Trends in Plant Science

CellPress OPEN ACCESS

Opinion

Digital twins for plant and crop improvement

Duke Pauli [1,2,*], Brennan Huppenthal[3], Diane Wang[4], James Schnable[5], and Bedrich Benes[6]

Plant digital twins (DTs) are dynamic, data-aware virtual representations that integrate plant structure, physiology, and environmental inputs to simulate plant function, growth, and performance. By combining process-based models, phenomics data, and environmental states in a continuous feedback loop linking physical and virtual plants, DTs offer a novel approach for investigating complex biological systems. In this opinion article, we outline the conceptual foundations of DTs and how they offer new opportunities to bridge the genotype-to-phenotype gap and aid in crop improvement. Despite their promise, major challenges remain in DT development, scalability, and accessibility. Addressing these challenges will be critical for enabling the broad adoption of DTs as transformative tools for accelerating plant science research and crop improvement.

Highlights

Digital twins are virtual plant models that have the potential to integrate genomic, phenomic, and environmental data to simulate plant function, growth, and responses to environmental changes over time.

Unlike traditional models, digital twins continuously update through real-time data assimilation, enabling adaptive learning and discovery of emergent biological properties as new data becomes available.

Integrating phenomics, AI-enabled data pipelines, and functional–structural plant models provides a scalable framework to bridge the genotype-to-phenotype gap.

Leveraging plant digital twins within crop growth models could reveal how plant-level traits scale to canopy function and impact performance and yield by increasing model fidelity and capturing heterogeneity.

Ultimately, digital twin technology could become a valuable research tool that enhances crop improvement by linking gene function to performance.

Need for digital twins

Digital twins (DTs) are an emerging technology designed to simulate and interrogate complex, dynamic systems such as those commonly encountered in biology. Although definitions vary across disciplines, DTs are broadly conceptualized as digital counterparts of physical entities, capable of mimicking their structure, function, and environmental responses, with their states continuously updated by data collected from the corresponding physical twin [1,2]. Originally developed and applied in engineering fields (Box 1), DTs show promise in life science fields with their ability to integrate various physical aspects of a real-world entity into a virtual representation.

DTs are not extensions of traditional or statistical modeling approaches but are more comprehensive and dynamic systems for integrating and interrogating models. They are a framework that integrates various functional components: data processing pipelines, including transmission, assimilation, and storage; mechanistic and process-based models; the ability to perform computational simulations utilizing these; and interfaces for monitoring the status of the DT. DTs are also composed of the software and cyberinfrastructure linking these functional components together [17]. By linking these functional components across temporal and/or spatial scales, DTs create an integrated system that enables the exploration of both predicted and unanticipated emergent behaviors that would otherwise not be captured by the individual components of the DTs [18–20]. Additionally, in contrast to current modeling approaches, DTs maintain a continuous data-feedback loop with their physical twin; as new data become available, such as phenotypic or environmental conditions, they are updated and recalibrated to refine their present state. By establishing a link between a physical and digital counterpart, we can directly compare and interrogate the system, which can reveal intricacies, unknown constraints, feedback loops, or new mechanistic properties that increase our understanding of underlying processes [21]. Because of these properties, DT technology is being widely explored across domains [22].

In biology, medicine is experiencing the most rapid adoption of DT technology. Notably, researchers have constructed DTs of human organs (http://projectbreatheasy.org/, https://onscale.com/blog/project-breatheasy-digital-twins-of-lungs-to-improve-covid-19-patients-outcomes/) [23–25]

[1]School of Plant Sciences, University of Arizona, Tucson, AZ, USA
[2]Center for Agroecosystem Research in the Desert (ARID), University of Arizona, Tucson, AZ, USA
[3]Department of Computer Science, University of Arizona, Tucson, AZ, USA
[4]Department of Agronomy, Purdue University, West Lafayette, IN, USA
[5]Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, USA
[6]Department of Computer Science, Purdue University, West Lafayette, IN, USA

*Correspondence: dukepauli@arizona.edu (D. Pauli).

