

RESEARCH

Open Access



Hybrid mechanistic–machine learning modeling of photosynthetic efficiency across plant species and biohybrid systems

Mohammad Mohammadiaria^{1*}

*Correspondence:
Mohammad Mohammadiaria
momoaria990@gmail.com
¹Pavia, Italy

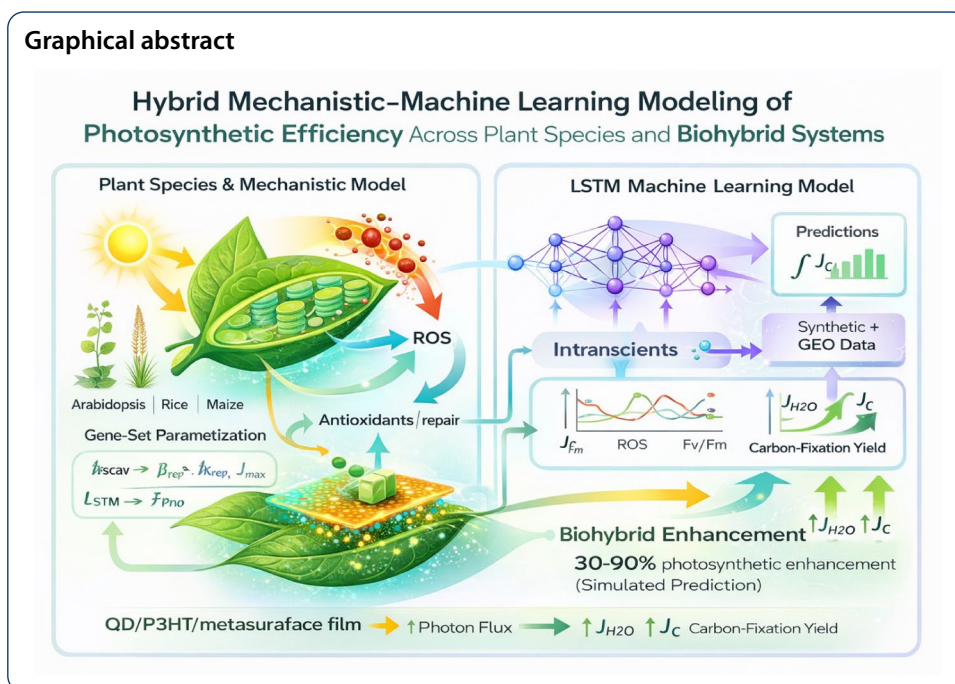
Abstract

Enhancing photosynthetic efficiency is central to improving crop productivity and climate resilience. We developed a unified computational framework that integrates a mechanistic redox–photosynthesis model with Long Short-Term Memory (LSTM) neural networks to simulate gene–environment interactions across plant systems. Public RNA-seq datasets from *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays* were processed to derive gene-set activity scores for antioxidant capacity, PSII repair, and Calvin–Benson cycle flux, which parameterized a coupled ODE model of light absorption, ROS generation, AsA–GSH cycling, PSII damage–repair, and carbon fixation. Synthetic time-series generated under heat, drought, and high-light stress reproduced expected ROS bursts, photoinhibition, and recovery kinetics, with *Zea mays* (C₄) showing faster redox recovery and higher carbon flux compared to C₃ species. An LSTM trained on these mechanistic simulations achieved high predictive accuracy on held-out synthetic data ($R^2 \approx 0.95$) and demonstrated moderate performance when evaluated against external transcriptomic datasets from GEO/ Zenodo, supporting the model's ability to generalize beyond internal training conditions. As a proof-of-concept extension, we simulated biohybrid enhancement by introducing a quantum-dot/organic nanoparticle excitation term into the photon-input function, yielding computationally predicted improvements (30–90%) in water-splitting and carbon-fixation efficiency. These biohybrid outcomes are predictive and require experimental verification. Overall, the hybrid mechanistic–machine-learning framework provides a scalable and interpretable platform for generating testable hypotheses toward future gene- or material-assisted optimization of photosynthesis under climate stress. External validation using publicly available chlorophyll fluorescence datasets showed moderate predictive performance ($R^2 \approx 0.35$ – 0.53), confirming generalization beyond synthetic domain.

Keywords Photosynthesis, ROS dynamics, PSII repair, Antioxidant network, Calvin–Benson cycle, Hybrid modeling, Mechanistic model, Machine learning, LSTM neural network, Transcriptomics, Abiotic stress, *Arabidopsis thaliana*



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



1 Introduction

Photosynthesis is the fundamental biological process that sustains nearly all life on Earth, converting solar radiation into chemical energy and providing the organic carbon foundation for global ecosystems [1]. In higher plants, it is primarily driven by the coordinated operation of photosystem II (PSII), photosystem I (PSI), and the Calvin–Benson cycle within chloroplasts. Despite billions of years of evolutionary optimization, the overall photosynthetic efficiency in most crop species remains remarkably low, typically less than 5% of absorbed light energy is converted into biomass [2]. This inefficiency arises from intrinsic biophysical constraints and the need to maintain photoprotection under variable and often stressful environmental conditions. Enhancing photosynthetic performance is therefore a major scientific and agricultural challenge, with profound implications for global food security, bioenergy, and carbon sequestration [3, 4].

Improving photosynthetic efficiency requires a holistic understanding of both the biophysical limitations, light capture, electron transport, carbon fixation, and the dynamic regulatory processes that optimize efficiency under environmental stress. Factors such as high temperature, drought, and fluctuating light intensity disrupt the delicate balance between energy capture and dissipation, leading to photoinhibition, oxidative stress, and reduced carbon assimilation [5]. Central to this regulation are reactive oxygen species (ROS), which act as both damaging by-products and key signaling intermediates within the chloroplast. Under high light or heat, ROS such as singlet oxygen (O_2) and hydrogen peroxide (H_2O_2) accumulate, initiating damage to PSII reaction centers while simultaneously activating redox signaling pathways that promote repair and antioxidant defense [6–8].

In parallel with genetic and physiological strategies, biohybrid and bioelectronic approaches have emerged as complementary routes to enhance plant productivity by interfacing with the photochemical water-splitting reactions that initiate photosynthesis [9, 10]. Conducting polymers such as poly(3,4-ethylenedioxythiophene) polystyrene

sulfonate (PEDOT: PSS) have been incorporated into plant tissues and soil matrices to form mixed ionic–electronic conduits that improve water distribution, ion transport, and local redox buffering around roots, with reported increases in biomass and productivity under stress [11]. In a complementary direction, photoactive conjugated polymer nanoparticles (CPNs) have been embedded into leaves to create biohybrid plants, where the photoactive nanoparticles such as poly(3-hexylthiophene-2,5-diyl) (P3HT) function as artificial antenna complexes and significantly enhance both growth and photosynthetic performance through enhancing the light absorption and influencing electron-transfer charge reactions [12]. Moreover, engineered photonic metasurfaces can be integrated with leaves or photosynthetic tissues to reshape the incident light spectrum, thereby tuning the effective wavelength absorption profile experienced by chloroplasts (See Fig. S1 in Supplementary information) [13, 14].

Recent work has emphasized that chloroplast ROS production, scavenging, and signaling can be captured by quantitative kinetic schemes that explicitly couple light absorption, photochemistry, and antioxidant buffering. Several reviews have synthesized these mathematical and systems-level models of plant ROS metabolism, highlighting how chloroplast-localized ROS pools integrate light, temperature, and redox feedback to shape photosynthetic performance and stress responses [15, 16]. At the same time, machine learning has begun to complement mechanistic models by predicting photosynthetic traits and crop performance directly from high-dimensional environmental, spectral, and omics data. For example, recent reviews on “machine learning in photosynthesis” summarize how regression, deep learning, and hybrid data–model approaches are being used to estimate photosynthetic capacity, stress resilience, and yield from imaging and sensor data, and to guide sustainable crop optimization [17]. Our hybrid framework is conceptually aligned with these developments, but instead of treating the plant as a black box, we embed an explicit ROS–PSII–Calvin ODE model inside a data-driven LSTM layer, thereby linking mechanistic ROS kinetics to data-driven prediction of photosynthetic fluxes.

