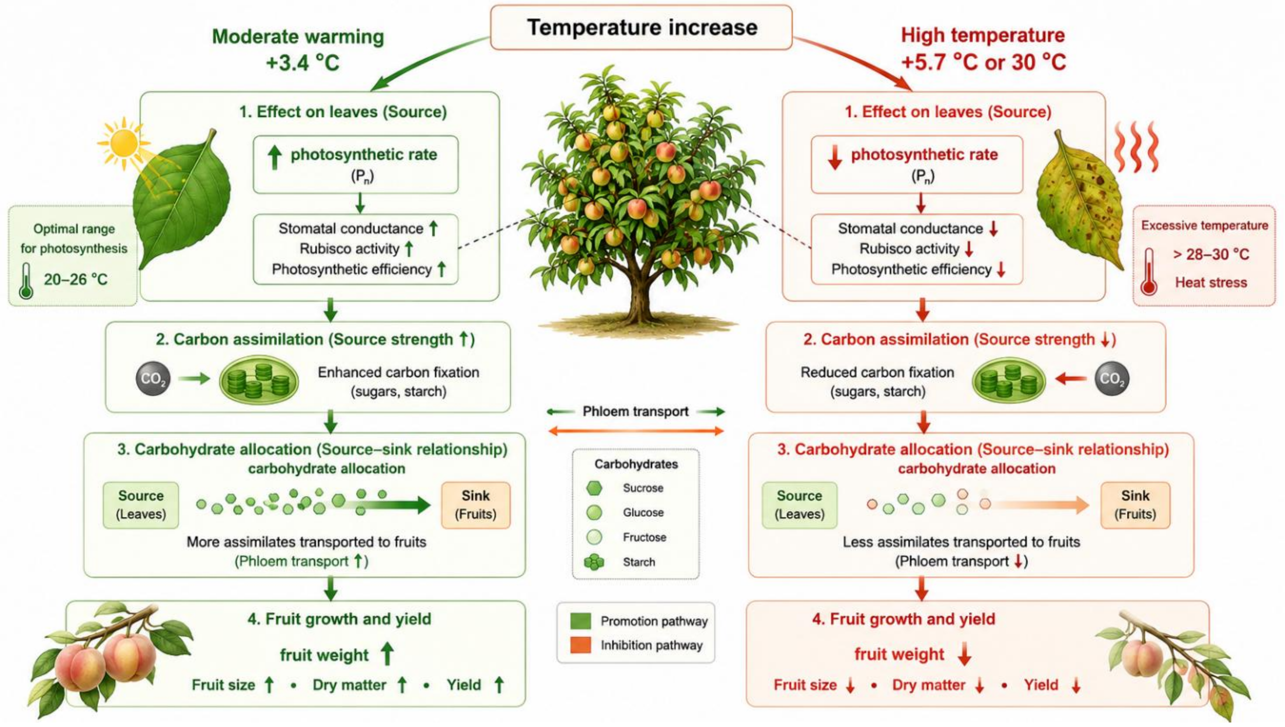

Figure 2 Effects of temperature variation on photosynthesis and source-sink dynamics in peach trees under controlled environments

5.3 Development of statistical and machine learning models

Once temperature indicators and physiological covariates are defined, they can be embedded in statistical and machine learning frameworks. For bud and bloom phenology, models combining chill accumulation (Dynamic model) with growing degree sums have predicted bud development stages across major Spanish peach regions with errors of about four days, outperforming forcing-only models and providing physiologically consistent thresholds for chill and heat. At finer scales, chill-heat models for individual cultivars have been fitted using sequential chill and GDH accumulation to estimate budbreak timing, offering simple yet robust tools for linking winter-spring temperatures to the onset of reproductive development (Cifuentes-Carvajal et al., 2023).

For yield-focused modeling, multivariate machine learning approaches appear especially promising. In ‘Esmeralda’ peach, k-nearest neighbors and stochastic gradient descent models trained on meteorological indices, foliar nutrients, and prior yield achieved accuracies up to 1.00 for several yield and quality indices, with chilling hours and degree-days emerging as top-ranked features (Nava et al., 2022). A broader yield prediction study comparing Random Forest, multiple linear regression, and support vector machines found that Random Forest provided the best performance, and identified hours of chilling, leaf K and N, and mean temperature as the most influential variables, confirming that nonlinear ML models can effectively learn complex temperature-nutrition-yield relationships when supported by well-engineered temperature indicators (Moura-Bueno et al., 2026).

6 Evaluation and Optimization of Temperature-Driven Prediction Models

6.1 Model performance comparison and selection

Evaluation of temperature-driven peach yield and quality models requires systematic comparison of alternative model structures and learning algorithms. For peach yield, a study comparing Random Forest, Multiple Linear Regression, and Support Vector Machine found that Random Forest achieved the highest predictive accuracy when using climatic, soil, and leaf nutrient data, with chilling hours and mean temperature among the most influential predictors (Moura-Bueno et al., 2026). Similarly, temperature-based phenology models for peach bloom (developmental rate, chill day, and new chill day models) were assessed with MAPE, R^2, and RMSE, and the new chill day model provided the best compromise between bias and precision across cultivars and sites, illustrating the value of multi-metric model comparison for temperature-driven processes.