Box 1. DTs history and usage

The concept of DTs dates back to the 1970s when NASA used physical replicas of spacecraft to test and prepare for missions [3,4]. Over time, the DT concept became more formalized as other industries started adopting it [5]. A main goal of DTs was to 'de-silo' the data and models of various functional areas describing a physical object so as to create autonomous objects that embody the information of their real-world counterparts—a so-called 'digital twin'. In manufacturing, products and entire production lines are modeled for performance optimization and predictive maintenance [6–9]. In healthcare, DTs replicate patient-specific organs or systems to support personalized treatment [10–13]. Urban planners are employing DTs of cities to test infrastructure changes and simulate the response to environmental stressors [14,15]. What distinguishes a DT from a traditional simulation is its foundation in (often) real-time or empirical data that enables the model to 'evolve' alongside its physical counterpart.

An example of DTs borrowed from engineering is the simulation of stress on a physical object, such as an aircraft wing. The DT of an existing wing can be coupled with a **finite element simulation**—a process that models how physical forces such as heat or vibration will impact an object. The key advantage of DTs is that they enable the exploration of 'what-if' scenarios without requiring the physical creation of the circumstances. DTs can therefore decrease the cost of experiments by enabling rapid simulation of processes that could be costly (damaging the wing through collision), uncovering insights that may be difficult or impossible to obtain through empirical trials such as extreme or rare environmental conditions (flying through a hurricane or other extreme winds), and provide results much faster than the real-world counterparts (e.g., simulating the 20-year usage of an airplane wing).

Plant DTs extend this concept to biological systems, representing plants or plant processes based on real data and aiming to capture the dynamic nature of biology. A plant DT is not just a visual 3D model of a plant; it is a data-driven system that integrates morphological, physiological, genetic, and environmental data [16]. The goal is to simulate how a real plant or collection of plants grow, function, and respond to their environment across time, including how different genotypes respond to the same set of growth parameters.

and are actively using them for predictive medicine and other applications [10,26–28]. In a particularly interesting development, researchers recently generated a set of 3800 cardiac DTs that revealed mechanistic functional patterns related to ischemic heart disease, which were emergent phenotypes of the simulated system, not detectable via imaging or electrocardiograms alone [25]. This example illustrates how DTs can provide new insights and understanding of complex systems not currently accessible through traditional approaches, and indeed highlights why researchers are pursuing DT technology to study other body systems and disorders [29].

Extending these efforts, scientists are incorporating genetic information into DT frameworks to link genotype to phenotype [25,30,31], a strategy that can be applied in plant systems to understand how genetics drive plant and crop performance under different growing environments. By bridging this so-called 'genotype-to-phenotype (G2P) gap', researchers and plant breeders can gain insight into how genes and gene functions scale to whole-plant performance across diverse environments [32]. With respect to crop improvement and agricultural productivity, closing this gap is fundamental to accelerating genetic gain for yield potential, enhancing climate resilience, and developing cultivars adapted to rapidly evolving production constraints [33]. Closing this gap in a timely manner is becoming critical for food security.

Modern agricultural systems have been highly successful—never have so few fed so many [34,35]. However, this success in food and crop production has come through the combination of high inputs of water, fertilizer, crop protection chemistries, and advanced agronomic practices, which are beginning to show diminishing returns [36,37]. Simultaneously, agricultural production environments are becoming increasingly volatile, with more frequent heat, drought, and extreme weather events destabilizing crop productivity, while fertilizer availability and costs are becoming unpredictable due to recurring supply constraints [38–41]. These pressures are also coalescing at a time when there is uncertainty about whether the current rates of crop improvement are capable of meeting projected global food demands. Indeed, the data suggest that our annual rate of genetic gain in many major crops remains at