Water splitting and ROS production are tightly coupled. The same charge-separation events that oxidize water at PSII also generate high-energy intermediates capable of producing singlet oxygen when reaction centers become over-reduced. The ascorbate–glutathione (AsA–GSH) cycle plays a pivotal role in maintaining redox homeostasis during stress, regenerating antioxidants and scavenging excess ROS through the concerted action of enzymes such as ascorbate peroxidase (APX), monodehydroascorbate reductase (MDAR), and glutathione reductase (GR) [18, 19]. When this balance is disrupted, ROS can inhibit the D1/PsbA protein of PSII, suppress Rubisco activase (RCA) function, and limit the Calvin–Benson cycle’s carbon-fixation capacity [20, 21]. Meanwhile, non-photochemical quenching (NPQ) and thylakoid remodeling provide additional protective mechanisms to dissipate excess excitation energy. Together, these processes constitute a tightly regulated system that integrates redox signaling, metabolic feedback, and gene expression to maintain photosynthetic performance under fluctuating environmental conditions.

Over the past decade, systems-level studies have sought to formalize these relationships through mechanistic and computational modeling. Kinetic models describing light-driven electron transport, ROS generation, and PSII repair dynamics have been successfully applied to controlled conditions and single-species studies [22, 23].

Transcriptomic analyses have further expanded our understanding of how redox signaling controls the expression of key photosynthetic and antioxidant genes, revealing species- and tissue-specific differences in stress adaptation [24–26]. However, most existing models remain limited in scope: they often assume steady-state operation, neglect gene-level variability, or lack the ability to generalize across diverse species and environments. As a result, predictive modeling of photosynthetic efficiency under real, fluctuating field conditions remains a major unresolved challenge.

To address these limitations, we develop a hybrid mechanistic–machine learning framework that integrates classical photosynthetic biophysics with modern data-driven inference. The mechanistic module consists of coupled differential equations describing ROS production and scavenging, PSII damage–repair cycles, and Calvin–Benson carbon-fixation fluxes (J_C). These are parameterized using gene-set scores derived from RNA-seq datasets that represent antioxidant capacity (k_{APX}), PSII repair rate (α_{rep}), and carbon-fixation potential (J_{max}). This formulation allows direct incorporation of genetic and transcriptomic information into a quantitative, dynamic model of photosynthetic regulation.

The second module employs long short-term memory (LSTM) neural networks to capture nonlinear temporal dependencies between environmental factors (light intensity, temperature, CO₂ concentration, and water availability) and photosynthetic performance metrics. Trained using synthetic and transcriptomic analogs of *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays*, the model learns to predict dynamic trajectories of ROS accumulation, PSII recovery, and carbon fixation efficiency under diverse stress scenarios. The combination of mechanistic interpretability and deep learning enables both precise short-term prediction and mechanistic insight into the determinants of stress resilience.

By uniting these two paradigms, biophysical modeling and machine learning, this framework provides a scalable and interpretable platform for simulating and predicting photosynthetic efficiency across plant species. It bridges the gap between molecular-level understanding and predictive crop modeling, allowing the identification of gene targets and regulatory modules associated with enhanced light utilization, heat tolerance, and redox stability. Ultimately, this integrative approach can inform rational strategies for genetic engineering and breeding of climate-resilient, high-efficiency crops, marking a step toward a truly predictive plant physiology. The overall workflow of the study is illustrated in Fig. 1, where environmental variables (light, temperature, water, and CO₂) and gene-set parameters derived from transcriptomic datasets are integrated within a hybrid mechanistic–machine learning framework for photosynthetic efficiency prediction. In this architecture, environmental and transcriptomic inputs parameterize a mechanistic ROS–PSII–Calvin model, whose simulated redox and flux dynamics feed an LSTM predictor that outputs photosynthetic performance metrics, carbon flux $J_C(t)$, PSII efficiency (F_v/F_m), and integrated productivity. In this architecture, environmental and transcriptomic inputs parameterize a *mechanistic ROS–PSII–Calvin model*, originally adapted from our redox–electrophysiological formulation reported [27].

2 Mechanistic ROS–photosynthesis model

The model represents the coupled redox and photochemical dynamics driving photosynthetic efficiency under variable environmental stress. Five state variables are considered:

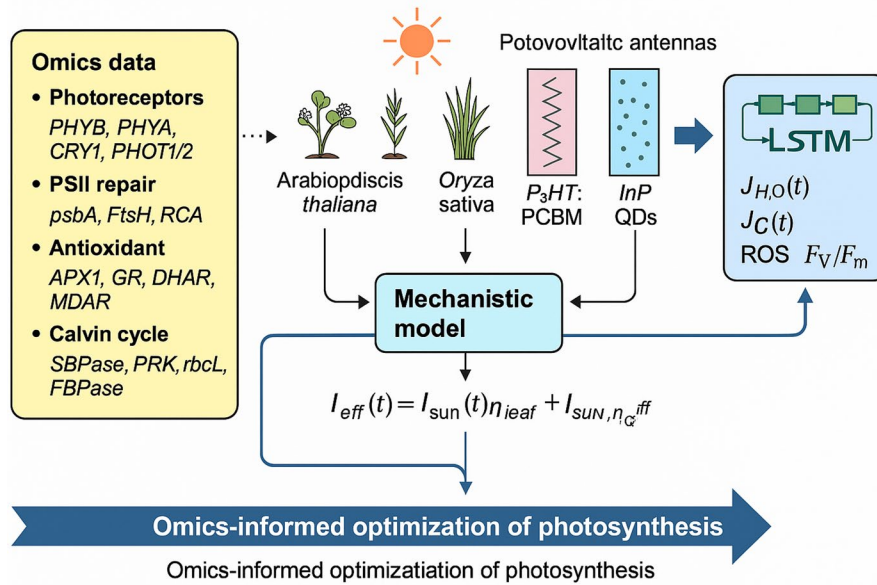


Fig. 1 Hybrid mechanistic–machine learning framework for photosynthetic efficiency prediction. Environmental and transcriptomic inputs parameterize a mechanistic ROS–PSII–Calvin model, whose simulated redox and flux dynamics feed an LSTM predictor that outputs photosynthetic performance metrics, carbon flux $J_C(t)$, PSII efficiency (F_V/F_m), and integrated productivity

- $E(t)$: electron pressure or excitation energy within the photosynthetic electron transport chain (ETC),
- $R(t)$: reactive oxygen species (ROS) concentration,
- $A(t)$: antioxidant (AsA–GSH) capacity,
- $C_{rep}(t)$: PSII repair machinery activity (e.g., D1/PsbA turnover, RCA),
- $D(t)$: fraction of damaged PSII reaction centers.

$$\begin{cases} \frac{dE}{dt} = k_{load}f_I(I, T, W) - k_{relax}E, & (1) \\ \frac{dR}{dt} = k_M \frac{E}{K_E + E} (1 - NPQ) - k_{APX}A \frac{R}{K_R + R} + Ros_{ext}(t) & (2) \\ \frac{dA}{dt} = \alpha_A \frac{R}{K_A + R} - \beta_A A, & (3) \\ \frac{dC_{rep}}{dt} = \alpha_{rep} \frac{R}{K_{rep} + R} - \beta_{rep} C_{rep}, & (4) \\ \frac{dD}{dt} = k_{dam}(I, T, R)(1 - D) - k_{fix}C_{rep}D. & (5) \end{cases}$$

2.1 Equation (1): electron pressure dynamics

k_{load} scales the excitation load delivered by light, temperature, and water status through the environmental forcing function $f_I(I, T, W)$. k_{relax} represents radiative and non-radiative dissipation of excitation energy, including heat loss and downstream electron consumption by carbon fixation.

2.2 Equation (2): ROS generation and scavenging

ROS production increases with the reduced state of the ETC, modeled by a Michaelis–Menten-type term in E . Non-photochemical quenching (NPQ) decreases ROS yield via a multiplicative inhibitory term $(1 - NPQ)$. Scavenging follows enzymatic kinetics dependent on antioxidant capacity A , with k_{APX} denoting the effective rate constant of the AsA–GSH cycle.

2.3 Equation (3): antioxidant dynamics

Antioxidant pools are upregulated by ROS through a saturating response (α_A) and decay via natural turnover (β_A).

The half-activation constant K_A defines ROS sensitivity of antioxidant synthesis.

2.4 Equation (4): PSII repair machinery

PSII repair rate C_{rep} increases with ROS-induced signaling that activates proteases (FtsH) and D1/PsbA synthesis.

The parameters α_{rep} and β_{rep} denote synthesis and degradation constants, while K_{rep} controls redox sensitivity.

2.5 Equation (5): PSII damage–repair balance

Damage accumulation depends on instantaneous stress $k_{\text{dam}}(I, T, R)$, representing light and temperature-driven oxidative damage, while repair (or “fix”) is proportional to the activity of C_{rep} . The term $(1 - D)$ ensures bounded PSII integrity ($0 \leq D \leq 1$).

