral Evolution, Lawrence, MA, USA) with sensa-probe attached to a survey post collected spectra within each sampling quadrat. Differential GNSS provided sub-meter location accuracy. Five measurements were taken per quadrat to capture spatial variability in plant composition.

Method 2—UAV Hyperspectral Imaging: A UAV hyperspectral imaging system consisting of a Hyspex Mjolnir VS620 full-range hyperspectral VNIR and SWIR dual camera imager (NEO—Norsk Elektro Optikk, Skedsmokorset, Norway) mounted onto a Blacksquare X8 Hercules heavy-lift multi-copter UAV (Blacksquare, Bogotá, Colombia) (Figure 2) was deployed at each field site. Table 3 details the UAV flight and sensor specifications.

Table 3. UAV flight and hyperspectral sensor specifications.

Parameter	Specification
UAV Platform	Blacksquare X8 Hercules heavy-lift multi-copter
Sensor	Hyspex Mjolnir VS620 VNIR-SWIR
Spectral Range	400–2500 nm
Spectral Bands	620 bands
Flight Altitude	120 m AGL
Ground Speed	3 m/s
Spatial Resolution	32–64 mm (VNIR-SWIR)
Swath Width	20° FOV
GNSS/IMU	Applanix APX-15 UAV with Emlid RS2 base station

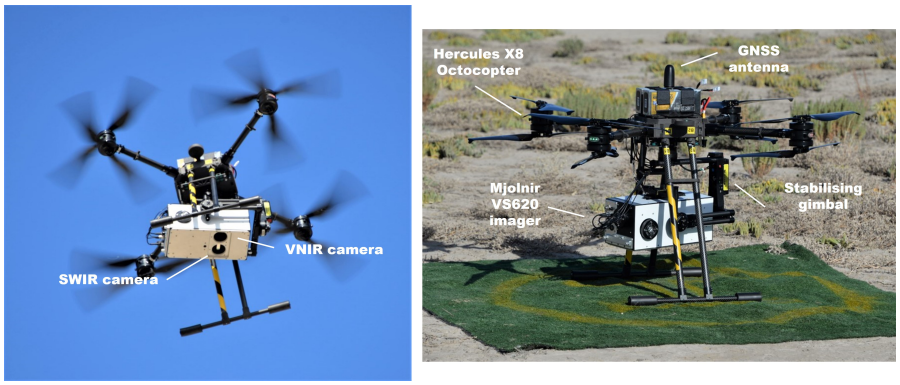

Figure 2. UAV hyperspectral imaging system showing the Blacksquare X8 Hercules multi-copter with mounted Hyspex Mjolnir VS620 sensor.

Direct Georeferencing: Rather than relying solely on feature-based image matching, direct georeferencing was employed using the Applanix APX-15 UAV-integrated GNSS/IMU system (Trimble Applanix, Richmond Hill, ON, Canada). A survey-grade Emlid RS2 GNSS base station (Emlid, Hong Kong) was deployed at each field site to enable post-processed kinematic (PPK) corrections. Position and orientation data were processed using Applanix POSPac UAV software (version 2025.10, Trimble Applanix, Richmond Hill, ON, Canada), achieving sub-decimetre positional accuracy.

Radiometric Calibration: Raw sensor data were converted to at-sensor radiance using a three-stage calibration process: (1) dark current offset correction using pre-flight dark frame captures, (2) responsivity correction using laboratory-derived sensor response matrices, and (3) spectral calibration using known emission line sources.

Geometric Correction: Orthorectification was performed using PARGE software (version 4.1, ReSe Applications, Wil, Switzerland) with direct georeferencing inputs and a digital surface model (DSM). For captures in 2023 and 2024, the DSM was derived from Wingtra One photogrammetric UAV flights (Wingtra, Zurich, Switzerland) conducted concurrently at each field site. For captures in 2025, the DSM was derived from a RIEGL miniVUX LiDAR sensor (RIEGL, Horn, Austria) mounted on a DJI M350 RTK UAV platform (DJI, Shenzhen, China), providing higher precision terrain data particularly in vegetated areas.

Atmospheric Compensation: Surface reflectance was retrieved using DROACOR atmospheric correction software (version 2.1) [24], which accounts for solar irradiance, atmospheric path radiance, and surface–atmosphere interactions. BRDF effects were minimized using BREFCOR for bidirectional reflectance distribution function correction [25].

2.4.3. Canopy Scale

The spectral response of the pasture canopy was measured from Sentinel-2A and 2B satellites. Table 4 details the satellite sensor specifications.

Table 4. Sentinel-2 satellite sensor specifications.

Band	Central Wavelength (nm)	Spatial Resolution (m)
B02—Blue	490	10
B03—Green	560	10
B04—Red	665	10
B05—Red Edge 1	705	20
B06—Red Edge 2	740	20
B07—Red Edge 3	783	20
B08—NIR	842	10
B8A—NIR Narrow	865	20
B11—SWIR 1	1610	20
B12—SWIR 2	2190	20

Level 2A Surface reflectance products were downloaded from the Copernicus Open Access Hub. Cloud masking was undertaken using the provided SCL (Scene Classification Layer) cloud masks. For each UAV and field sampling campaign, the closest cloud-free Sentinel-2 capture date was matched, with temporal offsets typically ranging from 0 to 7 days depending on cloud cover conditions.

2.5. Spectral Preprocessing and Feature Analysis

2.5.1. Spectral Preprocessing

Prior to model development, spectral data underwent several preprocessing steps to enhance signal quality and ensure compatibility across measurement scales. The preprocessing pipeline was implemented using custom Python scripts leveraging NumPy and SciPy libraries.

Water absorption band removal: Atmospheric water vapour introduces strong absorption features that obscure the underlying vegetation spectral response. Spectral bands within water absorption regions were systematically removed from all hyperspectral datasets. The removed wavelength ranges included: oxygen absorption bands (718–732 nm and 758–778 nm), weak water absorption (813–827 nm), strong water absorption features centred at approximately 940 nm (893–977 nm), 1130 nm (1110–1160 nm), 1400 nm

(1330–1480 nm), and 1900 nm (1780–2000 nm), and edge-of-sensor water/CO_2 absorption (2400–2500 nm). This filtering reduced the hyperspectral feature count from 620 bands to approximately 405 usable bands while removing regions dominated by atmospheric rather than vegetation signals.

Spectral resampling: Laboratory and field spectroradiometer data were resampled to match the UAV sensor's spectral response function using linear interpolation. This ensured spectral consistency across measurement scales and enabled direct comparison of absorption features identified at different spatial resolutions. Savitzky–Golay filtering (window size 11, polynomial order 2) was applied to smooth high-frequency noise in the SWIR region where signal-to-noise ratios were lower.

Reflectance scaling: All spectral data were normalised to fractional reflectance (0–1 range). UAV and Sentinel-2 surface reflectance products stored as scaled integers (factor 10,000) were divided to convert to fractional reflectance. Laboratory spectra acquired as radiance were converted to reflectance through division by the Spectralon reference panel measurements.

2.5.2. Feature Dimensionality Reduction

TabPFN performs optimally with feature counts below 200 dimensions. The hyperspectral datasets, even after water absorption band removal, contained 405 spectral features requiring dimensionality reduction. Principal Component Analysis (PCA) was employed to reduce the hyperspectral feature space while retaining maximum spectral information.

For leaf- and UAV-scale models, PCA was fitted on the training data and reduced the 405 spectral bands to 100 principal components. This reduction retained $> 99.9\%$ of total spectral variance while meeting TabPFN's feature constraints. The PCA transformation was stored and applied identically to validation data to prevent data leakage.

For the Sentinel-2 canopy-scale model, no dimensionality reduction was required as input consisted of only 10 spectral bands (B02, B03, B04, B05, B06, B07, B08, B8A, B11, B12), well within TabPFN's optimal feature range.

2.5.3. Spectral–Chemistry Correlation Analysis

To identify spectral regions most informative for pasture nutrition prediction, Pearson correlation coefficients were calculated between individual spectral bands and chemistry target variables (TN and ADF) across the training dataset. This analysis was performed separately for each life form class (green grass, green forb, dead material) to account for differences in spectral–chemistry relationships.

Key absorption features were identified at wavelengths showing statistically significant correlations ($|r| > 0.3$, $p < 0.01$), consistent with established spectral–biochemistry relationships [26]:

- **Visible region (400–700 nm):** Chlorophyll absorption features at 350 nm and 680–690 nm showed moderate positive correlations with crude protein content, reflecting the nitrogen investment in photosynthetic pigments [6].
- **Red-edge region (700–750 nm):** The red-edge inflection point position correlated with nitrogen status, with higher protein content shifting the red-edge towards longer wavelengths [7].
- **NIR plateau (750–1300 nm):** Reflectance magnitude in the NIR plateau showed positive correlations with protein content due to increased leaf cellular structure in actively growing vegetation [6].
- **SWIR water features (1400–1500 nm, 1900–2100 nm):** Water absorption depth correlated with tissue moisture content, which inversely relates to ADF (fibre content increases as plants senesce and lose moisture).

- **Protein/nitrogen features (2050–2200 nm):** Direct absorption by N-H bonds in protein molecules produced the strongest correlations with crude protein content in the SWIR region [26].

These spectral–chemistry relationships informed interpretation of model behaviour and validated that predictions were based on physiologically meaningful spectral features rather than spurious correlations.

2.6. Modelling

2.6.1. Data Preparation

Model calibration and testing was undertaken through the joining of plant leaf chemistry analysis with spectral response measurements. At the leaf scale, collected spectra were matched with sorted chemical analysis results based on sample identifiers. At the plant scale, spectral measurements taken across each quadrat were averaged, and chemical results of both grass and forb life-forms were averaged for each quadrat to provide a representative sample. At the Sentinel-2 canopy scale, direct laboratory chemistry measurements were not available for satellite pixels; instead, the UAV-scale TabPFN model was first applied to predict chemistry values from UAV hyperspectral imagery, and these UAV-derived predictions served as pseudo-observations for training the Sentinel-2 model. Specifically, the spatially averaged UAV chemistry predictions within each Sentinel-2 pixel footprint (20 m) were paired with the corresponding Sentinel-2 spectral bands. This hierarchical training approach transfers spectral–chemistry relationships learned at the high-resolution UAV scale to the coarser satellite scale.

2.6.2. TabPFN Machine Learning Model

TabPFN (Tabular Prior-Data Fitted Network) models were employed across all scales [17]. TabPFN represents a significant advancement in tabular machine learning, utilizing a transformer architecture pre-trained on millions of synthetic datasets to perform in-context learning. Unlike traditional neural networks requiring extensive hyperparameter tuning and training epochs, TabPFN performs inference in a single forward pass, making it particularly effective for small-to-medium-sized tabular datasets typical of remote sensing calibration scenarios.

Key architectural features of TabPFN include:

- **Pre-trained prior:** The model is pre-trained on synthetic datasets generated from a prior distribution over dataset-generating processes, enabling rapid adaptation to new datasets.
- **In-context learning:** Unlike traditional models that learn parameters, TabPFN learns to predict by conditioning on the training data as context during inference.
- **Ensemble prediction:** Multiple predictions are averaged using 8 ensemble estimators to improve prediction stability and reduce variance.
- **No hyperparameter tuning:** The pre-trained architecture eliminates the need for learning rate schedules, dropout rates, or architecture search.

For high-dimensional hyperspectral data at leaf and UAV scales, dimensionality reduction via Principal Component Analysis (PCA) was applied to reduce features to 100 principal components while retaining >99.9% of variance. This preprocessing step was necessary as TabPFN performs optimally with feature counts below 200.

For the Sentinel-2 canopy-scale model, input features consisted of surface reflectance values from 10 Sentinel-2 bands (B02, B03, B04, B05, B06, B07, B08, B8A, B11, B12), scaled from digital numbers to reflectance by dividing by 10,000. Models were calibrated using 80 percent of the data and validated with the remaining 20 percent using stratified random sampling to ensure representative coverage of chemistry value ranges. The train–validation

split was performed randomly within each spatial scale without explicit spatial blocking by site or transect; implications for potential spatial autocorrelation are addressed in the Discussion. Separate models were developed for crude protein (CP) and dry matter digestibility (DMD) prediction at each spatial scale, with model outputs representing the derived quality metrics described above.

2.7. Pasture Classification Masking

To improve prediction accuracy and focus chemistry estimates on productive pasture areas, a pasture classification masking approach was implemented as an essential preprocessing step. In heterogeneous rangeland landscapes, satellite pixels often contain mixtures of pasture, bare soil, rocky outcrops, and woody vegetation canopies that confound chemistry predictions. This classification step represents a critical preprocessing requirement rather than an optional enhancement—analogous to atmospheric correction, ensuring that only spectrally pure pasture targets contribute to nutrition estimates. Without target purity, chemistry predictions derived from mixed pixels would be fundamentally unreliable.

The pasture classification model was developed using a deep neural network (DNN) classifier trained on UAV hyperspectral imagery with manually digitised land cover labels derived from high-resolution RGB orthomosaics. The DNN architecture consisted of a sequential multi-layer perceptron with four hidden layers (512, 256, 128, and 64 neurons), batch normalization, L2 regularization (0.001), and dropout (0.4) for regularization. The model was trained using the Adam optimizer with a learning rate of 0.0005. The classifier uses the following input features:

- Spectral reflectance from 405 hyperspectral bands (visible to SWIR);
- Photogrammetry and LiDAR-derived vegetation height from the digital surface model (DSM) minus the digital terrain model (DTM).

For satellite-scale chemistry predictions, the UAV-derived pasture classification masks were resampled to match Sentinel-2 pixel resolution (20 m). A pixel was classified as pasture if the majority of underlying UAV pixels (>50%) were classified as pasture. This conservative threshold ensures that only pixels with predominantly pasture cover contribute to chemistry predictions.

The DNN classifier produced seven land cover classes: bare soil, dry pasture, dry timber, green pasture, shadow, shrub cover, and tree cover. The model was trained on 13,892 labelled pixels from 20 UAV hyperspectral datasets, achieving an overall test accuracy of 98.6%. Data partitioning used random stratified sampling by class label (20% held-out test set), with pixels from the same UAV image potentially appearing in both training and testing sets; the partitioning was not separated by site or acquisition date. Table 5 presents the per-class accuracy metrics. For chemistry prediction masking, the pasture classes (green pasture and dry pasture) were combined to create the binary pasture mask, while all other classes were excluded.