Glossary

Finite element simulation: a numerical method used to simulate and analyze how physical systems respond under various conditions, such as stress, heat, fluid flow, or electromagnetic fields, by reducing the structure into small, simple shapes (e.g., voxels, tetrahedra, and triangles) on which numerical processes can be performed.
Functional–structural plant model (FSPM): a computational model that simultaneously represents 3D plant architecture and physiological/developmental processes. These models capture interactions between plant form (morphology) and function (physiology), where changes in the architecture alter the resource capture and physiological performance of a plant, which in turn drives changes to the plant structure over time [58]. Unlike purely equation-based modeling paradigms, FSPMs often explicitly represent plant architecture in space using visually representable systems, such as L-systems or meshes.
Genomic prediction (GP): a modern breeding technique that uses genome-wide molecular marker data to estimate phenotypic or breeding values before extensive phenotyping or field evaluation has occurred.
Phenomics: the large-scale study of an organism's phenotypes, describing the observable physical and biochemical traits of an organism and how they change in response to differences in genetics, as well as the environment.
Process-based physiological model: a type of model that describes the biological processes within a plant, such as photosynthesis, respiration, water uptake, and nutrient transport, based on the underlying physiological mechanisms as represented by first-principle equations, such as Fick's law of diffusion or Darcy's law of permeability.
Ray-tracing simulator: a computational method that uses physics-based simulation to model scene illumination based on material properties and light. It accounts for multiple bounces of light through the space and its interactions with surfaces via transmittance and reflectivity. It is commonly used for calculating the amount of light reaching the surface of an object.
Transpiration efficiency (TE): the ratio of carbon gain to water loss during

approximately 0.5 to 1.5%, below the estimated 2 to 2.4% annual increase needed to meet the projected 2050 global demand [42–44]. At the same time, conventional breeding pipelines are resource- and time-intensive, often requiring years of multi-environment testing to evaluate performance, and may not deliver solutions at the pace demanded by global challenges [45,46]. Given these challenges, there is an increasing need for crop cultivars that are resource-efficient, environmentally resilient, and can be commercialized at a more rapid pace. DTs offer a potential framework to accelerate crop improvement by enabling virtual systems that can simulate the complexities of gene-by-environment interaction while minimizing time and resource expenditures. This type of DT environment would enable virtual experimentation on how specific physiological, architectural, and developmental traits interact under different growing environments. As a result, DTs could support the breeding of crops with improved drought resilience, canopy light-use efficiency, and adaptation to future climatic conditions.

plant gas exchange. At the leaf level, TE is expressed as the net photosynthesis divided by the transpiration rate as a function of stomatal conductance. At the whole plant or canopy level, TE is defined as the biomass per unit of water transpired.

To develop effective DT technology, information-rich phenotypic data sets are needed. In the past decade, the plant science community has developed various **phenomics** (see Glossary) technologies to provide phenotypic data on the same scale and volume as that of genomic data [33,47,48]. With these advancements, the ability to capture plant form and function on a scale not previously accessible is providing a wealth of data for researchers, both at spatiotemporally resolved levels as well as at the population level. We therefore posit that DTs offer a powerful means to integrate these phenomics advances with existing capacities in genomics, AI, modeling, and computer science, providing a pathway to more effectively bridge the G2P gap and accelerate plant biology research and crop improvement.

Critically, DTs offer a mechanistic pathway for closing the G2P gap by embedding genotype-specific parameters within dynamic process-based models. Genetic variation controls physiological responses, phenological development, and architectural traits that, in turn, govern light interception, carbon assimilation, and resource partitioning. These interactions unfold across space and time, generating emergent properties that are often obscured by reductionist analyses. By simulating and estimating how genotype-dependent parameters scale from organ function to canopy performance, and iteratively updating the DT with phenotypic data, DTs enable G2P mapping.

Bringing DTs to life

Although the above-described cardiac DT example is illustrative of the potential applications of DTs, plant- and crop-level DTs are still in the early stages of development, with practical demonstrations only beginning to emerge and often exhibiting more basic utility (e.g., [49–51]). We argue that the current limitations are not a flaw in the concept but rather stem from the innate biological complexity of plants, the lack of algorithms for extracting information from data, and the diversity of expertise required to integrate physiology, genomics, phenomics, modeling, and computer science. However, despite these challenges, we believe a foundational first step is to integrate plant phenomics data and physiology into the larger DT system (Figure 1, Key figure). This can be achieved by incorporating physiological equations into models that simulate explicit 3D plant growth, commonly referred to as **functional–structural plant models (FSPMs)** (Box 2) [57–59]. We believe FSPMs offer a way to integrate phenomics, physiological, and environmental data to represent the physical twin's plant growth and function; however, FSPMs should not be conflated with DTs themselves, as DTs provide functionality beyond modeling.