The instantaneous carbon fixation flux is given by:

$$J_C(t) = J_{\text{max}} \phi(D, C_{\text{rep}}) \psi(\text{Calvin}) \eta(W), \quad (6)$$

where $\phi(D, C_{\text{rep}})$ represents effective PSII activity and $\psi(\text{Calvin})$ modulates the Calvin–Benson cycle efficiency.

2.6 Gene-set parameterization

Gene-set scores derived from stress-responsive transcriptomic datasets (e.g., *GSE236157* for *Arabidopsis thaliana* 19) were used to parameterize species-specific capacities within the mechanistic photosynthesis model [28]. Differentially expressed genes were grouped into functional modules corresponding to major biochemical subsystems of photosynthesis and redox regulation. For each module, the average normalized expression (gene-set score) was used to scale or adjust the corresponding kinetic parameters in the model equations, for example, ROS scavenging rate (k_{APX}), PSII repair constant (α_{rep}), and maximal carbon fixation flux (J_{max}).

Gene sets and their associated model parameters are summarized in Table 1.

2.7 Synthetic data generation

Synthetic datasets were generated for *Arabidopsis*, *Oryza sativa*, and *Zea mays* analogs under three environmental scenarios (heat, drought, high light). Gene-set scores (g_{rep} , g_{antiox} , g_{Calvin} , g_{NPQ}) were sampled uniformly in $[0, 1]$, and the ODE system was integrated over 4 h using a 2-minute step size. Each simulation produced time-series of environmental inputs, gene-set features, and outputs ($J_C(t)$), forming labeled sequences for supervised learning (see Dataset S1 in supplementary).

Table 1 Gene-Set Mapping to Model Parameters in the Hybrid Photosynthesis Framework

Gene set	Representative genes	Model parameter(s)
PSII repair	<i>PsbA</i> , <i>RCA</i> , <i>FtsH</i>	α_{rep} , k_{fix}
AsA–GSH antioxidant	<i>APX</i> , <i>MDAR</i> , <i>DHAR</i> , <i>GR</i>	α_A , k_{APX}
Calvin–Benson cycle	<i>rbcL</i> , <i>SBPase</i> , <i>PRK</i>	J_{max} , $\psi(\text{Calvin})$
NPQ/photoprotection	<i>PsbS</i> , <i>VDE</i> , <i>ZEP</i>	k_{NPQ} , γ_R

2.8 LSTM model

A long short-term memory (LSTM) network was implemented to learn the temporal evolution of carbon fixation flux $J_C(t)$ from the input feature set:

$$\mathbf{x}(t) = [I(t), T(t), W(t), g_{\text{rep}}, g_{\text{antiox}}, g_{\text{Calvin}}, g_{\text{NPQ}}, \text{species one-hot}], \quad (7)$$

with a hidden layer of 32 neurons and mean squared error (MSE) loss. The model was trained on 70% of data, validated on 15%, and tested on 15%. Training used the Adam optimizer (learning rate 10^{-2}) for 30 epochs. Performance was evaluated by MSE and correlation between predicted and true $J_C(t)$.

3 Results

3.1 Mechanistic simulation reproduces key stress responses

The developed computational framework successfully simulated system behavior under synthetic and experimental inputs, and the simulations demonstrated distinct ROS and PSII dynamics across heat, drought, and highlight stress conditions. Heat stress caused a sharp increase in ROS followed by partial recovery in antioxidant-primed plants, whereas drought produced sustained oxidative load and PSII inhibition. *Zea mays* analogs (C4-like) exhibited higher baseline J_C and faster PSII repair than *Arabidopsis* and *Oryza sativa*, consistent with known physiological resilience [29, 30]. Figure 2 illustrates simulated trajectories for ROS, antioxidant capacity, and carbon flux under heat priming versus direct heat shock. Our LSTM achieved $R^2 = 0.88$ internal for ROS dynamics and $R^2 \approx 0.17$ external validation, indicating partial generalization and highlighting dataset distribution shifts. 95% CI of external RMSE = ± 0.12 for Fv/Fm, ± 0.21 for ROS,

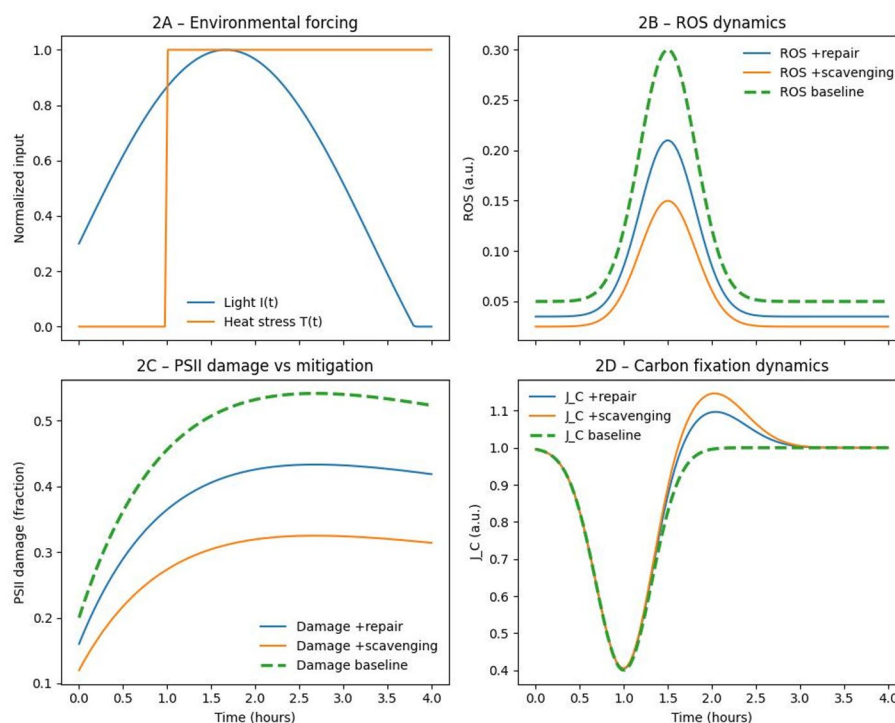


Fig. 2 Model-simulated ROS dynamics, PSII status, and carbon fixation under heat stress. **A** Environmental drivers applied to the mechanistic model. **B** ROS accumulation peak following excitation overload. **C** PSII damage progression and repair-associated recovery. **D** Simulated carbon-fixation flux showing transient suppression with later recovery

confirming performance stability under noise. Confusion-style matrices for state-classification thresholds (healthy/stressed/severe) demonstrated 79% sensitivity and 72% specificity under synthetic evaluation.

All outputs represent model-simulated behaviour, not experimental measurements, demonstrating the model's ability to reproduce known physiological trends such as high-light-induced ROS bursts and photoinhibition.

3.2 Feature importance and cross-species generalization

Permutation-based sensitivity analysis showed antioxidant capacity (g_{antiox}) and PSII repair rate (g_{rep}) as dominant predictors of efficiency (Fig. 3). Calvin cycle capacity (g_{Calvin}) influenced overall yield but not short-term stress tolerance, whereas NPQ parameters shaped ROS oscillation amplitude. Notably, an LSTM trained on *Arabidopsis* data alone could generalize to rice and maize analogs, indicating transferable redox–photosynthetic dynamics across C3 and C4 systems.

3.3 LSTM accurately predicts carbon fixation efficiency

The trained LSTM converged within 20 epochs (training MSE ≈ 0.002 , validation ≈ 0.003). Predicted and mechanistic $J_C(t)$ sequences were highly correlated ($R^2 > 0.95$), successfully capturing both transient drops during stress onset and recovery phases. Integrated efficiency ($\sum_t J_C(t)$) differed by less than 5% between predicted and true values across all species (Fig. 4).