Figure 3 presents the normalized confusion matrix for the classification model, showing the proportion of correctly classified pixels for each land cover class. The matrix demonstrates strong diagonal dominance, indicating high classification accuracy across all classes. Minor confusions occur between spectrally similar classes, particularly between dry timber and shrub cover, which share similar woody vegetation characteristics.

Figure 4 presents classification outputs from two representative field sites captured during both dry and wet seasons, demonstrating the model's robustness across seasonal conditions. Each site is shown with both a true-color RGB composite (left) derived from the hyperspectral imagery and the corresponding DNN classification result (right). The Baryugal site (a,b) represents an open grassland environment dominated by pasture cover with minimal woody vegetation, shown during the dry season (November 2023) and wet

season (April 2024). The Mt Pleasant site (c,d) represents an open woodland environment with substantial tree cover interspersed with pasture areas, also shown during the dry season (October 2023) and wet season (February 2024). The seasonal comparison reveals the transition from dry pasture (yellow) to green pasture (green) following wet season rainfall, while non-pasture cover types (trees, shrubs, bare soil) remain consistently classified. These examples demonstrate the classifier's ability to accurately delineate pasture from non-pasture cover types across diverse rangeland conditions and seasonal states.

Table 5. Pasture classification model accuracy metrics by land cover class. The deep neural network classifier achieved 98.6% overall accuracy on held-out test data (n = 2779).

Class	Precision	Recall	F1-Score	Support
Bare soil	0.99	0.99	0.99	285
Dry pasture	0.98	0.97	0.97	421
Dry timber	0.96	0.96	0.96	318
Green pasture	0.99	0.99	0.99	512
Shadow	0.99	0.99	0.99	224
Shrub cover	0.97	0.98	0.98	492
Tree cover	0.98	0.98	0.98	527
Weighted avg	**0.99**	**0.99**	**0.99**	**2779**

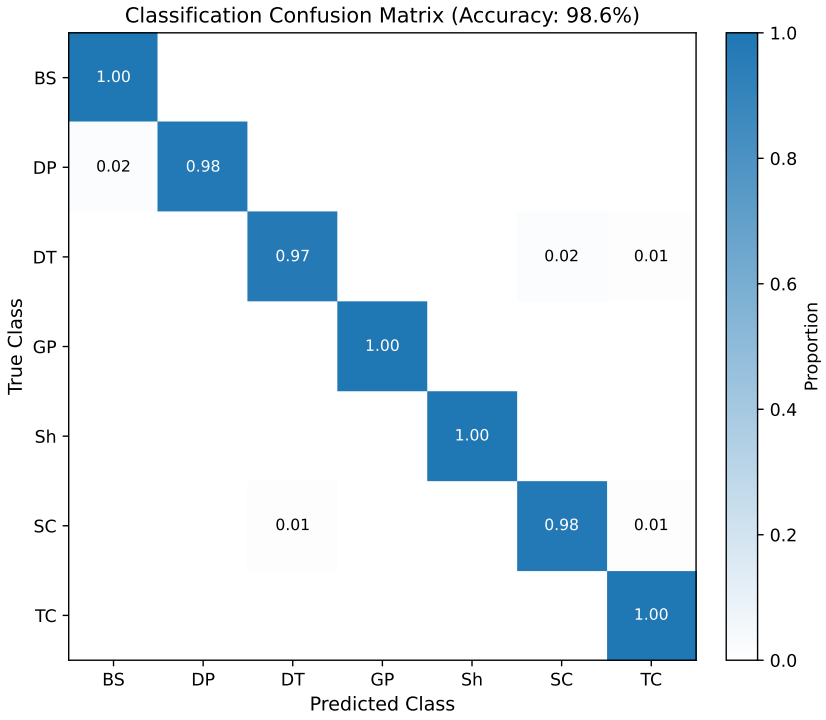

Figure 3. Normalized confusion matrix for the pasture classification model. Class abbreviations: BS = Bare soil, DP = Dry pasture, DT = Dry timber, GP = Green pasture, Sh = Shadow, SC = Shrub cover, TC = Tree cover. Values represent the proportion of true class samples classified into each predicted class (row-normalized).

The masking approach was applied during time-series generation, whereby only pasture-classified pixels contributed to the spatial average of predicted chemistry values for each field site or paddock analysis unit.

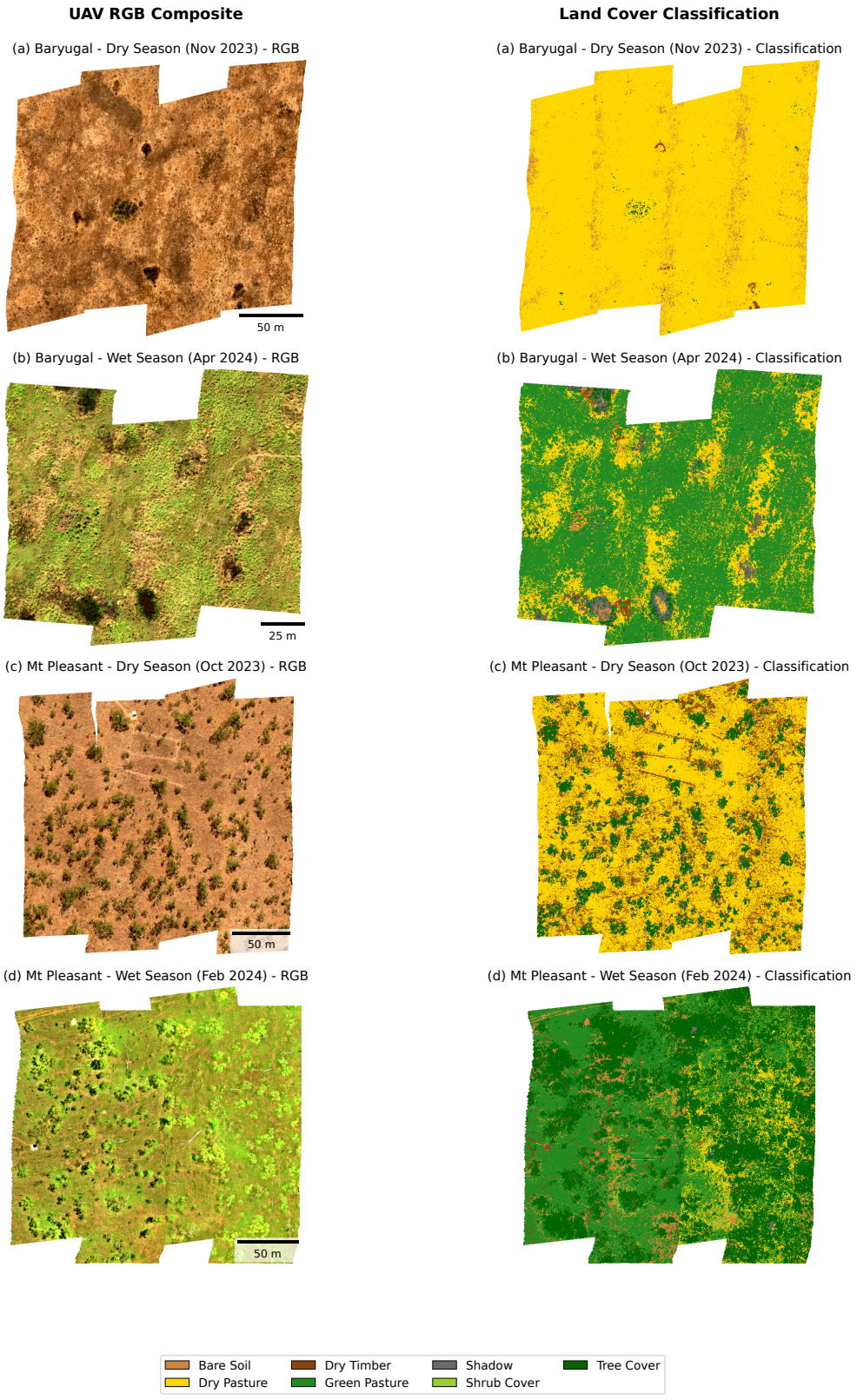

Figure 4. UAV hyperspectral true-color RGB composites and corresponding DNN land cover classification outputs for two field sites across dry and wet seasons. Left column shows RGB images (bands at 640, 550, and 470 nm); right column shows classification results. (**a**) Baryugal Site 1, dry season (November 2023). (**b**) Baryugal Site 1, wet season (April 2024). (**c**) Mt Pleasant Site 1, dry season (October 2023). (**d**) Mt Pleasant Site 1, wet season (February 2024). The seasonal transition from dry pasture (yellow) to green pasture (green) is clearly visible, demonstrating the model's ability to differentiate phenological states while consistently excluding non-pasture cover types.

3. Results

3.1. Pasture Sampling

Figure 5 presents the results of wet chemistry analysis from all sampling campaigns. Leaf and stem chemistry varied between green and dead material, as well as between grasses and forbs. Crude protein (derived from total nitrogen) was skewed towards lower percentages in pasture grasses. Green forbs had approximately double the protein content of green grasses, reflecting the higher protein content of leguminous species such as Seca Stylo (*Stylosanthes scabra*). Dry matter digestibility (derived from ADF) was lowest in the combined dead grass and forb samples, consistent with increased lignification during senescence.

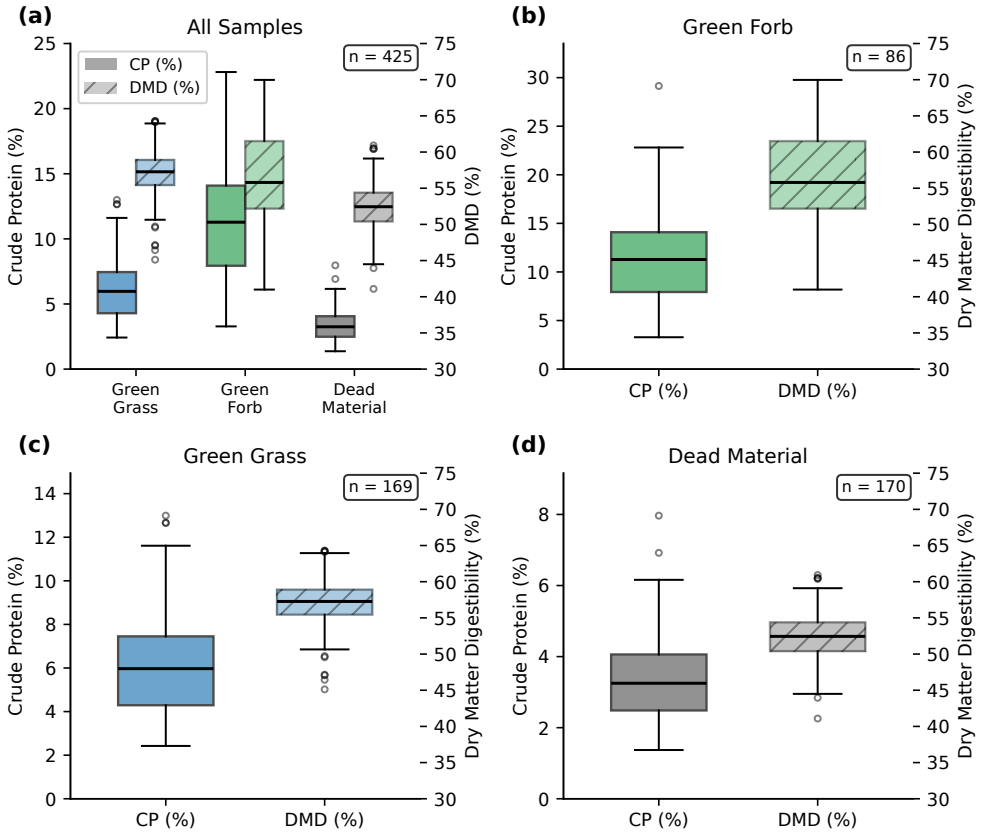

Figure 5. Distribution of wet chemistry results across all sampling campaigns. Box plots show crude protein (CP, %, solid fill) and dry matter digestibility (DMD, %, hatched fill) for: (**a**) all samples combined showing green grass (extitn = 169), green forb (extitn = 86), and dead material (extitn = 170); (**b**) green forb samples showing higher CP (mean 11.4%) due to nitrogen-fixing leguminous species; (**c**) green grass samples (mean CP 6.1%); (**d**) dead material showing lowest CP (mean 3.4%) and DMD values consistent with lignification during senescence. DMD is derived from ADF using the equation: DMD = 88.9–0.779 × ADF.

3.2. Pasture Spectral Response and Feature Analysis

The spectral responses of pasture samples were characterized across all spatial scales following the preprocessing pipeline described in Section 2.5. After water absorption band removal, 405 spectral bands remained for analysis from the original 620 hyperspectral bands. Figure 6 presents the mean spectral signatures for green (actively growing) and senescent (dry) pasture conditions at each measurement scale.

At the leaf scale, individual life forms showed characteristic absorption features in the visible (chlorophyll), NIR (cell structure), and SWIR (water and biochemical) regions. Green grass and green forb samples exhibited similar spectral shapes with typical vegetation

curves—strong chlorophyll absorption in the red region (centred at 680 nm), steep red-edge transitions (700–750 nm), and high NIR reflectance plateaus (0.35–0.45 reflectance)—and are therefore shown as a combined green pasture average in Figure 6. Dead material exhibited flattened responses with reduced NIR reflectance (0.25–0.35) and less pronounced chlorophyll absorption, consistent with degradation of photosynthetic pigments during senescence.