To initialize the DT, various types of sensor-derived phenotypic data would be collected on the plant: 3D point-cloud data capturing architecture [60]; estimates of stomatal conductance via thermal imagery [61]; estimates of photosynthetic parameters based on solar-induced

Key figure

Creation and workflow of a plant-level digital twin

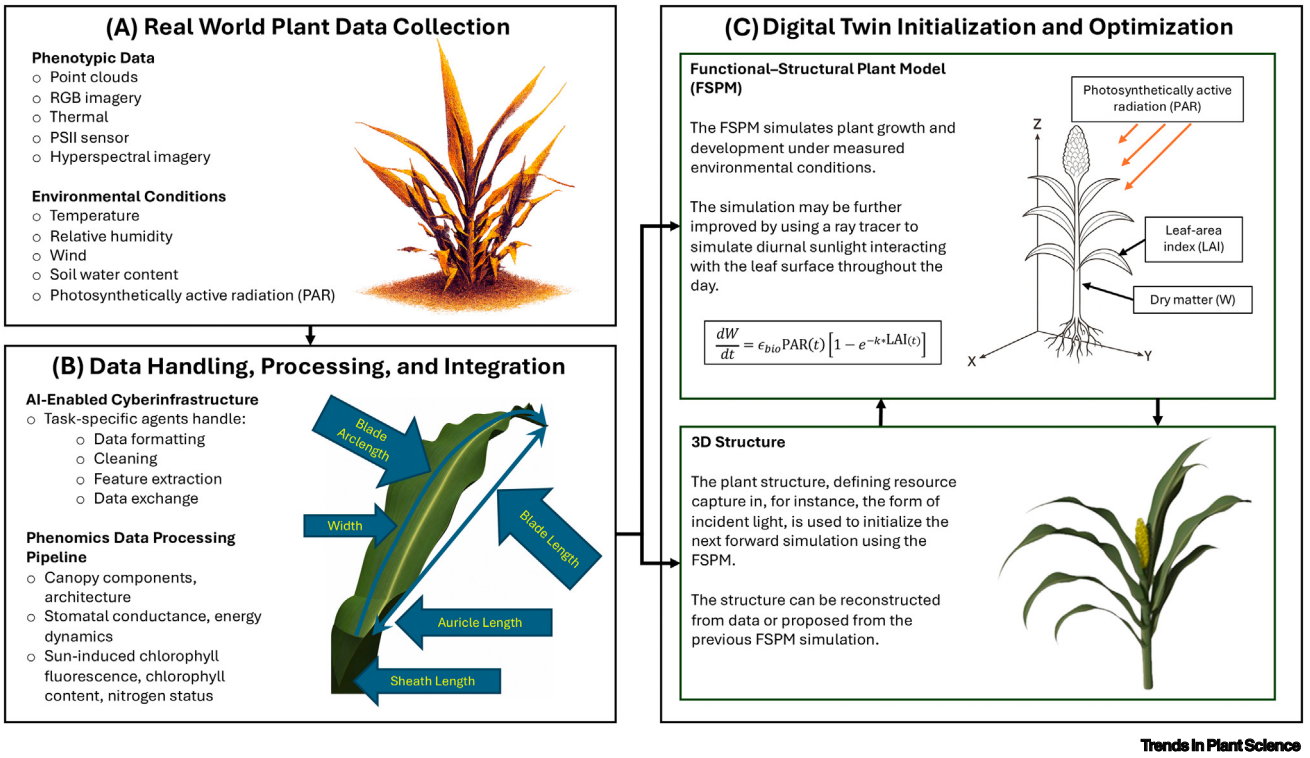

Figure 1. Simulations using digital twins are governed by functional–structural plant models (FSPMs), which codify how the 3D structure and function of a plant interact under given environmental conditions. There are three general steps involved in the creation and utilization of a digital twin to answer research questions. (A) First, raw data in the form of various sensor-derived phenotypic data and environmental conditions are measured and stored. (B) Next, this data is processed by an AI-enabled cyberinfrastructure, where task-specific agents handle data formatting, cleaning, filtering, and quality control. Upon completion, relevant phenotypic features are extracted by a phenomics data processing pipeline. (C) Finally, the 3D structure of the digital twin is parameterized by the extracted architectural measurements, which serve as the starting conditions for the FSPM simulation. The FSPM, in this case a radiation-driven biomass model, additionally takes as input the environmental conditions. Simulation consists of a feed-forward process in which the FSPM simulates growth and updates the plant's structure, which then defines how the plant interacts with the environment, providing new inputs to the model. These simulations can be further guided by providing additional phenotypic or environmental data collected at subsequent times.

fluorescence extracted from hyperspectral imagery [62,63]; and data on traits such as leaf count [64], phyllotaxy [65], and indices such as the normalized difference vegetation index (NDVI) [66] derived from red, green, and blue (RGB) or multispectral images. In parallel, weather data, soil moisture conditions, incoming sunlight, and other environmental data would also be collected. This raw data could be processed by AI-enabled cyberinfrastructure, where task-specific agents would handle formatting, filtering, perform quality control checks, temporal alignment of the data, point-cloud segmentation, and finally trait extraction. This process would convert heterogeneous measurements into meaningful parameters such as leaf area and angle distribution, branch length, volume, reflectivity, water status, and photosynthetic state.

Next, the DT would initialize an FSPM by instantiating a simplified but representative 3D model using the extracted architectural parameters after data processing and using extracted trait

Box 2. Modeling plants and crops