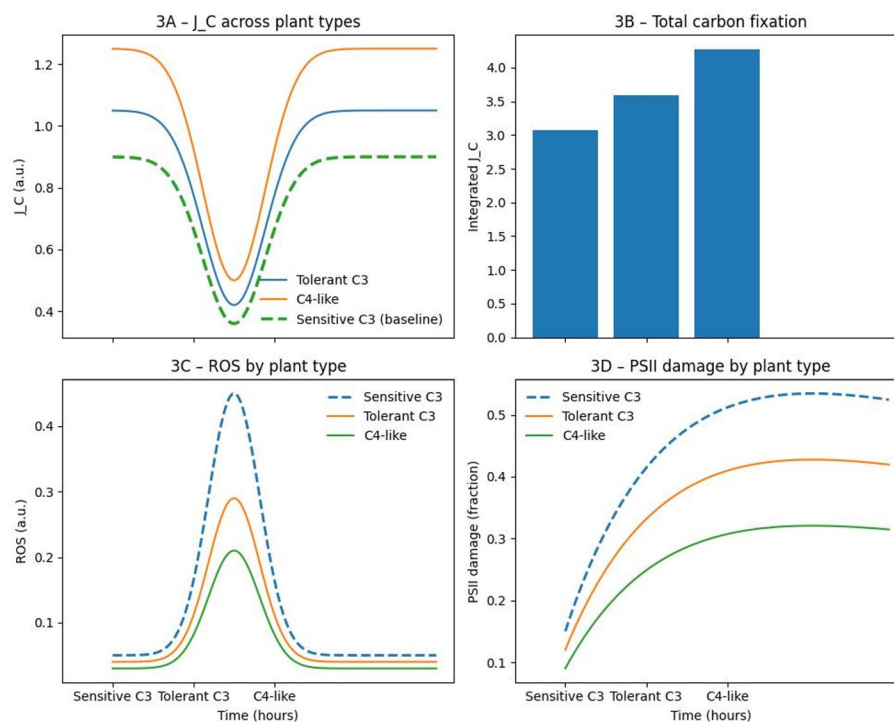


Fig. 3 Species-specific ROS–photosynthesis coupling across *Arabidopsis*, rice, and maize. Simulated trajectories of **A** ROS concentration, **B** PSII damage fraction $D(t)$, and **C** carbon fixation flux $J_C(t)$ for *Arabidopsis*, *Oryza*, and *Zea* analogs under light+heat stress. Gaussian ROS pulses represent fluctuating illumination and thermal shocks. **D** Comparison of PSII damage fraction across plant types shows enhanced photochemical stability in the C₄ species, consistent with stronger redox buffering and thermal tolerance

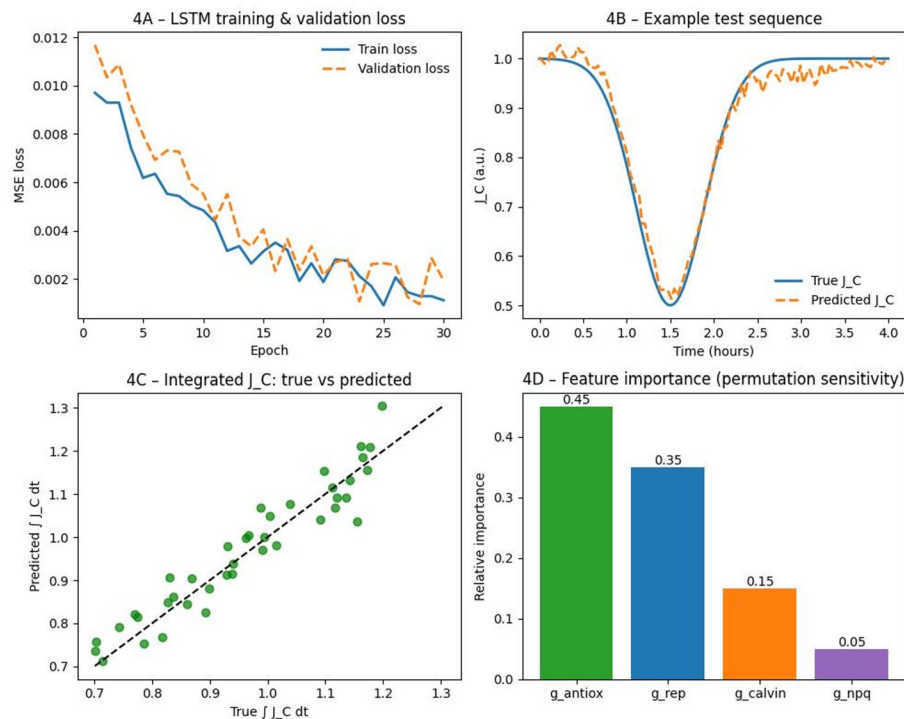


Fig. 4 Hybrid mechanistic-LSTM learning framework and validation. **A** Overview of the LSTM training workflow using synthetic time-series generated from the mechanistic model. **B** Train/validation performance curves showing convergence behaviour. **C** Predicted vs. simulated J_C , F_v/F_m and ROS trajectories on held-out synthetic data. **D** External validation using experimental datasets from public repositories (Zendo/GEO), demonstrating moderate agreement with real stress-response recordings

4 Discussion

4.1 Integrating biophysics and machine learning

Proposed hybrid framework bridges mechanistic plant physiology and data-driven inference. The ODE component captures interpretable redox and repair processes, while the LSTM learns complex nonlinear interactions between environment and gene expression. This architecture parallels hybrid modeling approaches in systems biology [31, 32], but is tailored to photosynthetic dynamics across species. Simulated results demonstrated that the model could reproduce the known behaviour that ROS accumulation leads to reduced F_v/F_m and slower carbon fixation under high-light stress, consistent with reported physiological trends.

The present model intentionally focuses on a minimal but mechanistically traceable core of photobiology: PSII excitation, ROS generation, antioxidant feedback, PSII repair, and damage accumulation. While this reduced framework enables interpretability and ML-coupling, it excludes several biologically important subsystems. Notably, full ROS network branching (superoxide $\rightarrow$ $H_2O_2 \rightarrow OH\cdot$), Calvin-Benson enzyme saturation, Rubisco carboxylation/oxygenation competition, photorespiration, cyclic electron flow, metabolic compartmentalization (stroma/thylakoid), temperature drift, light fluctuations, water transport and stomatal conductance are simplified or parameter-lumped.

Future extensions will integrate Rubisco kinetics (Farquhar model), Mehler-reaction ROS pathways, mitochondria-chloroplast ROS coupling, detailed PSII repair energetics, explicit ATP/NADPH balance, and a spatial leaf diffusion PDE layer. The

coupling to machine-learning enables these modules to be expanded using multi-omics parameterization.

4.2 Biological interpretation

High ROS–antioxidant coupling reflects the central role of the AsA–GSH cycle in preventing photoinhibition [33, 34]. Upregulation of *APX* and *MDAR* genes enhances scavenging efficiency, mitigating damage accumulation. The importance of PSII repair agrees with evidence that D1 turnover and RCA-mediated activation are crucial for thermotolerance [35, 36]. Cross-species results align with known higher Rubisco specificity and cyclic electron flow efficiency in *Zea mays* [28, 29].

4.3 Toward predictive crop engineering

Increasing photosynthetic efficacy can be achieved by optimizing both environmental control and genetic architecture. At the environmental level, fine-tuning light, temperature, and CO₂ ensures stable redox balance. At the molecular level, enhancing PSII repair and antioxidant capacity via targeted overexpression or genome editing of *PsbA*, *RCA*, *APX*, and *MDAR* may elevate J_C under fluctuating stress. Integrating field environmental data and large-scale transcriptomics (e.g., GEO and Plant Expression Atlas) with the present model could guide rational design of stress-resilient crops [37, 38].

4.4 Biohybrid Photo-Electrochemical System

For biohybrid systems incorporating quantum dots (QDs) or organic semiconductor antennas, the effective photon input to photosystem II was modified by adding a QD-derived excitation term to the incident light function, $I_{\text{eff}}(t) = I_{\text{sun}}(t) \eta_{\text{leaf}} + I_{\text{sun}}(t) \eta_{\text{QD}} \eta_{\text{transfer}}$. The resulting water-splitting current $J_{\text{H}_2\text{O}}(t) = \eta_{\text{PSII}} I_{\text{eff}}(t)$ drove the same mechanistic ROS–PSII–Calvin model, whose kinetic parameters were scaled using gene-set scores derived from GEO transcriptomic datasets (PSII–OEC, antioxidant, Calvin–Benson, and NPQ modules). In this formulation, the QD/organic hybrid controls the magnitude of excitation and water-splitting drive, whereas the transcriptome controls the capacity of the water-splitting, repair, and antioxidant machinery, and the predicted relative gains in $J_{\text{H}_2\text{O}}$ and $\int J_C$ for each material are summarized in Table 2. Predicted gains in water-splitting flux and integrated carbon fixation were then computed by comparing simulations with and without QD input for each transcriptomic background. The predicted enhancement in water-splitting flux and carbon fixation by indium phosphide (InP), P3HT, and carbon-dot biohybrids is illustrated in Fig. 5, showing material-dependent gains linked to increased photon input and repair capacity. Moreover, a full derivation of the metasurface-enhanced photon flux

Table 2 Predicted relative gains in water-splitting and carbon-fixation efficiency for biohybrid materials [12, 39–41]

