

greenhouse disease management (Lee and Yun, 2023). A central variable in these models is leaf wetness duration, because humidity affects disease not only through bulk air moisture but through condensation and persistence of wet surfaces. Early warning work for cucumber downy mildew treated leaf wetness duration as a key variable and combined leaf-wetness sensing with a relative-humidity threshold approach to create a practical estimation method when direct monitoring was difficult. Later greenhouse studies confirmed that leaf wetness duration is an important disease-model input related to infection and pathogen development, and showed that a back-propagation neural network estimated it more accurately than a simple relative-humidity model, reaching accuracies of 0.90 and 0.92 in two greenhouses (Liu et al., 2020).

Recent modeling has also moved from point estimates toward spatially explicit representations of humidity-related infection conditions inside greenhouses. A CFD transient model showed that canopy condensation caused by high humidity is a major cause of leaf wetness duration formation and could estimate the temporal and spatial distribution of wetness across cucumber canopies with good agreement to observations. Spatial heterogeneity studies likewise found that leaf wetness duration varied systematically within solar greenhouses, with longer wetness in the south and east zones and on rainy days, patterns that are closely related to the occurrence of high-humidity cucumber diseases (Liu et al., 2020). Another advance is the coupling of climate prediction with disease-process models so that warning systems can anticipate risk before unfavorable humidity conditions fully develop. A model-based methodology combined a mechanistic greenhouse climate model with a disease model, first predicting indoor climate 72 hours ahead and then using that forecast to detect downy mildew occurrence in advance. This integration is strengthened by machine-learning humidity forecasting, where a stacking ensemble for greenhouse indoor humidity achieved an R^2 of 0.96515 and high predictive precision, showing that accurate humidity prediction itself is now feasible enough to support downstream disease-risk models (Melal et al., 2024).

5.2 Machine learning prediction of disease risk

Machine learning has improved cucumber disease-risk prediction by exploiting the sequential structure of greenhouse environmental data rather than relying only on static thresholds. In cucumber downy mildew, an LSTM neural network was built from IoT-acquired time-series data on temperature, relative humidity, soil temperature, and solar radiation, and the resulting disease prediction model achieved 90% accuracy, 94% precision, 89% recall, and an AUC of 90.15%. The advantage of this approach is that classic machine-learning methods often handle long-term sequence dependence poorly, whereas LSTM better captures previous environmental feature information that is closely tied to disease onset (Liu et al., 2022). Prediction performance has improved further when environmental data are fused with direct disease indicators. A CNN-LSTM model for cucumber downy mildew integrated greenhouse and outdoor environmental records with airborne spore counts and observed diseased leaf area, reaching an R^2 of 0.9127 with low mean absolute and root mean square errors. The same study argued that future real-time prediction should integrate online spore detection with environmental big data, indicating that disease forecasting is moving toward multimodal monitoring rather than climate-only inference (Wang et al., 2024).

Broader deep-learning evidence suggests that humidity-based risk prediction can generalize beyond a single crop or pathogen when sufficient environmental data are available. A recent sequential deep-learning framework used previous growth-environment information, including air temperature, relative humidity, dew point, and CO_2 concentration, to predict crop disease risk and achieved an average AUROC of 0.917 across multiple crops and facilities. This wider applicability is plausible because environmental data are comparatively easy to collect in facility farms, and the authors argue that such data-driven learning frameworks could be used broadly for disease and pest prevention in controlled agriculture (Lee and Yun, 2023). At the same time, current machine-learning prediction still depends on stable management conditions and high-quality sensor streams. In the cucumber LSTM study, abrupt human-driven changes in greenhouse climate could disrupt data acquisition and degrade prediction accuracy, so venting schedules and field operations had to remain orderly for reliable modeling (Liu et al., 2022). Related greenhouse IoT reviews similarly note that high-detail monitoring in space and time is what enables more