FSPMs link explicit 3D growth with physiology. This class of models consists of two primary components: (i) the architecture part, which describes the types and spatial organization of plant organs, such as leaves, stems, and roots, and (ii) the process part, which simulates physiological activities such as photosynthesis, hydraulic processes, and source–sink relationships [52]. By employing mathematical representations of plant processes, as commonly done in **process-based physiological models** [53], valuable insight can be gained into the dynamics of plant growth patterns, resource allocation, and plant–environment interactions. A distinguishing feature of FSPMs is their ability to model feedback loops between structural development and physiological function—for example, an increase in leaf area enhancing photosynthesis. Higher fidelity representations may capture this interaction better, albeit at the expense of complexity and computational demand. Given this functionality, FSPMs are well suited for plant DTs, offering detailed, mechanistic, and spatially explicit representations of plant behavior. They also capture temporal dynamics, allowing growth and development to be updated as environmental conditions shift hourly, daily, and seasonally, and as sensor data are assimilated [54]. Operating at the organ level, FSPMs can simulate fine-scale biological processes that occur within the plant canopy microclimate but impact crop performance.

CGMs are also a class of process-based models. However, they simulate crop development, resource capture, and yield formation at larger spatial scales and represent physiology at a less granular level [55]. CGMs represent key plant processes such as phenological development, light interception, biomass accumulation, and water and nutrient dynamics using mechanistic equations or empirically derived parameter values. A key distinction from FSPMs is that rather than modeling individual organs or plants, they typically treat the crop canopy as a unified system. By simulating plant processes at the canopy scale, CGMs can link environmental drivers (sunlight and temperature), soil conditions (water and nitrogen availability), and crop management practices (planting date, irrigation, and population density) with traits such as yield or radiation use efficiency [56]. By modeling the crop at the systems-level and utilizing the simulation capabilities of DTs, CGMs can simulate differing agronomic scenarios and responses to climatic variability that extend beyond the scale of individual plants. CGMs can provide a broader field-scale and environmental context within which FSPMs can be embedded, supplying boundary conditions, resource constraints, and whole-field feedback, enabling DTs to bridge organ-level processes with crop performance.

and state variables (derived either from sensor data or from a previous simulation run) to parameterize the model and set the starting and boundary conditions for simulation. Then, the DT would run the FSPM to simulate organ growth and development, resource capture, photosynthesis, and carbon assimilation through mechanistic source–sink relationships. To replicate light dynamics, the DT would incorporate a bi-directional **ray-tracing simulator** to track solar movement and compute the diurnal irradiance, shading, and light interactions with the variable leaf surfaces. The computed values would be fed back into the FSPM module to recalculate carbon assimilation and update the equations governing source–sink relationships, as well as growth modules, to simulate plant development.