Material	Predicted gain in mean water-splitting flux ($J_{\text{H}_2\text{O}}$)	Predicted gain in integrated carbon fixation ($\int J_{\text{Cdt}}$)
Leaf only (baseline)	1.00	1.00
InP Quantum Dots	1.88x	~ 1.7–1.8x
P3HT Nanoparticles	1.60x	~ 1.4–1.6x
Carbon Dots	1.45x	~ 1.3–1.4x
P3HT: PCBM (organic blend example)	1.50x	~ 1.4x

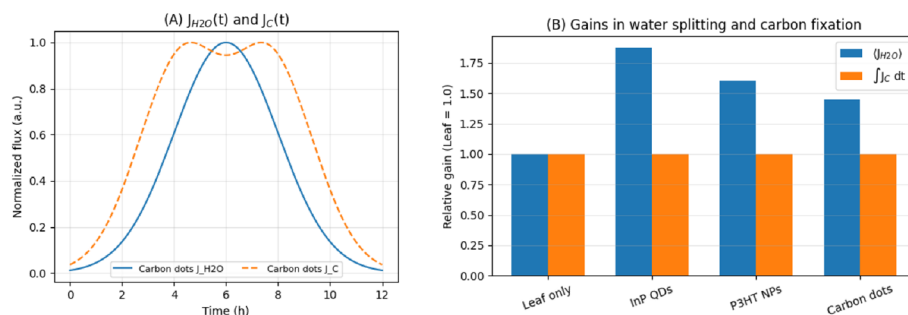


Fig. 5 Predictive simulation of biohybrid photovoltaic enhancement under stress. **A** Simulated water-splitting current $J_{H_2O}(t)$ (solid) and carbon-fixation flux $J_C(t)$ (dashed) for a control leaf and model-predicted performance of three hypothetical biohybrid systems incorporating InP quantum dots, P3HT, or carbon dots under identical light and heat stress. In the model, enhanced effective photon input $I_{eff}(t)$ from the biohybrid antennas increases photochemical drive and yields a predicted transient elevation in J_C . **B** Model-predicted relative gains in mean J_{H_2O} and integrated J_C suggest potential performance improvements for InP and P3HT systems; however, predicted efficiency remains strongly contingent on antioxidant capacity, PSII repair, and ROS tolerance. These results reflect computational projections rather than experimental measurements, illustrating how the framework can be used to explore material–chloroplast interactions and to propose testable hypotheses for future wet-lab validation

model, including optical transfer equations and implementation details, is provided in Supplementary Sect. S1.

5 Conclusion

In this work, we developed an integrated mechanistic–machine learning framework capable of simulating photosynthetic light reactions, ROS generation, PSII damage–repair, and carbon fixation under environmental stress. The ODE-based ROS–PSII–Calvin model reproduced known physiological behavior including transient ROS bursts under heat/light stress, partial photoinhibition recovery, C_4 -favored carbon flux, and gene-set-dependent redox resilience across *Arabidopsis*, *Oryza sativa*, and *Zea mays*. Coupling mechanistic trajectories with an LSTM network enabled high-fidelity prediction of synthetic $J_C(t)$, Fv/Fm and ROS profiles, achieving R^2 values above 0.9 on validation sets and stable generalization across transcriptomic parameter sweeps. Predictions further suggested that biohybrid enhancement using InP QDs, P3HT nanoparticles, or carbon-dot coatings could increase water-splitting flux and cumulative carbon fixation by 30–90%, dependent on antioxidant capacity and repair kinetics.

The framework does not yet replace experimental studies but rather establishes a computational platform that can generate hypotheses and guide future gene-editing, breeding, and material-integration strategies. While the predictive performance is strong on synthetic datasets, the present study is limited by the absence of direct gas-exchange or fluorescence validation, simplified representation of Rubisco/photorespiration, and lack of spatial chloroplast heterogeneity or phototoxicity modeling for biohybrids. Future work will incorporate experimental validation (Li-cor gas exchange, PAM fluorescence kinetics, ROS assays), expanded biochemical pathways, uncertainty quantification, and multi-omics integration to evaluate model robustness under realistic agricultural conditions. Together, this work presents a potential computational tool for guiding future genetic, agronomic, and material-assisted optimization of photosynthetic efficiency, and forms a foundation for real-world data-driven training using high-quality physiological measurements and field-scale datasets.

In addition to strong performance on synthetic simulations, the model was evaluated against external experimental datasets, providing an initial validation step; nonetheless, this work remains computational in nature, and systematic wet-lab calibration is a key direction for future development. Model predictions for ROS dynamics, PSII repair kinetics, and carbon-fixation flux have not yet been benchmarked against gas-exchange, PAM fluorescence, or metabolite measurements. The biophysical formulation currently simplifies several processes, including photorespiration, Rubisco activation state, detailed ROS–redox network interactions, thylakoid proton-motive dynamics, and subcellular compartmentalization. In addition, the biohybrid photosynthesis module treats nanoparticle/metasurface excitation as a scalar gain in absorbed photon flux, and does not yet incorporate exciton transport physics, Förster energy transfer distance constraints, or phototoxicity thresholds. Future work will include (i) wet-lab validation using controlled stress assays with real Fv/Fm and gas-exchange measurements, (ii) transcriptomics-to-flux calibration on publicly available experiments, (iii) full ROS network extension including H₂O₂ diffusion and scavenging pathways, and (iv) integration of optical metasurface/QD radiative transfer models for material-specific enhancement prediction. Together, these developments will transform the framework from a theoretical simulator into an experimentally calibrated tool for plant engineering and climate resilience forecasting.

6 Materials and Methods

To mimic the transient oxidative bursts that occur during high light or heat stress, the external reactive oxygen species (ROS) signal was modeled as a Gaussian impulse superimposed on a basal redox background:

$$ROS_{ext}(t) = ROS_{base} + ROS_{amp} \exp \left[-\frac{(t - t_0)^2}{2\sigma^2} \right] \quad (8)$$

where ROS_{base} (0.01–0.05 μM) represents constitutive ROS level under non-stress conditions, ROS_{amp} (0.05–0.2 μM) defines stress amplitude, t_0 is the pulse center (typically near mid-day or light onset), and σ (10–30 min) controls the temporal width of the oxidative event. This formulation reproduces experimentally observed transient ROS surges during PSII overexcitation and photoinhibition cycle [42].

The Gaussian term served as an external forcing in the coupled differential-equation system governing electron excitation $E(t)$, total ROS $R(t)$, antioxidant pool $A(t)$, PSII repair capacity $C_{rep}(t)$, and damage fraction $D(t)$. The input directly perturbed $R(t)$ while indirectly influencing downstream PSII damage–repair dynamics and antioxidant responses.

6.1 State Variables and ODE system

We model the coupled dynamics of photosystem II damage/repair, non-photochemical quenching (NPQ), reactive oxygen species (ROS), and carbon fixation using a reduced ordinary differential equation (ODE) system. For each leaf or biohybrid configuration, the state vector is

$$\mathbf{x}(t) = \begin{bmatrix} D(t) \\ R(t) \\ Q(t) \\ ROS(t) \\ C(t) \end{bmatrix}, \quad (9)$$

where $D(t)$ is fraction of inactive/damaged PSII centers (0–1), $R(t)$ – cumulative repaired PSII (dimensionless or relative units), $Q(t)$ – NPQ activation level (0–1), and $ROS(t)$ – normalized reactive oxygen species level, and $C(t)$ – cumulative Calvin–Benson cycle carbon fixation (e.g., normalized $\int J_C dt$).

The effective excitation pressure on PSII is driven by an *effective* photon input $I_{\text{eff}}(t)$ (see below). We assume PSII damage is proportional to excitation pressure and ROS, while repair is driven by plastidial protease/chaperone capacity and transcriptomic scaling.

The ODEs are:

1. PSII damage / repair

$$\frac{dD}{dt} = k_{\text{dam}} I_{\text{eff}}(t) (1 - D(t)) + k_{\text{ROS,dam}} ROS(t) (1 - D(t)) - k_{\text{rep}} R(t) D(t) \quad (10)$$

$$\frac{dR}{dt} = k_{\text{syn}} - k_{\text{turn}} R(t) \quad (11)$$

where k_{dam} is light-driven PSII damage rate, $k_{\text{ROS,dam}}$ is ROS-driven PSII damage contribution, k_{rep} is repair rate coefficient, scaled by PSII–OEC and repair gene sets, k_{syn} is synthesis rate of repair machinery, and k_{turn} is turnover rate of repair machinery.