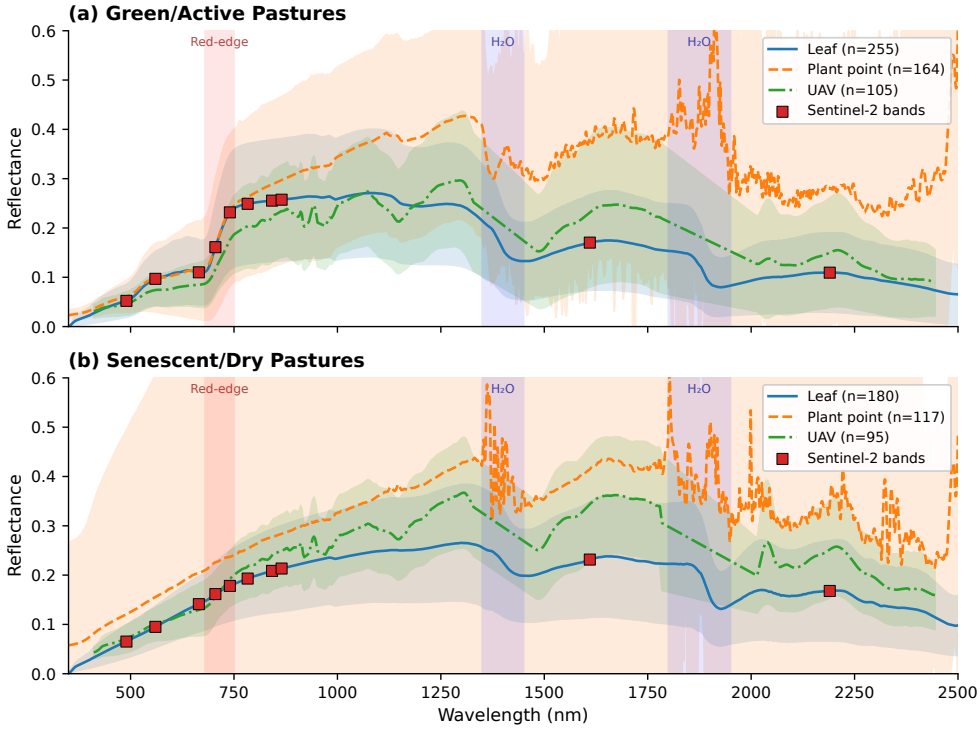

Figure 6. Average spectral response of pastures at different sampling scales. (**a**) Green/active pastures showing typical vegetation curves with strong NIR reflectance plateau and chlorophyll absorption in the red region. (**b**) Senescent/dry pastures exhibiting flattened responses with reduced NIR reflectance due to loss of cell structure. Shaded regions indicate ±1 standard deviation. Red-edge and water absorption regions are highlighted. Sentinel-2 band positions (squares) demonstrate the relationship between continuous spectral measurements and satellite band sampling.

3.2.1. Spectral–Chemistry Correlations

Correlation analysis between spectral reflectance and chemistry values identified key absorption features supporting the physiological basis for spectral prediction of pasture nutrition. Table 6 summarises the Pearson correlation coefficients between reflectance and chemistry values across key spectral regions.

Crude protein correlations: Strong positive correlations were observed between CP content and reflectance at 350 nm (UV region, $r = 0.42$), 540–560 nm (green peak, $r = 0.38$), and throughout the NIR plateau (750–1100 nm, $r = 0.35$–0.45). Strong negative correlations occurred at 680–690 nm (chlorophyll absorption, $r = -0.48$) reflecting the direct relationship between nitrogen content and chlorophyll concentration. In the SWIR region, correlations at 1390–1437 nm ($r = -0.36$) and 1870–2020 nm ($r = 0.40$–0.52) correspond to known nitrogen and protein absorption features, with the strongest correlations at 2050–2170 nm where N-H bond stretching produces direct absorption features.

Table 6. Summary of spectral–chemistry correlations by wavelength region. Pearson *r* values between reflectance and crude protein (CP) or dry matter digestibility (DMD). Positive values indicate higher reflectance with higher nutrient content; negative values indicate absorption.

Region	Wavelength (nm)	CP (*r*)	DMD (*r*)	Feature
UV	350	+0.42	–	UV absorption
Green peak	540–560	+0.38	–	Chlorophyll refl.
Red	680–690	−0.48	+0.35	Chlorophyll abs.
NIR plateau	750–1100	+0.35 to +0.45	–	Cell structure
SWIR	1390–1437	−0.36	−0.45 to −0.55	Water abs.
SWIR	1870–2020	+0.40 to +0.52	–	N-H stretching
SWIR	2050–2170	+0.45 to +0.55	–	Protein abs.
SWIR	2100–2300	–	−0.38 to −0.45	Lignin/cellulose

Dry matter digestibility correlations: ADF (the predictor of DMD) showed inverse correlation patterns to CP at most wavelengths, consistent with the negative relationship between fibre content and digestibility. Particularly strong negative correlations with DMD were observed in water absorption regions for green forb samples ($r = -0.45$ to -0.55), reflecting the relationship between tissue hydration status and cell wall maturity. The SWIR region (2100–2300 nm) showed consistent negative correlations ($r = -0.38$ to -0.45) with DMD corresponding to cellulose and lignin absorption features that increase with plant maturity.

3.2.2. Feature Reduction Performance

PCA dimensionality reduction from 405 bands to 100 principal components retained 99.94% of total spectral variance at the leaf scale and 99.91% at the UAV scale. The first three principal components captured the dominant modes of spectral variation: PC1 (78% variance) primarily represented overall brightness/albedo differences between samples, PC2 (12% variance) captured the contrast between visible and NIR regions corresponding to vegetation greenness, and PC3 (4% variance) represented SWIR features related to moisture content.

Figure 6 compares average spectral responses across sampling scales for both green/growing and dead/senescent pasture conditions.

3.3. TabPFN Model Results

Following the preparation of both the pasture leaf chemistry and associated spectral responses, training and testing of TabPFN models were undertaken for each sampling scale. Results of each scale's model calibration and validation are presented, followed by a summary and comparison of model performance metrics.

3.3.1. Leaf Scale

Results of the calibration and validation of the leaf-scale TabPFN model are presented in Figure 7. Model fitting metrics demonstrate strong performance, with R^2 values of 0.92 (calibration) and 0.76 (validation) for crude protein, and RMSE of 1.40%. Results of the digestibility model are lower, with R^2 values of 0.86 (calibration) and 0.47 (validation), and RMSE of 2.6%. The reduction in validation performance relative to calibration suggests some overfitting to the spectral characteristics of the training samples. Duplicate laboratory spectra (where identical spectra were assigned to multiple chemistry measurements from the same sample) were identified and averaged prior to modelling, reducing the training dataset from 391 to 370 samples and eliminating potential data leakage.

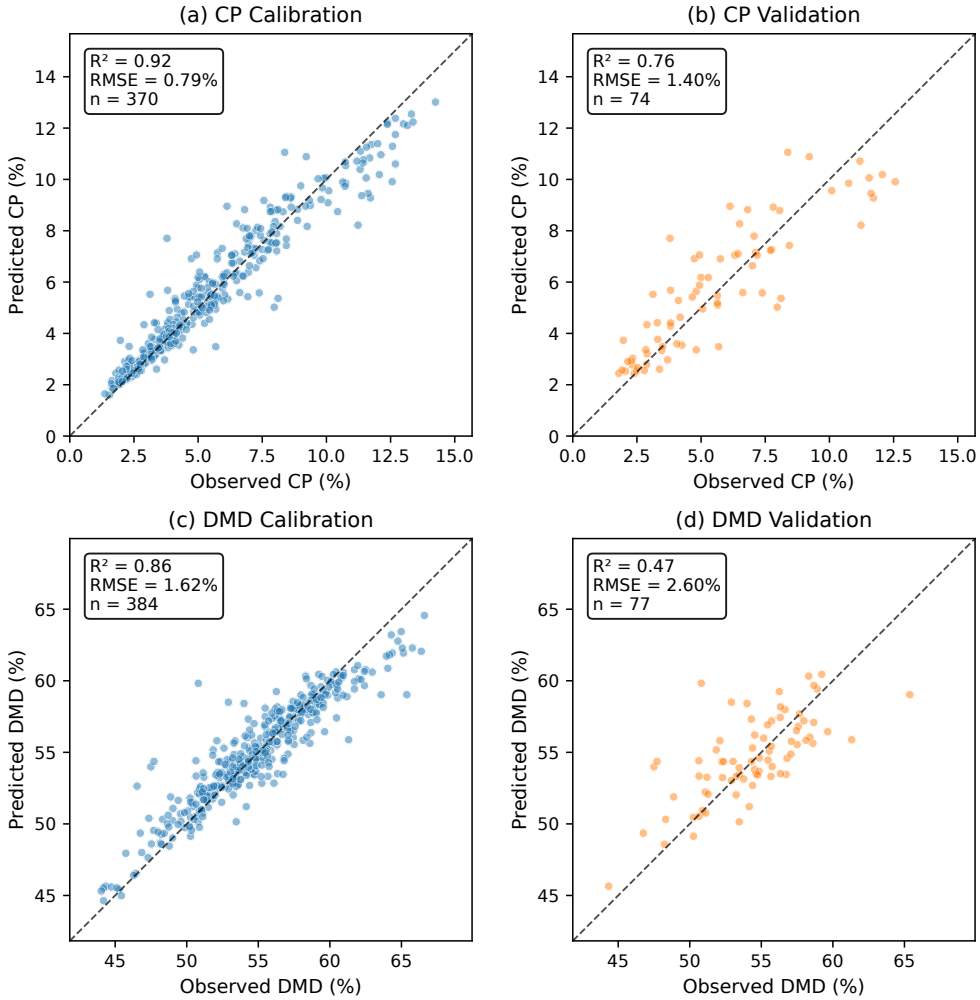

Figure 7. Leaf-scale TabPFN model calibration and validation results. Blue dots represent calibration samples, red dots represent validation samples, and the dotted line indicates the 1:1 relationship. (**Top row**): Crude protein (CP) showing calibration (**a**) R^2 = 0.92 and validation (**b**) R^2 = 0.76. (**Bottom row**): Dry matter digestibility (DMD) showing calibration (**c**) R^2 = 0.86 and validation (**d**) R^2 = 0.47. Models trained on deduplicated lab spectra (n = 296 calibration, n = 74 validation for CP; n = 307 calibration, n = 77 validation for DMD).

3.3.2. Plant-Point-Scale

Results of the plant-point-scale TabPFN model are presented in Figure 8. R^2 values were 0.74 (calibration) and 0.58 (validation) for crude protein, with an RMSE of 0.25%. Results of the digestibility model are lower, with R^2 values of 0.67 (calibration) and 0.13 (validation), and RMSE of 3.1%. The lower performance at this scale is attributed to the mixing of grass and forb spectral signatures at the quadrat level without the spatial averaging benefits of imaging sensors.

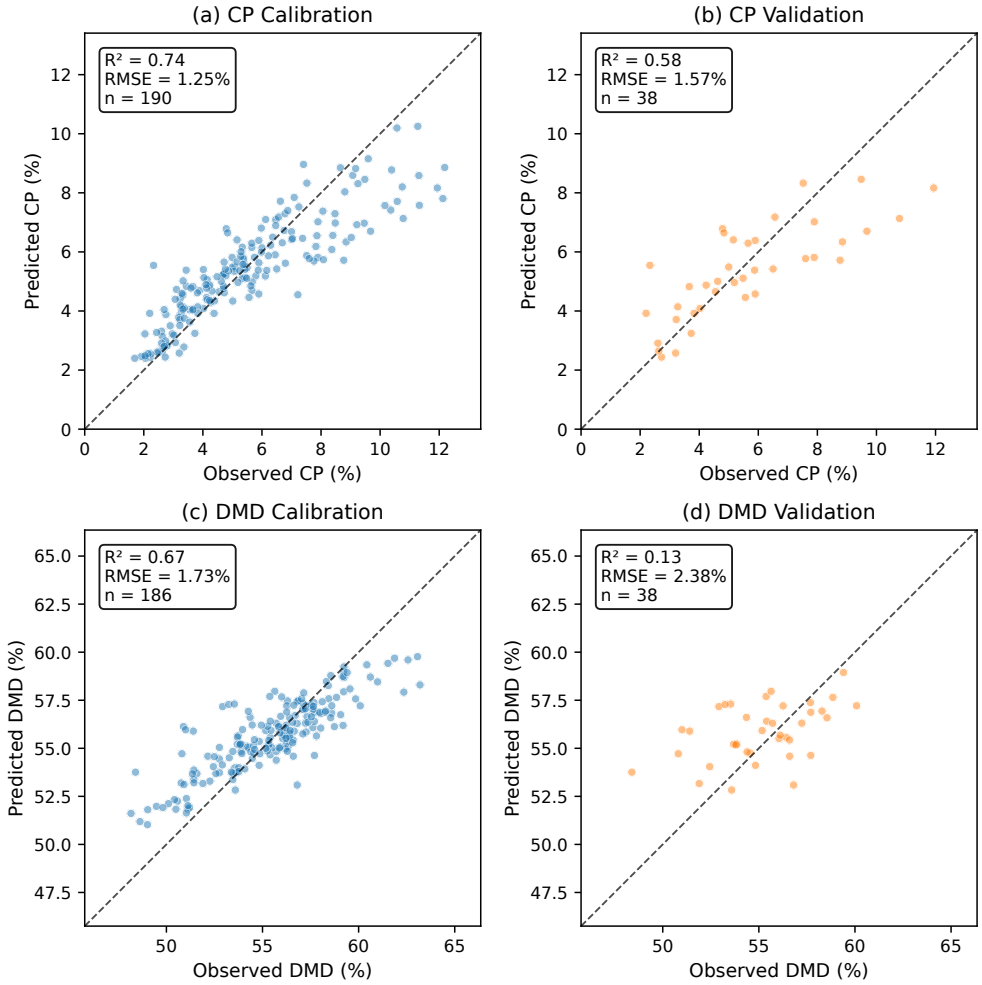

Figure 8. Plant-point-scale TabPFN model calibration and validation results. Blue dots represent calibration samples, red dots represent validation samples, and the dotted line indicates the 1:1 relationship. (**Top row**): Crude protein (CP) showing calibration (**a**) R^2 = 0.74 and validation (**b**) R^2 = 0.58. (**Bottom row**): Dry matter digestibility (DMD) showing calibration (**c**) R^2 = 0.67 and validation (**d**) R^2 = 0.13. Models trained on 190 samples (n = 152 calibration, n = 38 validation). $\hat{\sigma}$ represents mean model uncertainty derived from 80% prediction intervals.