At the next instance of data collection, the same trait parameters used for FSPM initialization would be updated, and a fit function would be evaluated. One simple example of a fit function is a weighted multi-trait sum of squares error, which provides a measure of how well the DT predicts multiple traits (such as leaf width, color, and yield) simultaneously. This would provide information on how well the DT replicates the physical twin, but more importantly, it could provide a key point for the integration of AI for DT improvement. One of the features that distinguishes DTs from traditional mechanistic models is their capacity for 'learning' and continuous updating through data assimilation. Through the use of AI, the DT could evolve alongside the physical twin and progressively refine its representation thereof [67]. This could occur through the use of agents or reinforcement learning that iteratively refine model parameter values so that processes not captured accurately by the FSPM can be identified and modified to more closely represent biological reality, as measured by the fit function. This feed-forward process could continue through the plant's lifecycle as data becomes available.

While there are certainly technical challenges to integrating the various components of the proposed system, the technical pieces, such as the phenomics data, AI-assisted data

processing pipelines, and FSPMs, are tools that are currently available [68,69]. Indeed, Smoleňová *et al*. [51] have successfully demonstrated a tomato DT that employs a similar architecture, demonstrating the utility and practicality of this approach to constructing functional DTs. The development and implementation of plant-level DTs are key first steps in leveraging this technology for crop improvement.

Bringing DT plants to the field

While single-plant DTs will be powerful tools for crop improvement, plants are not grown in isolation but instead form complex communities of individuals that must interact, share resources, and grow in coordinated fashion. As such, scaling from individual plant DTs to simulated crop communities is the logical next step (Figure 2). This advancement would help us better