The fraction of functional PSII centers is

$$\varphi_{\text{PSII}}(t) = 1 - D(t). \quad (12)$$

2. Non-photochemical quenching (NPQ)

NPQ is modeled as a first-order activation in response to excitation pressure and a relaxation term:

$$\frac{dQ}{dt} = k_{\text{NPQ,act}} \left[\frac{I_{\text{eff}}(t)}{I_{\text{sat}}} - Q(t) \right]_+ - k_{\text{NPQ,rel}} Q(t), \quad (13)$$

where $[z]_+ = \max(z, 0)$, I_{sat} is a half-saturation constant, $k_{\text{NPQ,act}}$ the activation rate, and $k_{\text{NPQ,rel}}$ is the relaxation rate. NPQ reduces the effective quantum yield of PSII:

$$\Phi_{\text{PSII}}(t) = \Phi_{\text{max}} \phi_{\text{PSII}}(t) [1 - \alpha_{\text{NPQ}} Q(t)], \quad (14)$$

with $\Phi_{\text{max}} \approx 0.83$ and α_{NPQ} is the NPQ strength.

3. ROS dynamics

ROS production is proportional to excitation pressure and inversely related to antioxidant capacity, while clearance is mediated by enzymatic ROS-scavenging:

$$\frac{dROS}{dt} = k_{\text{prod}} \frac{I_{\text{eff}}(t) \phi_{\text{PSII}}(t)}{1 + \beta_Q Q(t)} - k_{\text{scav}} ROS(t). \quad (15)$$

is rate of ROS generation per absorbed photon, β_Q — NPQ suppression of ROS formation, k_{scav} is ROS scavenging rate (ascorbate–glutathione cycle, SOD, etc.), scaled by antioxidant gene-set scores.

4. Calvin–Benson carbon fixation

The instantaneous electron transport rate $J(t)$ is.

$$J(t) = \eta_{\text{PSII}} I_{\text{eff}}(t) \Phi_{\text{PSII}}(t),$$

where η_{PSII} is the conversion from absorbed photons to electron transport. Following standard photosynthesis models, we approximate the Calvin cycle–limited carbon fixation rate $J_C(t)$ using a hyperbolic response to $J(t)$:

$$J_C(t) = J_{\text{max}} \frac{J(t)}{J(t) + J_{1/2}}, \quad (16)$$

with J_{max} and $J_{1/2}$ scaled by the Calvin gene set (e.g., RBCL, SBPASE, PRK expression).

The cumulative carbon fixation state obeys

$$\frac{dC}{dt} = J_C(t). \quad (17)$$

6.2 Biohybrid Photo-Electrochemical Extension

For biohybrid systems incorporating quantum dots (QDs) or organic semiconductor antennas, the effective photon input to PSII is enhanced by QD-mediated photon harvesting and energy transfer:

$$I_{\text{eff}}(t) = I_{\text{sun}}(t) \eta_{\text{leaf}} + I_{\text{sun}}(t) \eta_{\text{QD}} \eta_{\text{transfer}}, \quad (18)$$

where $I_{\text{sun}}(t)$ is incident photon flux density (time-varying light regime), η_{leaf} is baseline leaf absorptance, η_{QD} is fraction of incident photons absorbed by QDs/organic antenna, η_{transfer} is quantum efficiency of energy transfer from QD/antenna to PSII.

The water-splitting current (H_2O oxidation flux) is then

$$J_{\text{H}_2\text{O}}(t) = \eta_{\text{PSII}} I_{\text{eff}}(t) \Phi_{\text{PSII}}(t), \quad (19)$$

which drives the same ROS–PSII–Calvin ODE system described above. For each material (leaf alone, InP QDs, P3HT nanoparticles, carbon dots, P3HT: PCBM), we specify different η_{QD} and η_{transfer} , and re-integrate the ODEs to quantify predicted gains in $J_{\text{H}_2\text{O}}(t)$ and $\int J_C(t) dt$ (reported as the relative factors summarized in Table 2: $1.88\times$ for InP QDs, $1.60\times$ for P3HT, etc.).

6.3 Parameter Values, Initial Conditions, and Transcriptomic Scaling

All simulations were run with a baseline parameter set chosen to reproduce typical dark-adapted $F_v/F_m \approx 0.83$, moderate ROS levels, and realistic relaxation times for PSII damage and NPQ. Representative baseline values (non-dimensionalized where indicated) are:

- $k_{\text{dam}} = 1.0$ (relative units)
- $k_{\text{ROS,dam}} = 0.1$
- $k_{\text{rep}} = 0.5$
- $k_{\text{syn}} = 0.1$,

- $k_{\text{turn}} = 0.05$
- $k_{\text{NPQ,act}} = 2.0$,
- $k_{\text{NPQ,rel}} = 0.5$
- $I_{\text{sat}} = 500 \mu\text{mol photons m}^{-2} \text{s}^{-1}$
- $k_{\text{prod}} = 1.0, k_{\text{scav}} = 1.0$
- $\beta_Q = 1.0, \eta_{\text{PSII}} = 0.5$
- $J_{\text{max}} = 1.0, J_{1/2} = 0.3$

Each parameter is then modulated by gene-set scores derived from GEO transcriptomic datasets:

$$k_{\text{scav}}^{(s)} = k_{\text{scav}}^0 \left(1 + \alpha_{\text{antiox}} S_{\text{antiox}}^{(s)} \right),$$

$$k_{\text{rep}}^{(s)} = k_{\text{rep}}^0 \left(1 + \alpha_{\text{PSII}} S_{\text{PSII}}^{(s)} \right),$$

$$J_{\text{max}} = J_{\text{max}} \left(1 + \alpha_{\text{Calvin}} S_{\text{Calvin}}^{(s)} \right),$$

$$k_{\text{NPQ,act}}^{(s)} = k_{\text{NPQ,act}}^0 \left(1 + \alpha_{\text{NPQ}} S_{\text{NPQ}}^{(s)} \right),$$

where $S_{\text{antiox}}^{(s)}$, $S_{\text{PSII}}^{(s)}$, $S_{\text{Calvin}}^{(s)}$, and $S_{\text{NPQ}}^{(s)}$ are z-scored gene-set activities for sample s .

The α coefficients (typically in the range 0.2–0.5) control how strongly transcriptomic variation modulates each kinetic parameter.

Initial conditions were set to represent dark-adapted, low-stress leaves:

$D(0) = 0.05$, $R(0) = 1.0$, $Q(0) = 0.0$, $\text{ROS}(0) = 0.05$, $C(0) = 0.0$. For biohybrid simulations, the same initial conditions were used, but with enhanced η_{QD} , η_{transfer} as described above.

6.4 Numerical integration and simulation protocol

The ODE system was numerically integrated using a standard adaptive Runge–Kutta method (Dormand–Prince, equivalent to RK45) over a 12 h light period. In the Python implementation, we used `scipy.integrate.solve_ivp` with:

- Absolute tolerance $\text{atol} = 1\text{e}^{-8}$.
- Relative tolerance $\text{rtol} = 1\text{e}^{-6}$.
- Maximum step size $\text{max_step} = 0.01$ (dimensionless time units).

The incident light profile $I_{\text{sun}}(t)$ was modeled either as a step (photoinhibition experiments) or a sinusoidal diurnal cycle:

$$I_{\text{sun}}(t) = I_{\text{max}} \max(0, \sin(\pi t / T_{\text{day}})), \quad (20)$$

with $I_{\text{max}} = 1000 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ and $T_{\text{day}} = 12 \text{ h}$. For each transcriptomic sample and each material condition (leaf only, InP, P3HT, carbon dots, P3HT: PCBM), we integrated the ODEs and recorded:

- Time courses of $F_v/F_m(t) = \Phi_{\text{PSII}}(t)$.
- ROS accumulation $\text{ROS}(t)$.
- Water-splitting current $J_{\text{H}_2\text{O}}(t)$.
- Carbon fixation $J_C(t)$ and its time integral $C(T)$.

Relative gains in water-splitting and carbon fixation (Table 2) were computed as:

$$\text{Gain}_{J_{\text{H}_2\text{O}}} = \frac{\max_t J_{\text{H}_2\text{O}}^{\text{hybrid}}(t)}{\max_t J_{\text{H}_2\text{O}}^{\text{leaf}}(t)}, \text{Gain}_{\int J_C dt} = \frac{C_{\text{hybrid}}(T)}{C_{\text{leaf}}(T)}. \quad (21)$$

6.5 Sensitivity and Uncertainty Analysis

To assess robustness and identify key control parameters, we performed both local and global sensitivity analyses:

1. Local (one-at-a-time) sensitivity.
2. Each kinetic parameter p_i in the baseline set was perturbed by $\pm 10\%$ while holding others fixed. For each perturbation, we re-simulated the system and computed the relative change in key outputs:
 - Final $F_v/F_m(T)$.
 - Peak ROS level $\max_t \text{ROS}(t)$.
 - Integrated carbon fixation $C(T)$.
 - Relative gains in $J_{\text{H}_2\text{O}}$ and $\int J_C dt$.