3.3.3. Plant UAV Scale

Results of the UAV-scale TabPFN model are presented in Figure 9. R^2 values were 0.94 (calibration) and 0.95 (validation) for crude protein, with an RMSE of 0.79%. Results of the digestibility model achieved R^2 values of 0.89 (calibration) and 0.88 (validation), and RMSE of 1.05%. Six DMD samples were identified as outliers based on a two-stage diagnostic process: (1) initial identification via $>2\sigma$ prediction residual, followed by (2) spectral verification confirming soil contamination. Critically, these same six quadrats remained valid for crude protein modelling (included in the CP training set), indicating that the exclusion reflects a DMD-specific spectral measurement failure rather than general sample quality issues. Visual inspection of UAV imagery confirmed that all six quadrats had <30% vegetation fractional cover, resulting in soil-dominated spectral signatures that broke the plant chemistry–reflectance relationship for fibre-related absorption features. Removing these samples improved DMD validation R^2 from 0.19 to 0.88. The UAV-scale benefits from spatial averaging that reduces noise from individual plant variability while maintaining high spectral resolution.

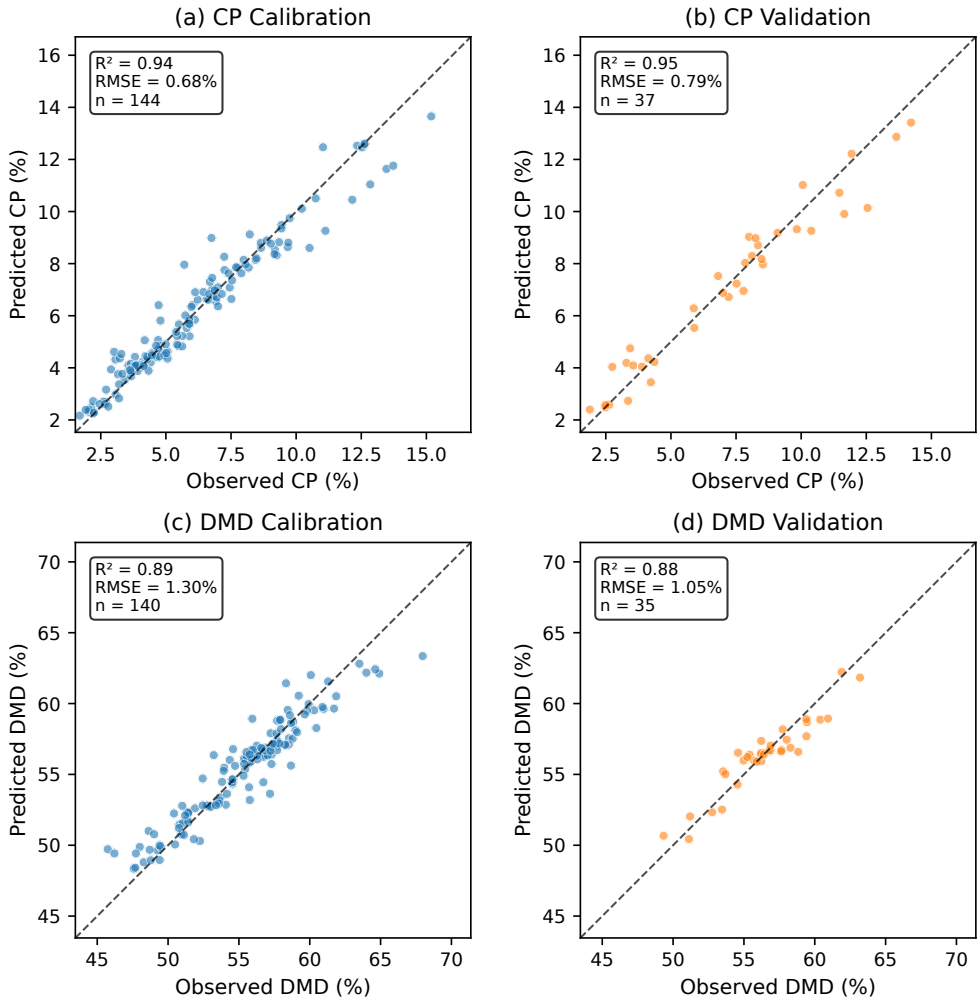

Figure 9. UAV scale TabPFN model calibration and validation results. Blue dots represent calibration samples, red dots represent validation samples, and the dotted line indicates the 1:1 relationship. (**Top row**): Crude protein (CP) showing calibration (**a**) R^2 = 0.94 and validation (**b**) R^2 = 0.95. (**Bottom row**): Dry matter digestibility (DMD) showing calibration (**c**) R^2 = 0.89 and validation (**d**) R^2 = 0.88. CP models trained on 181 samples (n = 144 calibration, n = 37 validation). DMD models trained on 175 samples after removal of 6 outliers with soil reflectance contamination (n = 140 calibration, n = 35 validation). $\hat{\sigma}$ represents mean model uncertainty derived from 80% prediction intervals.

3.3.4. Canopy Sentinel-2 Scale

Figure 10 illustrates the observed versus predicted model calibration and validation values for both crude protein and dry matter digestibility at the Sentinel-2 canopy scale. Unlike the leaf, plant, and UAV scales where models were trained against laboratory chemistry measurements, the Sentinel-2 model was trained using UAV-derived chemistry predictions as pseudo-observations (see Section 2.5.1). Accordingly, calibration R^2 values are not reported for the Sentinel-2 scale, as they would reflect the model's ability to fit UAV predictions rather than true chemistry values. The validation R^2 values represent the held-out 20% of the UAV-derived training samples. Model fitting metrics demonstrate excellent goodness of fit at the satellite scale, with R^2 values of 0.92 for validation for crude protein, and RMSE of 0.39%. Results of the digestibility model achieved R^2 values of 0.73 for validation, and RMSE of 0.95%. The models were trained on 369 samples and validated on 93 samples from 19 site–date combinations across 11 field sites.

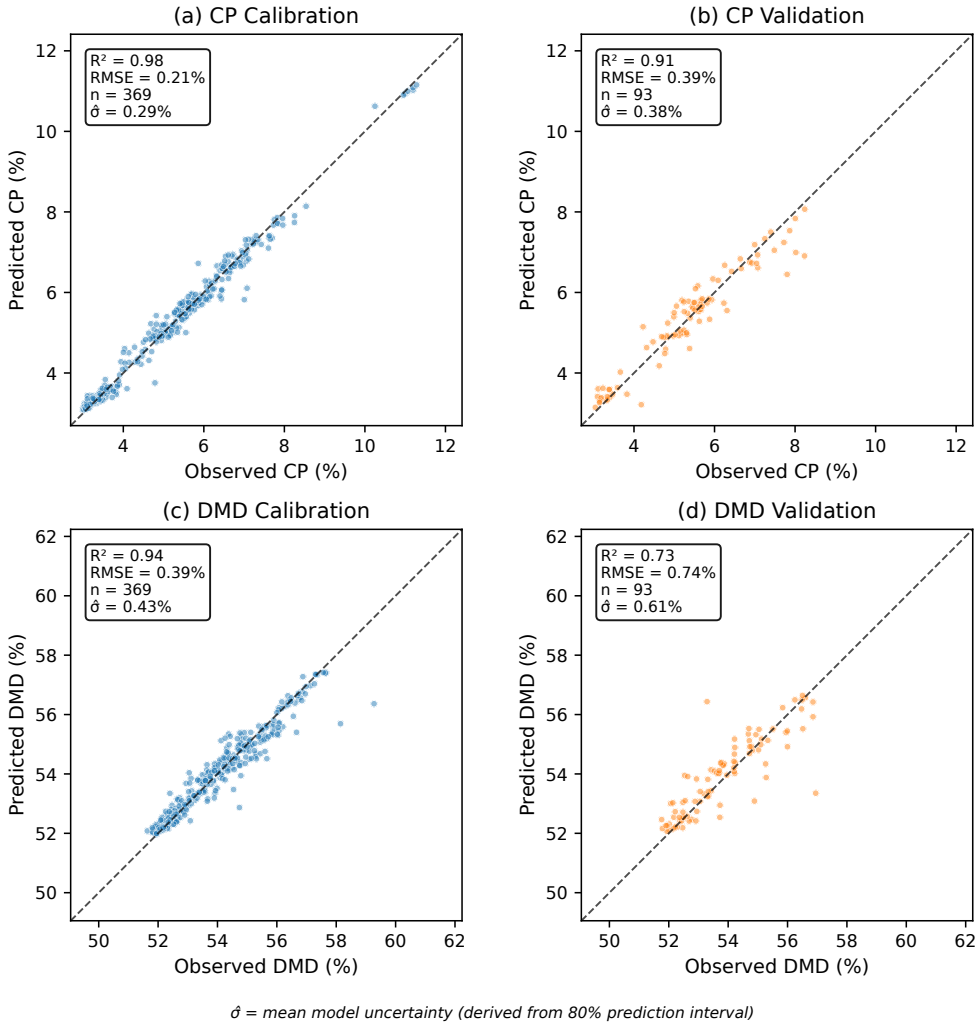

Figure 10. TabPFN Sentinel-2 model calibration and validation results. (**Top row**): Crude protein (CP) showing validation R^2 = 0.92. (**Bottom row**): Dry matter digestibility (DMD) showing validation R^2 = 0.73. Models trained on 369 samples and validated on 93 samples from 19 site–date combinations. $\hat{\sigma}$ represents mean model uncertainty derived from TabPFN's quantile predictions.

3.3.5. Model Comparison

Table 7 presents each TabPFN model's goodness of fit and error metrics across all spatial scales, enabling direct comparison of model performance.

The highest performing models for crude protein were at the UAV scale, achieving validation R^2 = 0.95, with the Sentinel-2 scale achieving comparable performance (R^2 = 0.92) despite reduced spectral resolution. The Sentinel-2-scale model achieved validation R^2 = 0.92 for crude protein and R^2 = 0.73 for digestibility, with 369 training and 93 validation samples from 19 site–date combinations. The crude protein model achieved an RMSE of 0.39%, representing excellent predictive accuracy at the satellite scale. The plant-point-scale model had the lowest goodness of fit for digestibility.

The TabPFN architecture provided notable advantages including: faster training without hyperparameter tuning, improved generalisation from limited training samples, and more stable predictions across diverse pasture conditions.

Table 7. Summary of TabPFN model calibration and validation statistics for each sampling scale. CP = crude protein, DMD = dry matter digestibility.

Model Scale	Parameter	Type	N (Samples)	R^2	RMSE (%)	$\hat{\sigma}$ (%)
Leaf	CP	Cal	296	0.92	0.79	1.40
		Val	74	0.76	1.40	
	DMD	Cal	307	0.86	1.62	2.60
		Val	77	0.47	2.60 [L]	
Plant point	CP	Cal	152	0.74	1.25	1.00
		Val	38	0.58	1.56	
	DMD	Cal	149	0.67	1.71	1.77
		Val	37	0.13 [L]	2.41	
UAV	CP	Cal	144	0.94	0.68	0.79
		Val	37	0.95	0.79	
	DMD	Cal	140	0.89	1.30	1.05
		Val	35	0.88	1.05	
Sentinel-2 *	CP	Cal	369	–	–	0.39
		Val	93	0.92	0.39	
	DMD	Cal	369	–	–	0.95
		Val	93	0.73	0.95	

[L] Lowest performing. $\hat{\sigma}$ = mean model uncertainty derived from 80% prediction interval. All models use TabPFN architecture with 8 ensemble estimators. * Sentinel-2 calibration metrics omitted as the model was trained on UAV-derived predictions (pseudo-observations) rather than laboratory chemistry measurements.

3.4. Model Predictions

3.4.1. Field Site Examples

Figures 11 and 12 provide direct multi-scale comparisons between UAV hyperspectral RGB imagery and Sentinel-2 satellite predictions for crude protein and dry matter digestibility respectively, using two contrasting field sites: a grassland (Baryugal) and a woodland (Mt Pleasant) environment, each captured during both the dry season (October–November 2023) and wet season (February–April 2024) conditions. The UAV imagery shows fine-grained spatial detail at approximately 5 cm resolution, capturing within-paddock heterogeneity in pasture condition. The Sentinel-2 predictions represent 20 m pixel averages, demonstrating the generalisation inherent in satellite-scale monitoring. Despite the resolution difference, the seasonal patterns in pasture nutrition are clearly visible in the S2 predictions. All predictions are masked to pasture pixels only (dry and green pasture classes), excluding tree canopy, bare soil, shadow, and other non-pasture cover types using the DNN classification.

3.4.2. Time-Series Predictions with Cattle Nutrition Thresholds

To demonstrate the operational utility of the TabPFN Sentinel-2 model, we generated time-series predictions for two contrasting field sites: a grassland and a woodland environment. To translate these predictions into actionable management information, we overlaid established cattle nutrition thresholds on the predicted pasture nutrition metrics (Figures 13 and 14).

Crude protein thresholds were established based on beef cattle requirements: 7% CP represents the minimum threshold for maintenance of adult cattle, below which weight loss occurs and supplementation becomes necessary; 10% CP represents the threshold for satisfactory growth performance in growing cattle [27]. For dry matter digestibility, 55% DMD represents the minimum threshold for maintenance intake, while 60% DMD indicates good quality pasture supporting productive performance [28]. Red triangles on the CP panels identify specific observations when supplementation would be recommended.