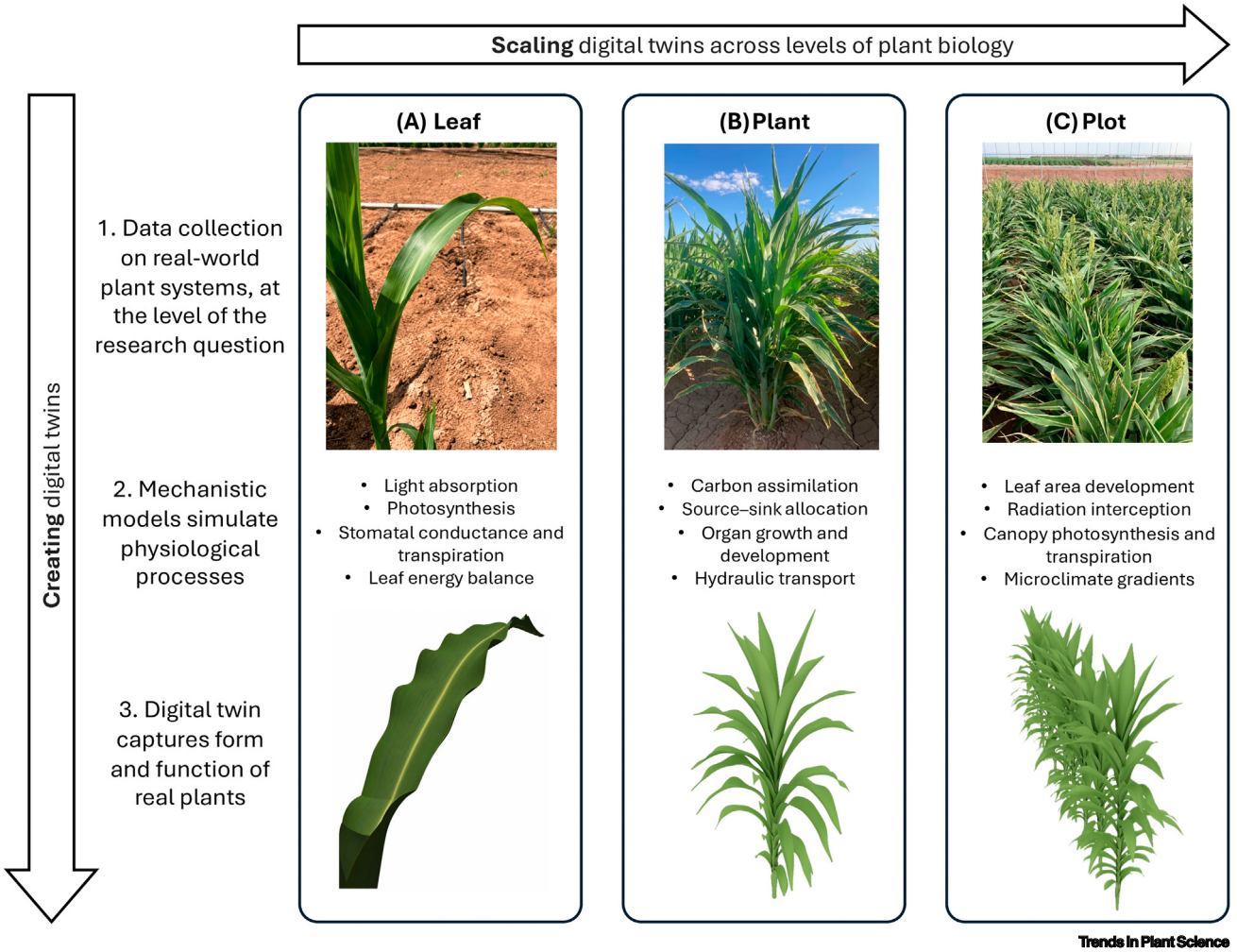

Figure 2. Application of plant digital twins across biological scales. Plant digital twins are virtual simulations of real plants that represent their structure, function, and dynamic response to environmental inputs in a feed-forward fashion [58,59,70]. From top to bottom, the steps for digital twin creation are (i) collecting morphological and physiological data from the real-world plant system, (ii) choosing which functional aspects are of interest and selecting an appropriate process-based physiological model, and (iii) utilizing the phenotypic data to instantiate and parameterize the physiological model, thus creating a virtual representation of the real-world plant. The three columns detail this process specifically for (A) leaf-level, (B) plant-level, and (C) plot-level plant systems, detailing the different functional aspects under consideration at each biological scale. Depending on the biological scale, different morphological traits and physiological states must be measured, and different functional aspects can be simulated using the mechanistic models, changing what inferences can be drawn from each level of digital twin.

understand how individual plant processes and traits scale to the crop level and give rise to emergent properties found in plant communities within a field [71]. By bridging from the plant to the community level, DTs could help reveal how the genetic basis of component traits (e.g., leaf elongation and expansion) impacts and contributes to higher-order traits such as canopy architecture and function (light interception, hydraulic, and photosynthesis gradients throughout the canopy). DTs that can represent plant-canopy dynamics would be valuable for developing plants that are more 'cooperative' and therefore more resource-efficient (Box 3). Additionally, DTs could help resolve disconnects between scales of biology. For example, while the mechanics of **transpiration efficiency (TE)** have been deeply characterized at the leaf level, efforts to translate that foundational knowledge to the plant and canopy levels have had limited success. In fact, plant traits that improve TE at the leaf level often result in yield penalties at the canopy scale [76]. The application of DTs to model and investigate the complex scaling of CO_2 and water vapor exchange in leaf-canopy systems could yield new insights into the physical processes that govern canopy TE dynamics.