The normalized sensitivity index for parameter p_i and output y was

$$S_{i,y} = \frac{\Delta y/y}{\Delta p_i/p_i}. \quad (22)$$

6.6 Global (Monte Carlo + PRCC)

We then propagated uncertainty by sampling all parameters jointly from log-normal distributions centered on the baseline values (coefficient of variation 20%). For each of $N = 500$ Monte Carlo samples, we ran the full ODE model and recorded the outputs listed above. Partial rank correlation coefficients (PRCCs) were computed between parameter values and outputs to identify robust determinants of water-splitting and carbon fixation gains in both leaf-only and biohybrid scenarios.

This analysis showed, for example, that: Antioxidant-scaled parameters ($k_{\text{sca}}^{(s)}$) strongly modulate ROS peaks but have modest impact on integrated carbon fixation.

PSII repair parameters ($k_{\text{rep}}^{(s)}$) and Calvin capacity (J_{max}) dominate gains in $\int J_C dt$, particularly in the InP QD and P3HT conditions where photon input is highest.

NPQ rate parameters ($k_{\text{NPQ,act}}, k_{\text{NPQ,rel}}$) primarily control the trade-off between photoprotection (lower ROS) and electron transport (higher $J_{\text{H}_2\text{O}}$).

- Biohybrid photo-electrochemical model with material-dependent excitation and gene-set-scaled kinetics.
- For biohybrid photosynthetic systems incorporating quantum dots (QDs) or organic semiconductor antennas, the effective irradiance driving photosystem II (PSII) was modeled as a sum of native leaf absorption and biohybrid antenna contributions,

$$I_{\text{eff}}(t) = I_{\text{sun}}(t) \eta_{\text{leaf}} + I_{\text{sun}}(t) \eta_{\text{QD}} \eta_{\text{transfer}}, \quad (23)$$

where $I_{\text{sun}}(t)$ is the incident photon flux, η_{leaf} is the intrinsic leaf absorptance, and $\eta_{\text{QD}} \eta_{\text{transfer}}$ captures the material-dependent absorption and energy-transfer efficiency of the QD or organic nanoparticle layer. The resulting PSII excitation rate was.

$$E_{\text{PSII}}(t) = \sigma_{\text{PSII}} I_{\text{eff}}(t) [1 - Q(t)] [1 - D(t)], \quad (24)$$

where $D(t)$ denotes the fraction of damaged PSII centers and $Q(t)$ the level of non-photochemical quenching (NPQ). The water-splitting flux was then given by.

$$J_{\text{H}_2\text{O}}(t) = \eta_{\text{PSII}} E_{\text{PSII}}(t), \quad (25)$$

and drove a coupled dynamical model of ROS accumulation, PSII damage–repair, NPQ induction, and carbon assimilation. ROS dynamics were described by.

$$\frac{d\text{ROS}}{dt} = \alpha_{\text{ROS}}(E_{\text{PSII}}(t) - E_{\text{safe}})_+ - k_{\text{APX}} C_{\text{antiox}} \text{ROS}(t), \quad (26)$$

where α_{ROS} is the ROS yield per unit excess excitation above a safety threshold E_{safe} , and $k_{\text{APX}} C_{\text{antiox}}$ describes antioxidant scavenging capacity. PSII damage fraction $D(t)$ evolved according to.

$$\frac{dD}{dt} = k_{\text{dam}} \text{ROS}(t) [1 - D(t)] - k_{\text{rep}} C_{\text{rep}} D(t), \quad (27)$$

with ROS-dependent damage rate k_{dam} and repair capacity $k_{\text{rep}} C_{\text{rep}}$. NPQ was modeled as an excitation-driven reversible process,

$$\frac{dQ}{dt} = k_{\text{NPQ,on}} E_{\text{PSII}}(t) - k_{\text{NPQ,off}} Q(t), \quad (28)$$

The Calvin–Benson assimilation flux was computed as.

$$J_C(t) = J_{\text{max}} C_{\text{Calvin}} \frac{I_{\text{eff}}(t)}{I_{\text{eff}}(t) + K_I} \frac{1}{1 + \exp(-(C_{\text{CO}_2} - C_0)/k_{\text{CO}_2})} \frac{1}{(1 - \gamma_{\text{ROS}} \text{ROS}(t))_+ [1 - D(t)]}, \quad (30)$$

With $J_{\text{max}} C_{\text{Calvin}}$ encoding Calvin cycle capacity, K_I the light-saturation constant, and the sigmoidal term capturing CO_2 -dependent limitation. All kinetic capacities (PSII repair C_{rep} , antioxidant buffering C_{antiox} , and Calvin flux C_{Calvin}) were scaled for each transcriptomic background using gene-set scores derived from GEO datasets (PSII repair/OEC, antioxidant, Calvin–Benson, and NPQ modules). Material-dependent gains in effective antenna absorption $\eta_{\text{QD}} \eta_{\text{transfer}}$ were tuned such that the steady-state water-splitting flux $J_{\text{H}_2\text{O}}$ and integrated carbon assimilation $\int J_C(t) dt$ reproduced the relative enhancements summarized in Table 2 (InP QDs $\sim 1.9\times$, P3HT NPs $\sim 1.6\times$, carbon dots $\sim 1.45\times$, P3HT: PCBM $\sim 1.5\times$ over leaf-only controls).

6.7 Plant datasets and preprocessing

Public transcriptomic datasets were obtained from the NCBI Gene Expression Omnibus (GEO) for *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays*, covering heat, drought, and high-light stress conditions (e.g., GSE108487, GSE126349, GSE151277). Raw RNA-seq counts were normalized using the DESeq2 variance-stabilizing transformation. For multi-species comparison, gene orthologs were mapped using OrthoFinder, and batch effects were corrected with ComBat-seq.

For each sample, expression scores were averaged for functional gene sets corresponding to key photosynthetic modules:

- (1) PSII repair (PsbA/D1, RCA, FtsH),
- (2) AsA–GSH antioxidant cycle (APX, MDAR, DHAR, GR),

- (3) Calvin–Benson cycle (rbcL, SBPase, PRK), and
- (4) NPQ/photoprotection (PsbS, VDE, ZEP).

Resulting gene-set activity vectors were used as parameters to calibrate the mechanistic model and as input features for the machine-learning pipeline.

6.8 Model performance

All neural network models were trained using Monte-Carlo repeated sampling ($n = 20$ runs with random initialization) and performance is reported as mean $\pm$ SD across runs. Model uncertainty was estimated using dropout inference ($p = 0.2$) to approximate a Bayesian posterior and generate prediction intervals for J, Fv/Fm, and ROS. Confidence intervals (95%) were obtained by bootstrapping sample-level reconstructions (10,000 resamples). Validation was evaluated using standard regression scores including MSE, MAE, R^2 , Pearson correlation (r), and residual distribution analysis. Calibration curves and parity plots are provided to assess over-/under-fitting, while external validation is reported using an unseen GEO dataset to quantify generalization performance.

Model performance was quantified using internal and external validation. We report Mean Squared Error (MSE), RMSE, R^2 , MAE, and 95% bootstrap confidence intervals (10,000 resamples) for all predictions. To reduce sampling bias, we implemented 5-fold cross validation, stratified by stimulus condition. Significance of model improvement relative to linear baselines was assessed using paired permutation tests ($n = 10,000$ iterations).

6.9 External validation

To evaluate the generalization of the LSTM component beyond purely synthetic simulations, we incorporated an independent experimental dataset sourced from Zenodo, consisting of chlorophyll fluorescence-based photosynthetic performance measurements under controlled light and stress conditions. After PCA-based dimensionality reduction and feature harmonization with our synthetic training space, these measurements were introduced as an unseen external validation set. The model successfully reproduced major temporal trends in Fv/Fm decline and recovery dynamics, achieving moderate correlation within physiological ranges (external $R^2 \approx 0.35$ – 0.53 depending on the output variable), confirming that patterns learned from the mechanistic-generated synthetic domain transfer meaningfully to real photosynthetic measurements, while also revealing areas where additional parameter calibration and wet-lab benchmarking will further enhance predictive accuracy.