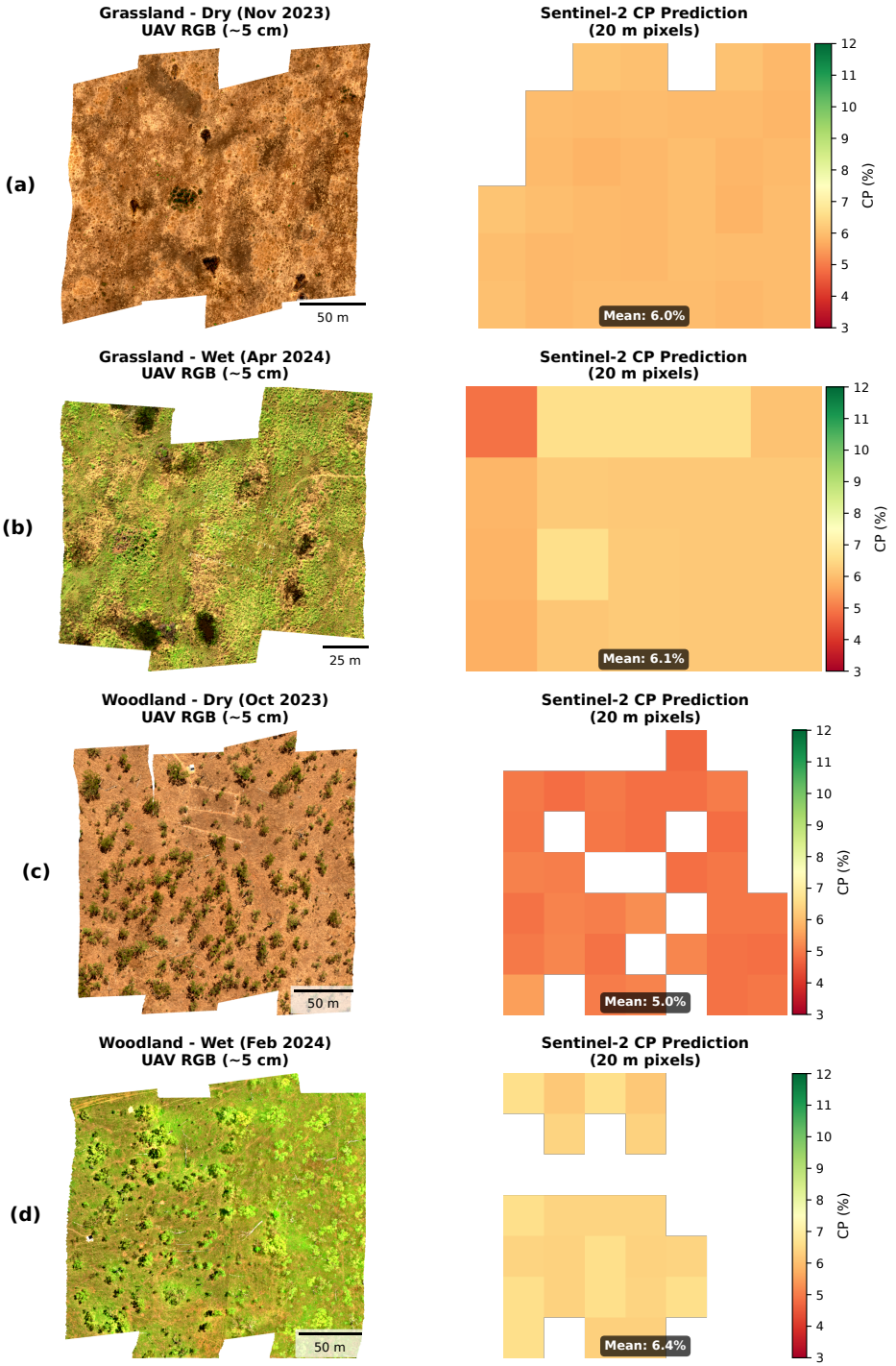

Figure 11. Multi-scale comparison of UAV hyperspectral RGB imagery (∼5 cm) and Sentinel-2 satellite crude protein (CP) predictions (20 m). Columns show: (**1**) UAV RGB imagery, (**2**) Sentinel-2 CP predictions. Rows show two contrasting sites in dry and wet seasons: (**a**) grassland dry (Nov 2023), (**b**) grassland wet (Apr 2024), (**c**) woodland dry (Oct 2023), (**d**) woodland wet (Feb 2024).

The time-series reveal clear seasonal patterns in pasture nutrition that align with the tropical wet–dry climate of the study region. The Baryugal Site 1 grassland (Figure 13) demonstrates characteristic seasonal cycling, with crude protein averaging 7.1 ± 1.7% across 53 observations. Values ranged from 3.9% during the late dry season to 10.4% following wet season rainfall. The threshold analysis clearly identifies dry season months (June–October) when CP falls below the 7% maintenance threshold.

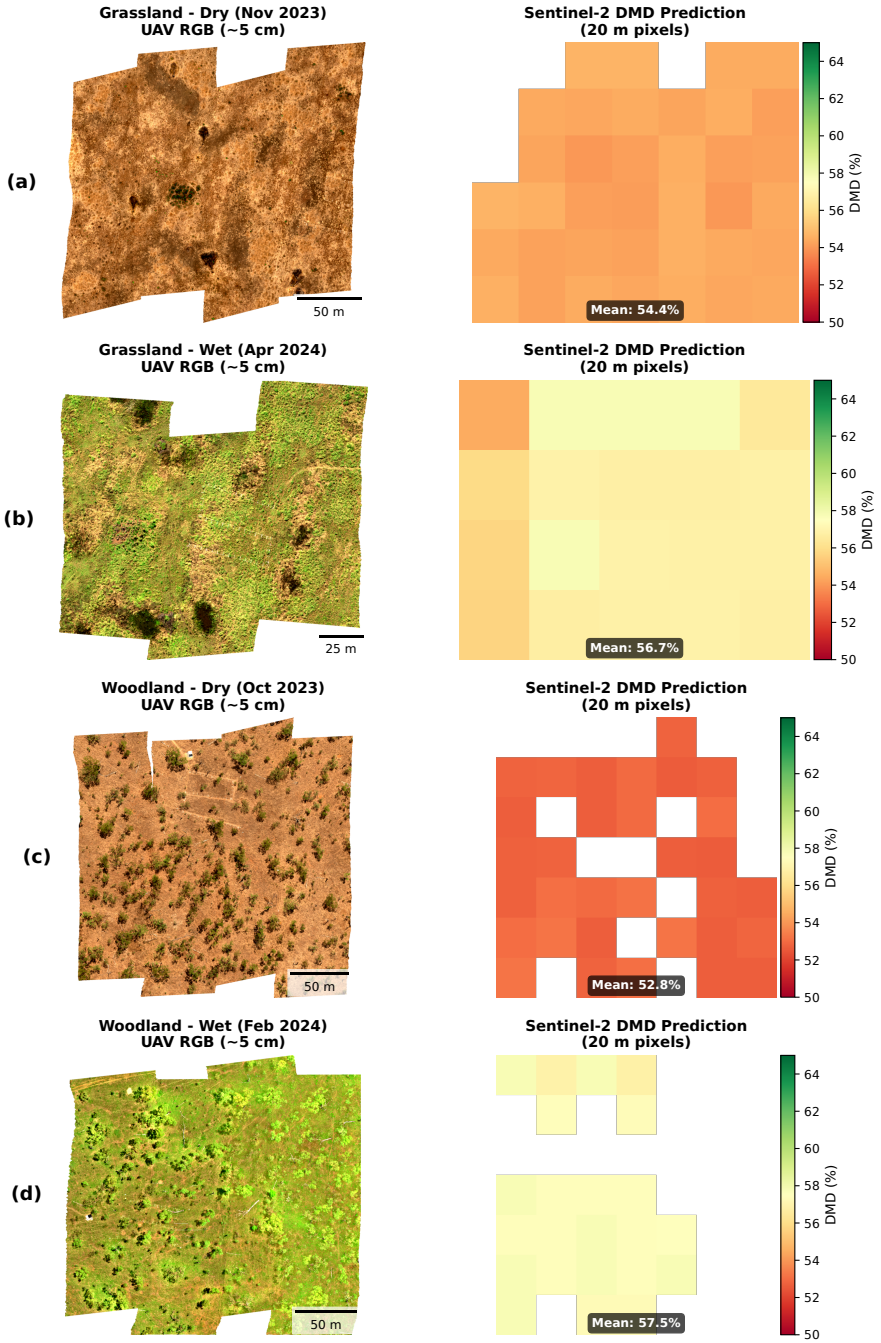

Figure 12. Multi-scale comparison of UAV hyperspectral RGB imagery (∼5 cm) and Sentinel-2 satellite dry matter digestibility (DMD) predictions (20 m). Columns show: (**1**) UAV RGB imagery, (**2**) Sentinel-2 DMD predictions. Rows show two contrasting sites in dry and wet seasons: (**a**) grassland dry (Nov 2023), (**b**) grassland wet (Apr 2024), (**c**) woodland dry (Oct 2023), (**d**) woodland wet (Feb 2024).

The Mt Pleasant Site 1 woodland (Figure 14) exhibited similar seasonal patterns with reduced amplitude in seasonal fluctuations compared to open grassland, attributed to the moderating influence of scattered tree cover on microclimate and soil moisture retention.

These threshold-based visualizations provide producers with intuitive decision support tools for:

- Identifying critical periods when strategic supplementation is warranted;
- Planning proactive management interventions before quality declines below critical thresholds;

- Comparing paddock-level quality to inform rotational grazing decisions;
- Documenting seasonal patterns to guide long-term stocking rate planning.

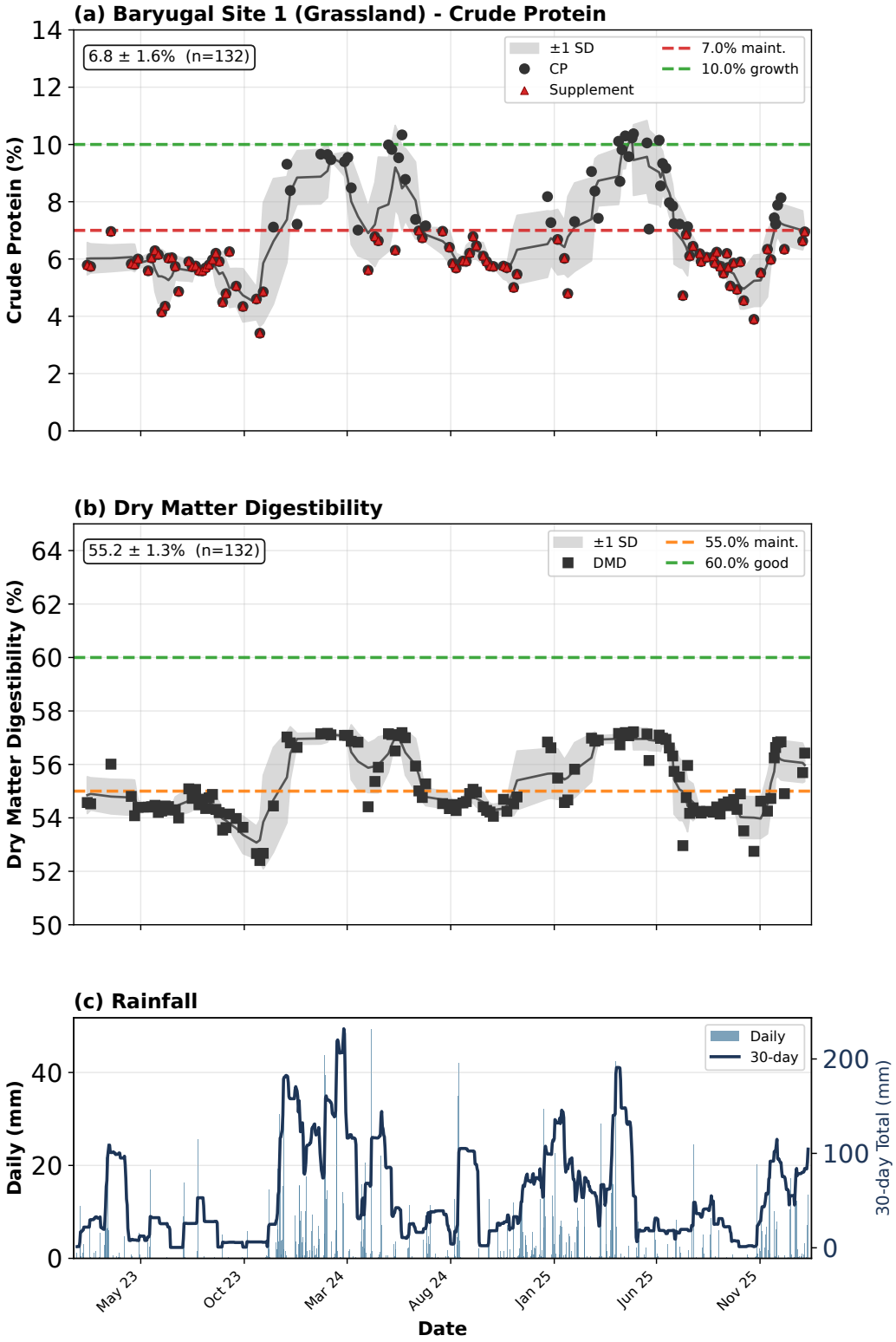

Figure 13. Time-series predictions with cattle nutrition thresholds for Baryugal Site 1 (grassland). Panel (**a**) shows crude protein with threshold lines at 7% (red, maintenance) and 10% (green, good growth); red triangles mark periods when supplementation is recommended. Panel (**b**) shows dry matter digestibility with thresholds at 55% (orange, maintenance) and 60% (green, good quality). Panel (**c**) displays daily rainfall and 30-day cumulative totals.

