Machine Learning Approaches for Climate Risk Assessment in Australian Agriculture

Sarah Chen

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Computer Science and Engineering
Faculty of Engineering
The University of New South Wales

March 2025

Supervisors: Dr. Michael Torres Prof. Elena Vasquez

Abstract

Abstract

Climate variability poses an escalating threat to agricultural productivity in Australia, where the sector contributes over $60 billion annually to the national economy and supports the livelihoods of more than 300,000 people. The Murray–Darling Basin alone accounts for approximately 40% of national agricultural output, yet its semi-arid climate renders crop yields acutely sensitive to shifts in temperature and rainfall patterns. Traditional statistical forecasting methods, grounded in linear assumptions, systematically underestimate the nonlinear feedback loops embedded in the atmosphere–ocean–land system.

This thesis develops a suite of machine learning frameworks tailored to the spatiotemporal complexity of Australian climate data. Drawing on a curated dataset of historical weather records spanning 1980–2023, provided by Nexa Analytics under a research partnership with the Bureau of Meteorology, the study evaluates Random Forest, XGBoost, LSTM, and CNN-LSTM architectures against a bespoke ensemble model. The ensemble approach, which combines gradient-boosted trees with a convolutional recurrent network, achieves a root mean square error of 7.8 mm for seasonal rainfall prediction — a 37% improvement over the current operational benchmark. Training and hyperparameter optimisation were conducted on the Synapse AI high-performance computing cluster at Nimbus Cloud Platform.

The principal contribution of this work is a validated, open-source pipeline that translates raw meteorological inputs into actionable risk scores for wheat, barley, and canola cropping systems across the Murray–Darling Basin. The framework is designed to integrate directly with existing decision-support tools used by agronomists and policy planners, bridging the gap between computational climate science and on-farm practice.

Keywords: climate risk, machine learning, agricultural forecasting, ensemble methods, deep learning, Murray–Darling Basin

1

Introduction

1.1 Background and Motivation

Agricultural systems are inherently coupled to climate dynamics. In Australia, the intersection of a highly variable climate regime and an export-oriented agricultural sector creates a uniquely challenging risk environment. The Murray–Darling Basin, encompassing over one million square kilometres across four states, produces approximately one-third of Australia's food supply and supports 40% of the nation's agricultural output by value. Wheat, the dominant winter crop, generated export revenues exceeding $7.2 billion in the 2022–23 season, while barley and canola contributed a further $5.8 billion combined. These figures underscore the economic imperative of accurate seasonal climate forecasting for the Basin's producers, traders, and policymakers.

The 2017–2019 drought, described by the Bureau of Meteorology as the most severe on record for the northern Basin, reduced aggregate crop production by 53% and cost the national economy an estimated $6.3 billion in foregone agricultural output. Subsequent analysis by the Australian Bureau of Agricultural and Resource Economics and Sciences revealed that existing seasonal prediction models, which rely predominantly on coupled ocean–atmosphere general circulation models, failed to capture the rapid onset and spatial heterogeneity of the drought's impact. This failure highlighted a critical gap: the need for data-driven methods capable of learning complex, nonlinear relationships from high-dimensional climate observations.

1.2 Problem Statement

Despite significant advances in numerical weather prediction, operational seasonal forecasting systems continue to exhibit systematic biases in key agricultural regions. Three interrelated limitations motivate this research:

(1) **Scale mismatch.** General circulation models operate at spatial resolutions of 50–100 km, whereas agricultural decision-making occurs at the paddock scale (0.1–10 km). Downscaling introduces uncertainty that compounds across the forecast horizon.

(2) **Nonlinear dynamics.** The relationship between large-scale climate drivers — the El Niño–Southern Oscillation, the Indian Ocean Dipole, and the Southern Annular Mode — and local rainfall is mediated by nonlinear teleconnection patterns that linear statistical models cannot represent.

(3) **Data heterogeneity.** Relevant datasets span disparate formats, temporal resolutions, and spatial grids: satellite-derived vegetation indices, gridded reanalysis products, in-situ weather station records, and soil moisture profiles from wireless sensor networks deployed by projects such as the Nimbus Climate Platform.

1.3 Research Objectives

This thesis pursues four primary objectives:

RO1: Develop a unified data ingestion and preprocessing pipeline for heterogeneous climate and agricultural datasets, capable of handling missing values, spatial misalignment, and temporal irregularity.

RO2: Design and benchmark a suite of machine learning architectures — spanning tree-based ensembles, recurrent neural networks, and hybrid convolutional-recurrent models — for seasonal rainfall prediction at the catchment scale.

RO3: Construct a stacked ensemble model that optimally combines the strengths of component architectures, with rigorous evaluation against operational baselines using cross-validated skill scores.

RO4: Package the validated pipeline as an open-source software toolkit, accompanied by documentation and reproducible workflows, to facilitate adoption by research groups and government agencies.

1.4 Comparative Model Performance

Table 1.1 summarises the predictive performance of the models evaluated in this study, based on 5-fold cross-validation over the period 2000–2023. All metrics are reported on the hold-out test folds.

Table 1.1: Predictive performance of candidate models for seasonal rainfall forecasting (mm), aggregated across all catchments in the Murray–Darling Basin study area.

Model	RMSE (mm)	MAE (mm)	R^2	Training (s)
Persistence baseline	18.3	14.1	0.41	—
Random Forest	12.4	8.7	0.82	45.2
XGBoost	10.1	7.3	0.88	32.8
LSTM	9.6	6.9	0.89	187.5
CNN-LSTM	8.2	5.8	0.92	245.1
Ensemble (proposed)	**7.8**	**5.4**	**0.93**	298.7

The ensemble model combines XGBoost and CNN-LSTM predictions through a stacked meta-learner trained via 5-fold cross-validation on the training partition. The 37% reduction in RMSE relative to the persistence baseline is statistically significant at $p < 0.001$

under a two-sided Diebold–Mariano test with the Newey–West estimator of the long-run variance.

1.5 Contributions

The primary contributions of this thesis are:

C1: A rigorously curated and documented dataset linking meteorological observations from 112 Bureau of Meteorology stations with agricultural yield statistics across 47 catchments in the Murray–Darling Basin, spanning 44 years (1980–2023).

C2: The first systematic comparison of tree-based and deep-learning architectures for seasonal rainfall prediction in an Australian agricultural context, with complete reproducibility through open-source code and containerised execution environments.

C3: A novel stacked ensemble framework that outperforms all individual architectures and the operational ACCESS-S2 seasonal prediction system by a margin of 19–37% across standard skill metrics.

C4: An interactive decision-support dashboard, developed in collaboration with agronomists at Vertex Agricultural Consulting, that translates model outputs into crop-specific risk scores accessible to non-specialist users.

1.6 Thesis Organisation

The remainder of this thesis is structured as follows. Chapter 2 surveys the literature on statistical and machine-learning approaches to seasonal climate forecasting, with emphasis on applications in semi-arid agricultural regions. Chapter 3 describes the data sources, quality-control procedures, and feature engineering methodology. Chapter 4 presents the model architectures and training protocols in detail. Chapter 5 reports the experimental results, including sensitivity analyses and ablation studies. Chapter 6 discusses the implications of the findings for operational forecasting and agricultural policy. Chapter 7 concludes with a summary of contributions and directions for future work.