management regimes are usually standardized within each trial; for example, integrated fertilization-tillage experiments in greenhouses applied defined NPK formulations and organic amendments across rotary tillage and plowing systems, and quantified biometric and yield traits under each treatment (Avasiloaiei et al., 2025). Together, such designs provide a template for specifying cultivars, structure, and management in temperature-yield modeling studies.

5.2 Temperature monitoring technologies and sensor deployment

Capturing temperature-yield relationships requires dense, reliable microclimate measurements rather than single-point records. Multi-year monitoring in commercial tomato greenhouses deployed multiple temperature-humidity sensors within the crop canopy, revealing spatial gradients up to about 3 °C in daily mean temperature and 0.6 kPa in vapour pressure deficit between locations, and linking these to local differences in stem and fruit growth (Šalagovič et al., 2024). Sensor networks are increasingly wireless: several studies describe custom wireless nodes or IoT platforms integrating temperature, relative humidity and sometimes CO_2 and light sensors, distributed at multiple horizontal positions and heights to resolve microclimate structure (Kolapkar and Sayyad, 2021). Such systems reduce cabling, facilitate relocation of nodes, and have been shown to detect microclimate layers between lower and upper canopy, as well as climate disturbances near walls or vents.

Recent work has combined distributed sensing with data fusion and model-based indices. In an Iranian commercial tomato greenhouse, a grid of 20 LoRaWAN wireless sensor nodes was installed on two horizontal planes at different heights, while an external weather station recorded outdoor conditions. Sensor calibration and validation were conducted offline in MATLAB/Simulink, and microclimate data were translated into an "optimality degree" index between 0 and 1 for temperature, RH and VPD, enabling direct assessment of how far local conditions deviated from crop comfort zones. Other wireless monitoring systems integrated fruit diameter sensors with 802.15.4-based temperature and radiation nodes and transmitted data via GPRS, achieving mean absolute temperature differences of only about 0.6 °C compared with wired systems, and data loss below 1%. Complementary approaches, such as compliant "plant wearables" measuring temperature and humidity directly on leaf surfaces, illustrate emerging options for ultra-localized microclimate characterization within greenhouse crops (Nassar et al., 2018).

5.3 Yield data collection and statistical preprocessing methods

Tomato yield data in greenhouse experiments are generally collected at plant or area level using standardized protocols, then subjected to statistical analysis and, in modeling studies, further preprocessing. Many agronomic trials quantify number of fruits per plant, individual fruit weight and total yield (e.g., $t \cdot ha^{-1}$ or $g \cdot plant^{-1}$) at one or more harvests, often alongside traits such as fruit size, firmness, soluble solids, and dry matter (Avasiloaiei et al., 2025; Ugbe et al., 2025). In cultivar or spacing-topping trials under greenhouse conditions, randomized or randomized complete block designs with three or more replications are analyzed using analysis of variance, with significance judged at $p < 0.05$ and treatment means separated by least significant difference or similar procedures. Microclimate-growth studies add growth rates of stems and fruits or truss mass at harvest, relating these to local temperature and VPD conditions over defined periods.

For data-driven modeling of temperature-yield relationships, more elaborate preprocessing is required. Yield-prediction studies using artificial neural networks and other machine-learning methods typically assemble datasets that combine environmental descriptors, management variables, and yield as inputs and outputs, then partition data into training and validation sets (Peng et al., 2023). A recent solar-greenhouse study compiled 390 datasets across multiple regions, each including planting density, organic and inorganic N, P, K rates, and effective accumulated temperature, with greenhouse tomato yield as the response; these variables were scaled and classified into different soil fertility levels before being used in neural-network models. In UAV-based yield prediction, ultra-high-resolution imagery was processed into hundreds of plant-level variables (e.g., means and higher-order statistics of vegetation indices), then reduced using feature-selection algorithms before model fitting (Tatsumi et al., 2021). Across these approaches, standard error metrics such as mean squared error, mean absolute error and coefficient of determination are calculated to evaluate predictive performance and to support sensitivity analysis