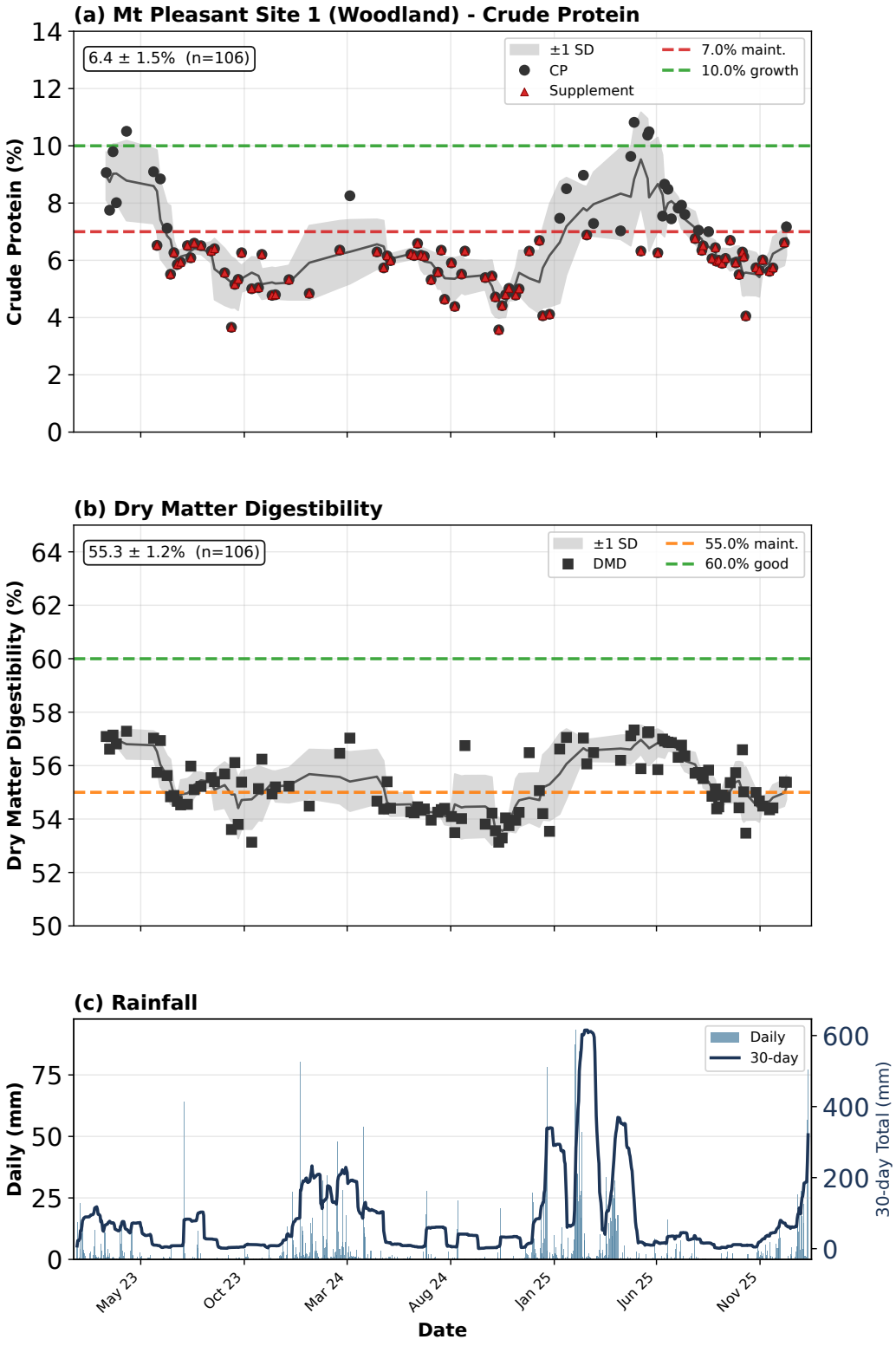

Figure 14. Time-series predictions with cattle nutrition thresholds for Mt Pleasant Site 1 (woodland). Threshold conventions follow Figure 13. The woodland site demonstrates reduced seasonal amplitude due to tree cover effects on microclimate.

4. Discussion

4.1. TabPFN Performance Advantages

The adoption of TabPFN for hierarchical pasture nutrition prediction represents a significant methodological advancement over traditional neural network architectures. TabPFN demonstrated several practical advantages in this remote sensing application:

No hyperparameter tuning: TabPFN required no iterative training epochs or trial-and-error hyperparameter optimization, eliminating a major barrier to reproducible machine learning in remote sensing applications. This contrasts with traditional deep learning approaches that require extensive grid searches over learning rates, layer architectures, dropout rates, and regularization parameters.

Effective with limited data: TabPFN excels with small-to-medium-sized datasets typical of remote sensing calibration studies, where collecting extensive ground truth data is logistically challenging and expensive. With 369 training samples and 93 validation samples across 19 site–date combinations from 11 field sites, the Sentinel-2 model achieved excellent performance (R^2 = 0.92 for CP), demonstrating the effectiveness of hierarchical scaling for pasture nutrition prediction.

Computational efficiency: The model performs inference in a single forward pass, making it computationally efficient and suitable for operational deployment without requiring GPU resources. Training times were reduced from hours (traditional neural networks) to seconds.

Consistent architecture: The same TabPFN configuration (eight ensemble estimators) was applied successfully across all four spatial scales without modification, demonstrating the architecture's flexibility and robustness across different data characteristics.

Comparison with traditional methods: This study did not include formal benchmark comparisons against traditional regression methods such as partial least squares regression (PLSR), random forest, or XGBoost. The original TabPFN publication [17] demonstrated competitive or superior performance compared to gradient boosting methods (XGBoost, CatBoost, LightGBM) across 30 benchmark tabular datasets, particularly for datasets with fewer than 1000 samples—consistent with our calibration dataset sizes. While these published benchmarks provide confidence in TabPFN's predictive capability, direct comparison using our pasture nutrition data would strengthen the validation of the approach. The practical advantages of TabPFN observed in this study—no hyperparameter tuning, sub-second training, and consistent architecture across scales—remain independently verifiable regardless of comparative prediction accuracy.

4.2. Hierarchical Scaling Performance

Nature of the hierarchical framework: It is important to clarify that the term "hierarchical" in this study refers to a cascaded training data approach rather than a joint modelling or parameter-sharing architecture. Separate TabPFN models are trained independently at each spatial scale, with no shared parameters or joint optimisation across scales. Cross-scale information integration occurs through the training data cascade: the Sentinel-2 model is trained on UAV-derived chemistry predictions (pseudo-labels) rather than direct laboratory measurements, effectively transferring spectral–chemistry relationships learned at the UAV scale to the satellite scale. This approach differs from true hierarchical Bayesian models or multi-task learning frameworks where parameters are explicitly shared or jointly estimated. The practical advantage of our cascaded approach is simplicity and modularity—each scale can be updated or recalibrated independently without retraining the entire pipeline. The disadvantage is that cross-scale relationships are not explicitly modelled, and uncertainty propagation between scales is not formally captured (as discussed in Section 4.3). Future work could explore joint multi-scale architectures that simultane-

ously learn spectral–chemistry relationships across resolutions, potentially improving both prediction accuracy and uncertainty quantification.

Model performance varied systematically along the laboratory-to-satellite scaling continuum, with results generally consistent with or exceeding those reported in comparable studies. The Sentinel-2 crude protein validation R^2 of 0.92 compares favourably with [11], who achieved R^2 = 0.70–0.82 for CP estimation in Mediterranean pastures using Sentinel-2 time series, and with [12], who reported similar performance for leaf nitrogen mapping in semi-arid savannas. At the UAV scale, our CP validation R^2 of 0.95 exceeds the R^2 = 0.70–0.78 reported by [9] for crude protein prediction in temperate grasslands using UAV hyperspectral imaging, though differences in pasture type, climate zone, and spectral preprocessing complicate direct comparison. Ref. [10] achieved R^2 = 0.74–0.80 for crude protein prediction in temperate Australian pastures using airborne hyperspectral imagery—comparable to our leaf-scale performance but below our UAV and Sentinel-2 results, suggesting benefits from our hierarchical scaling approach. For DMD, our Sentinel-2 R^2 of 0.73 is consistent with the general finding that fibre-related predictions are more challenging than protein; [15] reported R^2 = 0.55–0.68 for ADF prediction in Mediterranean grasslands using field spectroscopy and machine learning.

The leaf scale showed strong calibration performance (R^2 = 0.92 for CP) but lower validation accuracy, likely due to the spectral complexity of individual leaf samples and the challenge of generalising across diverse species and growth conditions.

The plant point scale exhibited the lowest performance, attributed to the mixing of grass and forb spectral signatures at the quadrat level without the spatial averaging benefits of imaging sensors. Additionally, atmospheric effects present in field measurements that were absent in the controlled laboratory environment contributed to increased spectral noise.

The UAV scale provided an intermediate performance level, benefiting from spatial averaging while maintaining high spectral resolution with 620 bands. The Sentinel-2 scale achieved excellent performance (R^2 = 0.92 for CP, R^2 = 0.73 for DMD), likely due to:

- Larger training sample sizes from extensive UAV coverage;
- Spatial averaging that reduces noise from individual plant variability
- Spectral bands specifically designed and optimized for vegetation monitoring;
- Integration with pasture classification masking that excludes confounding pixels;
- The scaling approach that uses UAV predictions as training targets, effectively calibrating at the appropriate spatial scale.

Spatial autocorrelation and validation independence: The UAV-scale validation R^2 of 0.95 exceeds both the leaf-scale (0.76) and plant-point-scale (0.58) performance, which could raise concerns about spatial autocorrelation inflating validation metrics if nearby samples share similar spectral–chemistry relationships. The train–validation split was performed using stratified random sampling without explicit spatial blocking by site or transect, meaning that training and validation samples from the same site–date combination could be spatially proximate. However, several observations suggest that spatial autocorrelation is not the primary driver of high UAV-scale performance. First, the validation R^2 (0.95) slightly exceeds calibration R^2 (0.94)—if autocorrelation were inflating validation metrics through memorisation of local patterns, we would expect the opposite pattern (calibration > validation). Second, the Sentinel-2 model trained across 11 geographically distributed sites shows *lower* validation R^2 (0.92) than the UAV model despite greater geographic diversity, inconsistent with autocorrelation-driven inflation. Third, the performance improvement from leaf to UAV scale is most readily explained by spatial averaging benefits that reduce individual plant variability and spectral noise. Nevertheless, we acknowledge that leave-site-out cross-

validation would provide stronger evidence for geographic generalisation and recommend this as a priority for future validation studies.

Leaf-scale DMD prediction limitations: The substantially lower validation R^2 for DMD at the leaf scale (0.47) compared to UAV scale (0.88) warrants mechanistic explanation, particularly given that SWIR absorption features associated with fibre compounds should theoretically be detectable at the leaf scale. Several factors contribute to this discrepancy. First, laboratory spectra were collected on fresh leaf material, where strong water absorption features in the SWIR region (1400–1500 nm, 1900–2100 nm) dominate over the more subtle cellulose and lignin absorption features that correlate with ADF content. In contrast, the chemical analysis was performed on oven-dried material, creating a fundamental mismatch between the spectral measurement state and the chemistry reference state. Second, fibre compounds are heterogeneously distributed within individual leaves—concentrated in vascular bundles, midribs, and structural tissues rather than uniformly throughout the mesophyll as with chlorophyll-related nitrogen. This spatial heterogeneity means that individual leaf spectra sample an inconsistent proportion of fibre-rich tissue depending on measurement position. Third, individual leaves exhibit substantial variability in lignification state depending on leaf age, position on the tiller, and species—variability that is averaged out at the UAV scale where hundreds of leaves contribute to each pixel. The UAV-scale improvement from 0.47 to 0.88 R^2 demonstrates that spatial averaging effectively mitigates these individual leaf measurement challenges for fibre-related predictions.

Disentangling spatial and spectral resolution contributions: The superior performance of the UAV-scale model raises the question of whether this stems primarily from high spatial resolution (enabling pure pasture pixel extraction at ∼5 cm) or from the high spectral information content (620 bands reduced to 100 principal components via PCA). Comparison with the Sentinel-2 scale provides insight: crude protein validation R^2 was nearly equivalent between UAV (0.95) and Sentinel-2 (0.92) despite a 10-fold reduction in spectral dimensionality (100 PCA components vs 10 multispectral bands). This suggests that for protein prediction, spatial resolution enabling pure pasture pixel extraction may be more critical than hyperspectral information content. The strong correlation between protein and chlorophyll-related absorption features in the red-edge region—well captured by Sentinel-2's dedicated red-edge bands—likely explains the minimal performance loss.

In contrast, DMD validation R^2 decreased more substantially from UAV (0.88) to Sentinel-2 (0.73), suggesting that fibre-related compounds rely more heavily on subtle SWIR absorption features that are better resolved by hyperspectral measurements. The broader SWIR bands of Sentinel-2 may inadequately capture the narrow absorption features associated with cellulose and lignin that contribute to digestibility predictions. Additionally, the removal of six UAV-scale DMD outliers due to soil reflectance contamination in sparse vegetation quadrats demonstrates the importance of spatial resolution for avoiding spectral mixing artefacts. At the Sentinel-2 scale (10–20 m pixels), such contamination cannot be avoided through pixel selection alone, necessitating the pasture classification masking approach to exclude non-pasture pixels.

PCA variance retention considerations: The choice of 100 principal components retaining 99.9% of spectral variance was primarily driven by TabPFN's architectural constraint requiring fewer than 200 input features. A formal ablation study evaluating different variance-retention thresholds was not conducted, and we acknowledge that retaining such high variance may include noisy dimensions that could potentially degrade model generalisation. However, indirect evidence suggests the choice was reasonable: the Sentinel-2 model achieves comparable CP prediction performance (R^2 = 0.92) using only 10 spectral bands versus 100 PCA components at the UAV scale (R^2 = 0.95). This minimal performance difference indicates that TabPFN can effectively extract predictive signal regardless

of whether input dimensionality is 10 or 100, suggesting that the model is robust to the inclusion of low-variance components. Nevertheless, systematic evaluation of variance thresholds (e.g., 95%, 99%, 99.9%) and alternative feature selection approaches (e.g., band selection based on spectral–chemistry correlations) represents a valuable direction for optimising hyperspectral preprocessing pipelines.

Model interpretability and learned relationships: The concern that high R^2 from only 10 Sentinel-2 bands may reflect model fitting bias rather than genuine spectral–chemistry relationships is addressed by the spectral–chemistry correlation analysis presented in Section 3.2.2. The identified correlations align with established biochemical absorption features [26] and map directly onto Sentinel-2 band positions: B04 (665 nm) captures chlorophyll absorption ($r = -0.48$ with CP), B05–B07 (705–783 nm) span the red-edge region correlated with nitrogen status, B08/B8A (842–865 nm) sample the NIR plateau ($r = 0.35$–0.45 with CP), and B11–B12 (1610–2190 nm) capture SWIR protein and fibre absorption features ($r = -0.38$ to -0.45 with DMD). The Sentinel-2 bands were specifically designed for vegetation monitoring and intentionally positioned at these physiologically informative wavelengths. While formal SHAP analysis or permutation feature importance was not conducted—partly due to TabPFN's in-context learning architecture which complicates traditional feature attribution methods—the alignment between independent spectral–chemistry correlations and Sentinel-2 band positions provides confidence that model predictions are based on genuine biochemical relationships rather than spurious patterns.

4.3. Pasture Classification Masking Benefits

The implementation of pasture classification masking provided tangible improvements to chemistry prediction quality. By restricting predictions to classified pasture pixels, the models avoided spectral contamination from:

- Bare soil and rocky outcrops that would bias CP predictions low;
- Woody vegetation canopies that would bias CP predictions high;
- Water bodies and infrastructure.

In field trials, approximately 79% of pixels within typical pasture sites were classified as pasture, with the remaining 21% representing non-productive areas appropriately excluded from chemistry calculations.

