At larger scales and across crops, feature-importance and explainable-AI analyses consistently rank temperature, rainfall, and macronutrient levels among the most influential predictors of yield, revealing strong interactions between climate drivers and NPK supply (Meng et al., 2021; Mohan et al., 2025). In a maize yield prediction framework integrating fertilizer systems with multi-source data, random forest feature importance highlighted fertilizer variables, maximum temperature, and precipitation as key determinants, with different fertilizer systems shifting which factors were most limiting under given climatic conditions (Meng et al., 2021). Together, these results indicate that for eggplant, dominant yield-limiting factors are likely to be inadequate or poorly balanced N (and Ca), interacting with temperature and water availability, rather than single inputs considered in isolation.

7 Case Study on Regional Eggplant Yield Prediction

7.1 Overview of the selected experimental region

The experimental region represents a semi-arid to arid environment where eggplant production is constrained by high evaporative demand, limited and seasonally concentrated rainfall, and strong sensitivity of yield to microclimate modification. In cold and arid oasis conditions, such as the Hexi irrigation area of northwest China, annual precipitation is only about 183-285 mm, evaporation exceeds 1600 mm, and sunshine duration approaches 3000 h, creating a dry atmosphere where irrigation and fertilization strategies are critical to sustain eggplant productivity (Li et al., 2025). Comparable semi-arid vegetable regions, for example Carnarvon in Western Australia, face high temperatures and intense solar radiation during spring-summer, which damage crops and shorten the production season unless protective cultivation is adopted (Nguyen et al., 2022).

Within these environments, protected and controlled systems are increasingly used to create favorable microclimates for eggplant. Shade-net houses in Carnarvon, using moderate shade factors around 21%, altered light intensity and microclimatic conditions in ways that promoted taller, bushier plants and higher fruit yield compared with open-field cultivation (Figure 2) (Nguyen et al., 2022). Similarly, controlled and semicontrolled greenhouse systems in arid regions have shown that adjusting temperature, light, and nutrient sources (inorganic fertilizers, compost, and their mixtures) can strongly influence eggplant growth, yield, and water-use efficiency, providing locally specific data to calibrate yield models for such climates (Abbas et al., 2025).

7.2 Application of the prediction model to field data

Regional yield prediction relies on integrating field experiments that quantify responses to water and fertilizer regimes under real climate variability. In the Hexi oasis, split-plot experiments across two seasons with three irrigation levels (50%-60%, 60%-70%, 70%-80% field capacity) and three nitrogen rates (215, 270, 325 kg/ha) generated detailed yield, quality, and resource-use data, enabling identification of an optimal mild water deficit (60%-70% FC) with moderate nitrogen (270 kg/ha) under mulched drip irrigation (Li et al., 2025). These structured datasets, including soil properties and multi-year climate records, are well suited for training and validating regional prediction models that link fertilization and climate variables to eggplant yield.

Advanced modeling frameworks in other crops illustrate how multi-layered, multi-farm datasets can be used to forecast yield at field and regional scales. In Western Australia, yield monitor data for wheat, barley, and canola over three seasons were combined with weather and soil-related predictors to build random forest models at 100 m resolution, achieving concordance correlation coefficients of 0.89-0.92 and RMSE of 0.36-0.42 t/ha . Applying similar machine learning workflows to eggplant, using experimental and commercial field data from protected and open-field systems, allows spatially explicit yield forecasts that support regional fertilizer and irrigation decisions.

7.3 Implications for precision fertilization and farm management

Results from regional case studies highlight that optimal eggplant yield can be achieved with water- and nitrogen-saving strategies tailored to local climate, providing a basis for precision fertilization. In the cold, arid Hexi region, mild water deficit with moderate nitrogen significantly increased yield, fruit quality, and water- and nitrogen-use efficiency relative to unfertilized, fully irrigated controls, demonstrating that blanket high-input strategies are neither necessary nor efficient (Li et al., 2025). Parallel work in deficit drip irrigation on sandy clay loam soils showed that maintaining soil moisture at 75% field capacity maximized yield (≈39 t/ha) and irrigation water-use efficiency, with further increases in water supply reducing both yield and efficiency (Ouma et al., 2024).