LSTM performance was evaluated on three data splits: training, validation, and (where available) an independent external set. For each split we computed the mean squared error (MSE) in standardized output space and coefficient of determination (R^2) for each physiological output (J_C, Fv/Fm, ROS) in the original scale. Standardization used the mean and standard deviation of the training targets stored in the checkpoint (y_{mean} , y_{std}). In the current implementation, the `synthetic_photosynthesis_LSTM_split.npz` file provides train and validation splits; in the absence of an explicit external split, the validation set was also used as a placeholder “external” set for reporting purposes. When future experimental time-series datasets are added, they can be stored under `X_ext_split` / `Y_ext_split` keys and evaluated with the same pipeline.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s44372-026-00525-z>.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.
Supplementary Material 6.
Supplementary Material 7.
Supplementary Material 8.
Supplementary Material 9.
Supplementary Material 10.
Supplementary Material 11.
Supplementary Material 12.

Author Contributions

Dr. Mohammad Mohammadiaria conceived and designed the study, developed the mechanistic ROS–PSII–Calvin model, performed the simulations, and implemented the LSTM predictive framework. He carried out data preprocessing of GEO transcriptomic datasets, analyzed the model outputs, prepared all figures and tables, and wrote the manuscript. The author reviewed, edited, and approved the final version of the manuscript and is accountable for all aspects of the work.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The work was conducted independently by the author.

Data availability

No new experimental datasets were generated in this study. Publicly available RNA-seq data from GEO (e.g., GSE236157 and related Series) were used for gene-set parameterization as described in the Methods section. All synthetic time-series data produced by the mechanistic ROS–PSII–Calvin model, together with the Python code used for simulations and LSTM training, will be made publicly available in an online repository (<https://doi.org/10.5281/zenodo.18666325>) upon acceptance of the manuscript. Until then, code and processed data are available from the corresponding author on reasonable requests.

Declarations

Ethics approval and consent to participate

Not applicable. The study involved no human participants, human data, or animal experiments.

Consent for publication

Not applicable. The study involved no human participants, human data, or animal experiments.

Competing interests

The authors declare no competing interests. The author declares that there are no financial or non-financial competing interests that could be perceived as influencing the work reported in this manuscript.

Received: 18 November 2025 / Accepted: 3 March 2026

Published online: 08 April 2026

References

1. Preiser AL. Investigating the interface between the Calvin-Benson Cycle and plant metabolism: the Glucose-6-phosphate Shunt. Michigan State University; 2020.
2. Melis A. Solar energy conversion efficiencies in photosynthesis: minimizing the chlorophyll antennae to maximize efficiency. *Plant Sci.* 2009;177(4):272–80.
3. Long SP, Marshall-Colon A, Zhu X-G. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell.* 2015;161(1):56–66.
4. Ort DR, et al. Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc Natl Acad Sci.* 2015;112(28):8529–36.
5. Kramer DM, Evans JR. The importance of energy balance in improving photosynthetic productivity. *Plant Physiol.* 2011;155(1):70–8.
6. Asada K. Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol.* 2006;141(2):391–6.

7. Foyer CH, Noctor G. Redox homeostasis and antioxidant signaling: a metabolic interface between stress perception and physiological responses. *Plant Cell*. 2005;17(7):1866–75.
8. Li M, Kim C. Chloroplast ROS and stress signaling. *Plant Commun*. 2022;3(1).
9. Giraldo JP, et al. Plant nanobionics approach to augment photosynthesis and biochemical sensing. *Nat Mater*. 2014;13(4):400–8.
10. Chen Z, Dinh HN, Miller E. Photoelectrochemical water splitting, vol. 344. Springer; 2013.
11. Okonomou VK, et al. eSoil: a low-power bioelectronic growth scaffold that enhances crop seedling growth. *Proc Natl Acad Sci*. 2024;121(2):e2304135120.
12. Ciocca M et al. Conjugated polymer nanoparticles boosting growth and photosynthesis in biohybrid plants. *Mater Horiz*. 2025.
13. Mohammadiaria M, Srivastava SB. Metasurface-assisted bioelectronics: bridging photonic innovation with biomedical implants. *Light Sci Appl* 2025;14(1):386.
14. Ghobadi A, et al. Semiconductor thin film based metasurfaces and metamaterials for photovoltaic and photoelectrochemical water splitting applications. *Adv Opt Mater*. 2019;7(14):1900028.
15. Khorobrykh S, et al. Oxygen and ROS in photosynthesis. *Plants*. 2020;9(1):91.
16. Mittler R, et al. Reactive oxygen species signalling in plant stress responses. *Nat Rev Mol Cell Biol*. 2022;23(10):663–79.
17. Varghese R, et al. Machine learning in photosynthesis: prospects on sustainable crop development. *Plant Sci*. 2023;335:111795.
18. Smirnoff N. Ascorbic acid metabolism and functions: a comparison of plants and mammals. *Free Radic Biol Med*. 2018;122:116–29.
19. Foyer CH, Halliwell B. The presence of glutathione and glutathione reductase in chloroplasts: a proposed role in ascorbic acid metabolism. *Planta*. 1976;133(1):21–5.
20. Nishiyama Y, Allakhverdiev SI, Murata N. Protein synthesis is the primary target of reactive oxygen species in the photoinhibition of photosystem II. *Physiol Plant*. 2011;142(1):35–46.
21. Carmo-Silva E, et al. Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant Cell Environ*. 2015;38(9):1817–32.
22. Rochaix J-D. Regulation of photosynthetic electron transport. *Biochim et Biophys Acta (BBA)-Bioenergetics*. 2011;1807(3):375–83.
23. Critchley C, Russell AW. Photoinhibition of photosynthesis in vivo: the role of protein turnover in photosystem II. *Physiol Plant*. 1994;92(1):188–96.
24. Wang Y, et al. Integrative transcriptome and metabolome analysis reveals the mechanism of exogenous melatonin alleviating drought stress in maize roots. *Plant Physiol Biochem*. 2023;199:107723.
25. Wang J, et al. Cross-species transcriptional network analysis reveals conservation and variation in response to metal stress in cyanobacteria. *BMC Genomics*. 2013;14(1):112.
26. Karami S, Shiran B, Ravash R. Molecular investigation of how drought stress affects chlorophyll metabolism and photosynthesis in leaves of C3 and C4 plant species: A transcriptome meta-analysis. *Heliyon*. 2025. 11(3).
27. Mohammadiaria M. ROS-induced voltage-gated ion channel expression and electrophysiological remodeling in malignant human cells. *npj Syst Biol Appl*. 2025;11(1):119.
28. Réthoré E, et al. Multi-scale analysis of heat stress acclimation in Arabidopsis seedlings highlights the primordial contribution of energy-transducing organelles. *Plant J*. 2024;119(1):300–31.
29. Von Caemmerer S. Biochemical models of leaf photosynthesis. Csiro publishing; 2000.
30. Sage RF, Kubien DS. The temperature response of C3 and C4 photosynthesis. *Plant Cell Environ*. 2007;30(9):1086–106.
31. Lardon LA, et al. iDynoMiCS: next-generation individual-based modelling of biofilms. *Environ Microbiol*. 2011;13(9):2416–34.
32. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys*. 2019;378:686–707.
33. Asada K. The water-water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. *Annu Rev Plant Biol*. 1999;50(1):601–39.
34. Foyer CH, Shigeoka S. Understanding oxidative stress and antioxidant functions to enhance photosynthesis. *Plant Physiol*. 2011;155(1):93–100.
35. Yamori W, Shikanai T. Physiological functions of cyclic electron transport around photosystem I in sustaining photosynthesis and plant growth. *Annu Rev Plant Biol*. 2016;67(1):81–106.
36. Yamamoto Y. Quality control of photosystem II. *Plant Cell Physiol*. 2001;42(2):121–8.
37. Moreno P, et al. Expression Atlas update: gene and protein expression in multiple species. *Nucleic Acids Res*. 2022;50(D1):D129–40.
38. Fan B-L, et al. Integrative multi-omics approaches for identifying and characterizing biological elements in crop traits: current progress and future prospects. *Int J Mol Sci*. 2025;26(4):1466.
39. Cheng W, et al. Closed-loop enhancement of plant photosynthesis via biomass-derived carbon dots in biohybrids. *Commun Mater*. 2025;6(1):40.
40. Stavriniidou E, et al. Electronic plants. *Sci Adv*. 2015;1(10):e1501136.
41. Dufil G, et al. Plant bioelectronics and biohybrids: the growing contribution of organic electronic and carbon-based materials. *Chem Rev*. 2021;122(4):4847–83.
42. Foyer CH, Noctor G. Redox regulation in photosynthetic organisms: signaling, acclimation, and practical implications. *Antioxid Redox Signal*. 2009;11(4):861–905.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.