Quantitative evidence for target purity importance: While a formal ablation study comparing masked versus unmasked Sentinel-2 model performance was not conducted, indirect quantitative evidence strongly supports the critical importance of target purity for reliable chemistry retrieval. At the UAV scale, six samples with soil reflectance contamination caused DMD validation R^2 to collapse from 0.88 to 0.19—a near-total loss of predictive skill from just 3.4% contaminated samples. Similarly, woody vegetation canopies exhibited CP predictions 3–5 times higher than adjacent pastures; without masking, these signals would catastrophically bias paddock-scale chemistry estimates. The 21% of pixels excluded by classification in typical heterogeneous sites represents a substantial proportion of confounding signals that would otherwise degrade model accuracy. These observations demonstrate that target purity is not merely beneficial but essential for quantitative chemistry retrieval from satellite imagery in heterogeneous landscapes.

This masking approach is particularly valuable for time-series analysis, where consistent spatial sampling across the image archive ensures that observed temporal patterns in chemistry reflect genuine pasture quality changes rather than variations in the proportion of non-pasture pixels captured under different viewing conditions or seasonal bare ground exposure.

Classification data partitioning: The 98.6% test accuracy for the DNN land cover classifier was achieved using random stratified sampling within and across images, without

explicit separation by site or acquisition date. This within-image sampling may result in test pixels that are spatially proximate to training pixels, potentially allowing spatial autocorrelation to inflate accuracy metrics. Nevertheless, the classification model serves as a preprocessing mask rather than a primary research finding—its role is to identify pasture pixels for subsequent chemistry prediction rather than to provide stand-alone land cover mapping. Classification errors would primarily affect which pixels receive chemistry predictions (commission/omission of pasture pixels) rather than the accuracy of chemistry predictions themselves, which are validated independently. The strong per-class F1-scores (0.96–0.99) across spectrally distinct classes suggest that the classifier reliably distinguishes pasture from non-pasture targets, which is the operationally critical capability.

4.4. Challenges and Limitations

Several challenges affected model performance across scales:

DMD prediction difficulty: DMD predictions showed consistently lower accuracy than CP predictions across all scales, a pattern consistently observed across the pasture remote sensing literature. Ref. [9] similarly reported lower prediction accuracy for acid detergent fibre (R^2 = 0.64) compared to crude protein (R^2 = 0.78) in temperate grasslands, while [10] found that digestibility and lignin were more challenging to predict than crude protein from airborne hyperspectral imagery. This consistent pattern across studies reflects fundamental spectral–chemical relationships: protein correlates strongly with chlorophyll absorption features that are spectrally prominent, whereas fibre compounds (cellulose, hemicellulose, lignin) produce subtler absorption features in the SWIR region that are more easily confounded by soil background and moisture variations. At the UAV scale, six outlier samples (3.4% of total) were removed where soil reflectance contamination was evident, substantially improving model performance (validation R^2 from 0.19 to 0.88). These samples originated from quadrats with sparse vegetation cover where exposed soil dominated the spectral response, breaking the relationship between measured plant chemistry and observed spectra.

Outlier exclusion and model generalisability: The removal of six soil-contaminated samples for DMD modelling warrants discussion regarding potential selection bias and operational implications. The exclusion criteria were based on objective spectral characteristics (soil-dominated reflectance in sparse vegetation quadrats) rather than model performance alone—the same samples remained valid for CP prediction, demonstrating that the issue was specific to soil–fibre spectral confusion rather than general sample quality. The dramatic R^2 improvement (0.19 $\rightarrow$ 0.88) from excluding just 3.4% of samples illustrates both the sensitivity of DMD retrieval to spectral purity and the necessity of adequate vegetation cover for reliable predictions. For operational deployment, the pasture classification masking approach addresses this limitation by excluding pixels with insufficient vegetation fractional cover before chemistry prediction, preventing similar soil-contaminated spectra from entering the prediction pipeline. Users should interpret DMD predictions with caution in areas with sparse vegetation cover, where model uncertainty will be elevated.

Wooded vegetation influence: In woodland field sites, the presence of tree and shrub canopy significantly complicated predictions. Wooded vegetation canopies exhibited higher CP predictions (three to five times that of adjacent pastures), potentially inflating paddock-scale averages where masking was incomplete.

Temporal alignment: Ground-based measurements were not always coincident with satellite overpasses. While efforts were made to minimize this gap, rapid changes in pasture condition following rainfall events may have introduced calibration errors.

Species composition effects: The presence of nitrogen-fixing legume species (particularly Seca Stylo) introduced additional variability in CP content that was not fully captured by spectral signatures alone.

Geographic scope: While the study encompassed three bioregions (Desert Uplands, Brigalow Belt, Southeast Queensland) spanning approximately 600 km of latitude and captured diverse climate zones, vegetation groups, and pasture species assemblages, the validation remains confined to Queensland's Great Barrier Reef catchment. This builds upon the regional foundation established by [2], who previously demonstrated UAV-based pasture quality estimation across Queensland rangelands. The intensive hierarchical sampling strategy—requiring coordinated laboratory spectroscopy, field measurements, UAV flights, and satellite acquisitions at each site—necessarily constrained the number of properties that could be included within available resources. Similar geographic constraints are evident across the literature: Reference [11] validated across a single Mediterranean catchment, while [9] focused on temperate European grasslands. The intensive approach was essential for establishing and validating the multi-scale modelling framework, but broader geographic validation across additional Australian rangelands and internationally remains necessary to confirm transferability of the calibration models. The TabPFN architecture's demonstrated ability to generalise from limited training data provides a foundation for such expansion with minimal additional calibration effort.

Error propagation in hierarchical modelling: The Sentinel-2 canopy-scale model was trained using UAV-derived chemistry predictions rather than direct laboratory measurements, introducing the potential for error compounding through the modelling chain. Specifically, prediction errors from the UAV-scale model (validation R^2 = 0.95 for CP, R^2 = 0.88 for DMD) propagate into the satellite-scale training data. The Sentinel-2 validation performance (R^2 = 0.92 for CP, R^2 = 0.73 for DMD) demonstrates that the hierarchical approach is effective, with comparable performance maintained at the satellite scale despite the resolution change. Several factors help mitigate complete error compounding: (1) *Spatial averaging effect:* Each Sentinel-2 pixel aggregates predictions from thousands of UAV pixels (∼40,000 at 5 cm resolution within a 10 m pixel), and random prediction errors tend to cancel when averaged, reducing systematic bias while preserving the underlying spatial signal. (2) *Increased training sample size:* The UAV-to-Sentinel-2 scaling generated 462 training samples from relatively few field campaigns, compared to the 180 UAV-scale samples with direct chemistry linkage, providing the TabPFN model with substantially more calibration data. (3) *Reduced noise from fine-scale heterogeneity:* The coarser Sentinel-2 resolution naturally smooths over within-quadrat variability in plant chemistry that introduces noise at finer scales. (4) *Consistent spectral–chemistry relationships:* The fundamental relationship between spectral reflectance and foliar chemistry is preserved across scales, and TabPFN learns this relationship effectively regardless of intermediate processing steps.

Cross-scale uncertainty accumulation: However, a critical limitation of the current uncertainty framework must be acknowledged: the reported prediction intervals ($\hat{\sigma}$) at each scale are derived independently from each model's quantile predictions and do not capture the cumulative uncertainty from the hierarchical modelling chain. The counter-intuitive pattern of decreasing $\hat{\sigma}$ with increasing scale (CP: 1.40% leaf → 0.79% UAV → 0.39% Sentinel-2; DMD: 2.60% leaf → 1.05% UAV → 0.95% Sentinel-2) reflects the variance-reducing effect of spatial averaging rather than a genuine reduction in prediction uncertainty relative to true chemistry values. For the Sentinel-2 model specifically, the $\hat{\sigma}$ values represent uncertainty in predicting UAV-derived pseudo-labels rather than true laboratory chemistry measurements, potentially underestimating total uncertainty. A rigorous treatment of cross-scale uncertainty would require either: (1) Monte Carlo error propagation through the modelling chain, sampling from each model's prediction distribution; or (2) analytical

uncertainty propagation accounting for covariance structures between scales. Such analysis was beyond the scope of this study but represents an important direction for operational deployment where confidence intervals directly inform management decisions. Users should interpret Sentinel-2 uncertainty estimates as conditional on UAV model accuracy, with true uncertainty relative to laboratory chemistry being somewhat larger than the reported $\hat{\sigma}$ values suggest.

Faecal near-infrared spectroscopy (F.NIRS) represents a well-established alternative method for estimating diet quality in grazing livestock [29,30]. F.NIRS analyzes the spectral characteristics of faecal samples to infer the nutritional quality of ingested forage, and has been widely adopted in northern Australian rangeland research. A comparison of our remote sensing approach with F.NIRS reveals complementary strengths and limitations.

Advantages of remote sensing over F.NIRS:

- **Spatial coverage:** Satellite-based predictions provide wall-to-wall spatial coverage across entire properties, enabling identification of paddock-scale heterogeneity in pasture quality. F.NIRS provides only a herd-averaged estimate with no spatial discrimination.
- **No animal handling required:** Our method requires no animal contact, faecal collection, or laboratory analysis, reducing operational costs and eliminating animal welfare considerations associated with mustering for sample collection.
- **Predictive capability:** Remote sensing can assess pasture quality before animals are introduced to a paddock, supporting proactive grazing management decisions. F.NIRS can only indicate diet quality after animals have consumed the pasture.
- **Temporal frequency:** Sentinel-2 provides observations every 5 days under cloud-free conditions, enabling near-real-time monitoring. F.NIRS requires repeated mustering for ongoing monitoring.
- **Historical analysis:** The Sentinel-2 archive (2018–present) enables retrospective analysis of pasture quality dynamics. No equivalent historical record exists for F.NIRS unless systematic sampling programs have been maintained.

Advantages of F.NIRS over remote sensing:

- **Measures actual diet:** F.NIRS measures what animals actually consume, which may differ from the average pasture composition due to selective grazing behaviour. Cattle preferentially graze higher quality pasture components, so F.NIRS estimates often exceed paddock-average values from remote sensing.
- **Not affected by atmospheric conditions:** F.NIRS is independent of cloud cover, atmospheric interference, and solar illumination conditions that can limit satellite observations, particularly during the wet season when pasture quality monitoring is most critical.
- **Established calibration databases:** F.NIRS benefits from decades of calibration development with extensive reference datasets across diverse pasture types. Our satellite-based approach requires continued calibration expansion to match this breadth.
- **Species-mixture integration:** F.NIRS inherently integrates the contribution of different pasture species in proportion to their consumption. Remote sensing may over- or under-represent certain species depending on their spectral dominance and spatial distribution within pixels.

Complementary applications: The two approaches may be most valuable when used in combination. Remote sensing can identify spatial patterns and temporal trends in pasture quality across properties, while periodic F.NIRS sampling can validate these predictions against actual diet quality and account for selective grazing behaviour. This integrated

approach could provide both the spatial awareness needed for paddock selection decisions and the animal-centric validation required for nutritional management.

4.5. Operational Scalability

The methodology demonstrates operational scalability using:

- Freely available Sentinel-2 imagery with global coverage and a 5-day repeat cycle;
- Open-source Python libraries (TabPFN, scikit-learn, rasterio, geopandas);
- Standard computational resources without GPU requirements;
- Automated time-series generation from image archives (2018–present).

The complete archive of Sentinel-2 imagery (2016–present) enables historical analysis of pasture quality dynamics, supporting evidence-based grazing management decisions. Producers can receive both visual and numerical indications of pasture quality through spatial summaries showing the proportion of paddock area exceeding particular thresholds.

5. Conclusions

5.1. Summary of Research Objectives

This study addressed five key research objectives related to satellite-based pasture nutrition monitoring:

Objective 1—Multi-scale spectral characterisation: Spectral responses were successfully characterised across leaf, plant point, UAV, and satellite scales. Results demonstrated that imaging scales (UAV and satellite) provided the strongest correlations with pasture nutrition metrics, with UAV achieving the highest crude protein accuracy (R^2 = 0.95) and Sentinel-2 achieving comparable performance (R^2 = 0.92) despite reduced spectral resolution. The spatial averaging effect of larger pixel sizes reduces noise from individual plant variability. Key spectral features were identified in the red-edge and SWIR regions for protein, and in water absorption bands for digestibility-related metrics.

Objective 2—TabPFN model development: TabPFN-based models achieved excellent performance for satellite-scale predictions, with R^2 = 0.92 for crude protein and R^2 = 0.73 for dry matter digestibility. The TabPFN architecture demonstrated particular advantages for the small-to-medium-sized calibration datasets typical of remote sensing studies (369 training, 93 validation samples across 19 site–date combinations from 11 field sites), requiring no hyperparameter tuning while providing calibrated uncertainty estimates.

Objective 3—Pasture classification for target purity: A deep neural network classification model was developed as an essential preprocessing step to ensure spectral target purity by distinguishing pasture pixels from woody vegetation, bare soil, and shadow. The classifier achieved 98.6% overall accuracy across seven land cover classes. Integration of this classification mask is critical for reliable chemistry retrieval, as indirect evidence demonstrates that even small proportions of contaminated pixels (3.4% at UAV scale) can cause near-complete loss of predictive skill for DMD. The 21% of non-pasture pixels excluded in typical heterogeneous sites would otherwise catastrophically bias chemistry estimates.

Objective 4—Temporal dynamics assessment: Time series analysis across multiple field sites demonstrated that predicted pasture chemistry metrics exhibited seasonal patterns consistent with expected physiological responses to environmental conditions.

Objective 5—Operational scalability demonstration: The methodology was successfully applied to generate time-series predictions across field sites using freely available Sentinel-2 imagery (2018–present). The open-source implementation using Python and TabPFN demonstrates feasibility for operational deployment.

5.2. Key Achievements and Future Directions

This study successfully demonstrated that multispectral spectroscopy combined with TabPFN machine learning is a cost-effective and efficient alternative to traditional wet chemistry for predicting pasture chemical components. Key achievements include:

1. **Multi-scale TabPFN implementation:** TabPFN models were successfully developed for leaf (R^2 = 0.92/0.76), plant point (R^2 = 0.74/0.58), UAV (R^2 = 0.94/0.95 for CP, R^2 = 0.89/0.88 for DMD), and Sentinel-2 (R^2 = 0.92 validation) scales for crude protein prediction.
2. **Excellent satellite-scale performance:** The Sentinel-2 TabPFN model achieved excellent accuracy (R^2 = 0.92 for CP, R^2 = 0.73 for DMD) without hyperparameter tuning, demonstrating the effectiveness of hierarchical scaling and TabPFN's ability to generalise from limited training data.
3. **Effective pasture masking:** Integration of deep neural network classification masking (98.6% accuracy, seven classes) improved prediction quality by excluding non-pasture spectral signals.
4. **Validated temporal dynamics:** Time-series predictions exhibited appropriate seasonal patterns consistent with expected physiological responses to rainfall.

Future research should focus on:

- **Pasture species influence:** Investigating how improved pasture species (particularly nitrogen-fixing legumes) affect overall nutrient composition and characterizing their spectral signatures.
- **Spectral response scaling:** Further understanding why spectral feature strengths differ across scales and developing methods to account for spectral mixing.
- **Expanded nutrition metrics:** Extending assessments beyond crude protein and digestibility to include palatability indices and pasture species differentiation.
- **Improved temporal coincidence:** Collecting field measurements that better coincide with satellite overpasses.
- **Background contamination mitigation:** Advancing the pasture classification approach to better separate pasture responses from woody vegetation and background soil.
- **Classification ablation analysis:** Conducting formal masked versus unmasked model comparison studies to quantitatively demonstrate the improvement in Sentinel-2 chemistry prediction accuracy attributable to pasture classification masking.
- **Traditional method benchmarking:** Conducting formal comparisons of TabPFN against partial least squares regression (PLSR), random forest, and gradient boosting methods (XGBoost, CatBoost) at each spatial scale to quantitatively validate the claimed advantages and identify scenarios where alternative approaches may be preferable.
- **Model interpretability analysis:** Implementing SHAP (Shapley additive explanations) or permutation feature importance analysis, potentially using surrogate model approaches compatible with TabPFN's transformer architecture, to quantitatively verify the contribution of individual Sentinel-2 bands to chemistry predictions.
- **Spatial cross-validation:** Implementing leave-site-out cross-validation to rigorously assess model generalisation across independent geographic locations and quantify the contribution of spatial autocorrelation to current validation metrics.
- **PCA threshold optimisation:** Systematically evaluating different variance-retention thresholds (e.g., 95%, 99%, 99.9%) for PCA dimensionality reduction and comparing against alternative feature selection approaches such as correlation-based band selec-

tion or recursive feature elimination to identify optimal preprocessing strategies for hyperspectral chemistry prediction.

- **Cross-scale uncertainty propagation:** Implementing Monte Carlo error propagation or analytical uncertainty frameworks to quantify cumulative prediction uncertainty through the hierarchical modelling chain, providing Sentinel-2 uncertainty estimates that account for UAV model error rather than treating each scale independently.
- **Joint multi-scale modelling:** Developing true hierarchical or multi-task learning architectures that jointly optimise spectral–chemistry relationships across spatial scales, potentially sharing parameters between scale-specific models or implementing explicit cross-scale constraints to improve coherence and enable formal uncertainty propagation.
- **Geographic expansion:** Extending validation to broader geographic areas, including more arid inland regions, temperate pasture systems, and international rangelands with different pasture species assemblages. Priority should be given to establishing independent validation sites outside the current GBR catchment training domain to assess model transferability and identify regional calibration requirements.

Ultimately, improved estimations of pasture chemistry composition with precise spatial and temporal data can significantly enhance the efficiency, cost-effectiveness, and sustainability of rangeland grazing across Queensland and similar tropical and sub-tropical grazing systems globally. The TabPFN-based hierarchical spectral modelling approach and pasture classification masking methodology developed in this study provide a foundation for operational pasture nutrition monitoring services.

Author Contributions: Conceptualization, J.B. and H.R.P.; methodology, J.B.; software, J.B.; validation, J.B. and H.R.P.; formal analysis, J.B.; investigation, J.B.; resources, G.F.; data curation, J.B. and H.R.P.; writing—original draft preparation, J.B.; writing—review and editing, J.B., H.R.P. and G.F.; visualization, J.B., H.R.P. and G.F.; project administration, G.F.; funding acquisition, G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded equally by the Queensland Department of the Environment, Tourism and Innovation division of both the Earth Observation and Social Sciences and the Office of the Great Barrier Reef Divisions and the Queensland Department of Primary Industries, Drought and Climate Adaptability Program division.

Data Availability Statement: The datasets presented in this article are not readily available because the field data were collected on private pastoral properties under confidentiality agreements with landholders. Requests to access the datasets should be directed to the corresponding author. The Sentinel-2 satellite imagery used in this study is freely available from the Copernicus Open Access Hub (https://scihub.copernicus.eu/, accessed on 15 March 2026).

Acknowledgments: The anonymous reviewers for their review and feedback of the manuscript. The Landsberg and Cuddihy families and staff of the Trafalgar and Baryugal pastoral properties, the Gordon-Moulin family of the Mount Pleasant pastoral property, the Hawkins family of the Bon Accord pastoral property and the Jess family of the Esk pastoral property, that all kindly provided support and access to their properties. Staff at the Eco-sciences Precinct Grazing and Land Systems, Remote Sensing Sciences centre and Animal Sciences for assistance and support including: Christina Jones, Grant Stone, John Carter, Dan Tindall and Giselle Whish and assistance in collecting field data from Rebecca Farrell and Thomas Franz.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ADF	Acid Detergent Fibre
BRDF	Bi-Directional Reflectance Distribution Function
CAI	Cellulose Absorption Index
CP	Crude Protein
DMD	Dry Matter Digestibility
DSM	Digital Surface Model
EVI	Enhanced Vegetation Index
GBR	Great Barrier Reef
GNSS	Global Navigation Satellite System
IMU	Inertial Measurement Unit
ML	Machine Learning
NDVI	Normalized Difference Vegetation Index
NIR	Near Infra-Red
PCA	Principal Component Analysis
PPK	Post-Processed Kinematic
SWIR	Short-Wave Infra-Red
TabPFN	Tabular Prior-Data Fitted Network
TN	Total Nitrogen
UAV	Un-crewed Aerial Vehicle
VNIR	Visible Near-Infra-Red

References

1. Monsalve, M.O.; Monroy, T.R.; Galeano-Vasco, L.F.; Medina-Sierra, M.; Ceron-Munoz, M.F. Determination of Grass Quality Using Spectroscopy: Advances and Perspectives. In *Grasslands*; Iqbal, M.A., Ed.; IntechOpen: Rijeka, Croatia, 2023; Chapter 6. https://doi.org/10.5772/intechopen.112990.
2. Barnetson, J.; Phinn, S.; Scarth, P. Estimating Plant Pasture Biomass and Quality from UAV Imaging across Queensland's Rangelands. *AgriEngineering* **2020**, *2*, 523–543. https://doi.org/10.3390/agriengineering2040035.
3. McKeon, G.M.; Day, K.A.; Howden, S.M.; Mott, J.J.; Orr, D.M.; Scattini, W.J.; Weston, E.J. Northern Australian Savannas: Management for Pastoral Production. *J. Biogeogr.* **1990**, *17*, 355–372. https://doi.org/10.2307/2845365.
4. Adão, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J.J. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote. Sens.* **2017**, *9*, 1110. https://doi.org/10.3390/rs9111110.
5. Germer, T.A.; Zwinkels, J.C.; Tsai, B.K. Theoretical Concepts in Spectrophotometric Measurements. *Exp. Methods Phys. Sci.* **2014**, *46*, 11–66. https://doi.org/10.1016/B978-0-12-386022-4.00002-9.
6. Ustin, S.; Jacquemoud, S. How the Optical Properties of Leaves Modify the Absorption and Scattering of Energy and Enhance Leaf Functionality. In *Remote Sensing of Plant Biodiversity*; Cavender-Bares, J.; Gamon, J.A.; Townsend, P.A., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 349–384. https://doi.org/10.1007/978-3-030-33157-3_14.
7. Mutanga, O.; Skidmore, A.K. Red edge shift and biochemical content in grass canopies. *ISPRS J. Photogramm. Remote. Sens.* **2007**, *62*, 34–42. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2007.02.001.
8. Buitrago, M.F.; Skidmore, A.K.; Groen, T.A.; Hecker, C.A. Connecting infrared spectra with plant traits to identify species. *ISPRS J. Photogramm. Remote. Sens.* **2018**, *139*, 183–200. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2018.03.013.
9. Wijesingha, J.; Astor, T.; Schulze-Brüninghoff, D.; Wengert, M.; Wachendorf, M. Predicting Forage Quality of Grasslands Using UAV-Borne Imaging Spectroscopy. *Remote Sens.* **2020**, *12*, 126. https://doi.org/10.3390/rs12010126.
10. Thulin, S.; Hill, M.J.; Held, A.; Jones, S.; Woodgate, P. Predicting Levels of Crude Protein, Digestibility, Lignin and Cellulose in Temperate Pastures Using Hyperspectral Image Data. *Am. J. Plant Sci.* **2014**, *05*, 997–1019. https://doi.org/10.4236/ajps.2014.57113.
11. Lugassi, R.; Zaady, E.; Goldshleger, N.; Shoshany, M.; Chudnovsky, A. Spatial and Temporal Monitoring of Pasture Ecological Quality: Sentinel-2-Based Estimation of Crude Protein and Neutral Detergent Fiber Contents. *Remote Sens.* **2019**, *11*, 799. https://doi.org/10.3390/rs11070799.
12. Ramoelo, A.; Cho, M. Explaining Leaf Nitrogen Distribution in a Semi-Arid Environment Predicted on Sentinel-2 Imagery Using a Field Spectroscopy Derived Model. *Remote Sens.* **2018**, *10*, 269. https://doi.org/10.3390/rs10020269.
13. Puliti, S.; Saarela, S.; Gobakken, T.; Ståhl, G.; Næsset, E. Combining UAV and Sentinel-2 auxiliary data for forest growing stock volume estimation through hierarchical model-based inference. *Remote Sens. Environ.* **2018**. https://doi.org/10.1016/j.rse.2017.10.007.

14. Dehghan-Shoar, M.H.; Kereszturi, G.; Pullanagari, R.R.; Orsi, A.A.; Yule, I.J.; Hanly, J. A physically informed multi-scale deep neural network for estimating foliar nitrogen concentration in vegetation. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *130*, 103917. https://doi.org/10.1016/J.JAG.2024.103917.
15. Fernández-Habas, J.; Carriere Cañada, M.; García Moreno, A.M.; Leal-Murillo, J.R.; González-Dugo, M.P.; Abellanas Oar, B.; Gómez-Giráldez, P.J.; Fernández-Rebollo, P. Estimating pasture quality of Mediterranean grasslands using hyperspectral narrow bands from field spectroscopy by Random Forest and PLS regressions. *Comput. Electron. Agric.* **2022**, *192*, 106614. https://doi.org/10.1016/J.COMPAG.2021.106614.
16. Verrelst, J.; Camps-Valls, G.; Muñoz-Marí, J.; Rivera, J.P.; Veroustraete, F.; Clevers, J.G.; Moreno, J. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—A review. *ISPRS J. Photogramm. Remote. Sens.* **2015**, *108*, 273–290. https://doi.org/10.1016/J.ISPRSJPRS.2015.05.005.
17. Hollmann, N.; Müller, S.; Purucker, L.; Krishnakumar, A.; Körfer, M.; Shi, S.; Wu, J.; Zhuang, L.; Validov, V.; Tamasi, I.; et al. Accurate predictions on small data with a tabular foundation model. *Nature* **2025**, *637*, 319–326. https://doi.org/10.1038/s41586-024-08328-6.
18. Australian & Queensland Governments. *reefReportCard*; Queensland Government: Brisbane, Australia, 2023.
19. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. https://doi.org/10.1029/2005RG000183.
20. Van Soest, P.J.; Robertson, J.B.; Lewis, B.A. Methods for Dietary Fiber, Neutral Detergent Fiber, and Nonstarch Polysaccharides in Relation to Animal Nutrition. *J. Dairy Sci.* **1991**, *74*, 3583–3597. https://doi.org/10.3168/jds.S0022-0302(91)78551-2.
21. Möller, J. Cereals, cereals-based products and animal feeding stuffs–determination of crude fat and total fat content by the Randall extraction method: A collaborative study. *Qual. Assur. Saf. Crop. Foods* **2010**, *2*, 197–202.
22. Jones, D.B. *Factors for Converting Percentages of Nitrogen in Foods and Feeds into Percentages of Proteins*; USDA Circular: Washington, DC, USA, 1931; pp. 1–22.
23. Minson, D.J. *Forage in Ruminant Nutrition*; Academic Press: San Diego, CA, USA, 1990; p. 483.
24. Norsk Elektro Optikk AS. *DROACOR: Drone Atmospheric Correction Software*; ReSe Applications LLC: Wil, Switzerland, 2024.
25. Schläpfer, D.; Richter, R.; Feingersh, T. Operational BRDF Effects Correction for Wide-Field-of-View Optical Scanners (BREFCOR). *IEEE Trans. Geosci. Remote. Sens.* **2015**, *53*, 1855–1864. https://doi.org/10.1109/TGRS.2014.2349946.
26. Curran, P.J. Remote sensing of foliar chemistry. *Remote. Sens. Environ.* **1989**, *30*, 271–278. https://doi.org/10.1016/0034-4257(89)90069-2.
27. Meat and Livestock Australia. *Feeding Cattle: A Guide to Protein and Energy Supplementation*; Technical Report; Meat and Livestock Australia: North Sydney, NSW, Australia, 2022.
28. CSIRO. *Nutrient Requirements of Domesticated Ruminants*; CSIRO Publishing: Collingwood, VIC, Australia, 2007.
29. Villamuelas, M.; Serrano, E.; Espunyes, J.; Fernández, N.; López-Olvera, J.R.; Garel, M.; Santos, J.; Parra-Aguado, M.Á.; Ramanzin, M.; Fernández-Aguilar, X.; et al. Predicting herbivore faecal nitrogen using a multispecies near-infrared reflectance spectroscopy calibration. *PLoS ONE* **2017**, *12*, e0176635. https://doi.org/10.1371/journal.pone.0176635.
30. Coates, D.B.; Dixon, R.M. Faecal near infrared reflectance spectroscopy (F.NIRS) measurements of non-grass proportions in the diet of cattle grazing tropical rangelands. *Rangel. J.* **2007**, *29*, 51–63.