To scale up from individual plants to communities, plant-level processes captured by FSPMs need to be linked to canopy-level traits, such as those typically captured by crop growth models (CGMs) (Box 2) [59,70]. Because CGMs represent crops as spatially averaged systems rather than explicit individual plants, they are better able to represent community-level processes that reflect agronomic outcomes [77]. However, a limitation is their inability to capture within-canopy heterogeneity, organ-level processes, and structural plasticity inherent to crop canopies. As a result, genotype-specific traits are approximated using empirical parameters rather than emerging mechanistically from plant structure and interactions with the within-canopy microclimate, as is the case with FSPMs. We argue that combining FSPMs and CGMs through the framework of DTs would provide a powerful tool for studying the complexity of crop canopies and their associated biology, which arises from the individual plants themselves.

Box 3. Optimizing crop canopy architecture

Optimizing crop canopy architecture and the physiological processes underlying photosynthesis (PS) remains a central strategy for improving crop productivity and resource use efficiency [72,73]. Sunlight interactions with the plant canopy are heterogeneous with respect to time, space, and light quality/quantity due to the interactions of leaf size, orientation, number, and spatial arrangement within the canopy [74]. Consequently, crop canopies frequently exhibit light use inefficiencies, with upper leaves operating beyond PS saturation while simultaneously restricting light penetration to lower, photosynthetically competent tissues. Because of this, the concept of 'smart canopies' has emerged [72]. These are canopies that optimize both canopy architecture and PS processes to reduce intra-canopy competition and improve PS efficiency throughout the canopy [75].

Because DTs provide a framework which couples plant architecture, physiology, dynamic light environments, and data assimilation into a unified whole, they are natural candidates for scaling from individual plants to larger plant systems, such as crop canopies. A canopy-level DT could comprise multiple whole-plant DTs, each representing a single plant within the canopy. Using data collected from a physical canopy, extracted architectural and physiological parameters could be used to initialize individual whole-plant DTs which are consistent with the canopy-level observations. With the whole-plant DTs initialized, the canopy-level DT could simulate the dynamic light environment throughout the canopy using ray tracing. The resulting irradiance would then be propagated to each whole-plant DT to update the individual FSPM models, enabling simulation of emergent within-canopy dynamics. Crucially, this would enable the quantification of how canopy architecture traits modulate intra-canopy light gradients and drive heterogeneous physiological responses. One such physiological response is how PS saturation in the upper leaves of the canopy, which experience lower relative humidity (RH), impacts the lower canopy leaves, which experience lower light conditions but higher RH. Simulations could also enhance understanding of how light transmission properties (reflectance, absorbance, and transmission) impact PS along this gradient.

Further, because the DT continually integrates sensor data and serves as a virtual twin of a physical canopy, it could be used to understand the robustness of canopy performance by investigating a range of canopy scenarios that can be replicated in the field. By iteratively perturbing architectural and physiological parameters of the DT *in silico*, we could perform many simulations in parallel to understand how sensitive canopy performance is to small changes in leaf orientation, architectural, or planting configurations, and how these changes influence PS efficiency.

The creation of canopy-level DTs will leverage the use of whole-plant DTs (Figure 3). Specifically, the framework we envision involves using a parameterized whole-plant DT and then creating multiple instances of it to establish a canopy structure. By using measured data on the physical twin to capture the variation of respective traits (e.g., leaf angle, length, and number), each plant instance could represent a randomized parameterization of the plant DT to reflect the natural variation of plant architecture, thereby creating an *in silico* crop canopy. We therefore argue that a nested FSPM–CGM linked through an AI interface would mechanistically connect plant- and canopy-level physiology [69,70,78,79]. This would be accomplished by first initializing and running the FSPM to estimate plant-level physiological trait parameters (e.g., CO_2 assimilation rate and intercepted light), and then passing these parameters to the CGM to initialize traits such as radiation use efficiency, leaf area, and climate drivers. The CGM would then be run and provide updated growing conditions back to the FSPM, which would, in turn, be run to simulate growth. Data exchange between models could be coordinated through a shared computational environment that supports interoperability among models, data streams, and analytical tools. Emerging AI frameworks and interoperability protocols (e.g., large multimodal environments and model context protocols) provide mechanisms for data formatting, communication,

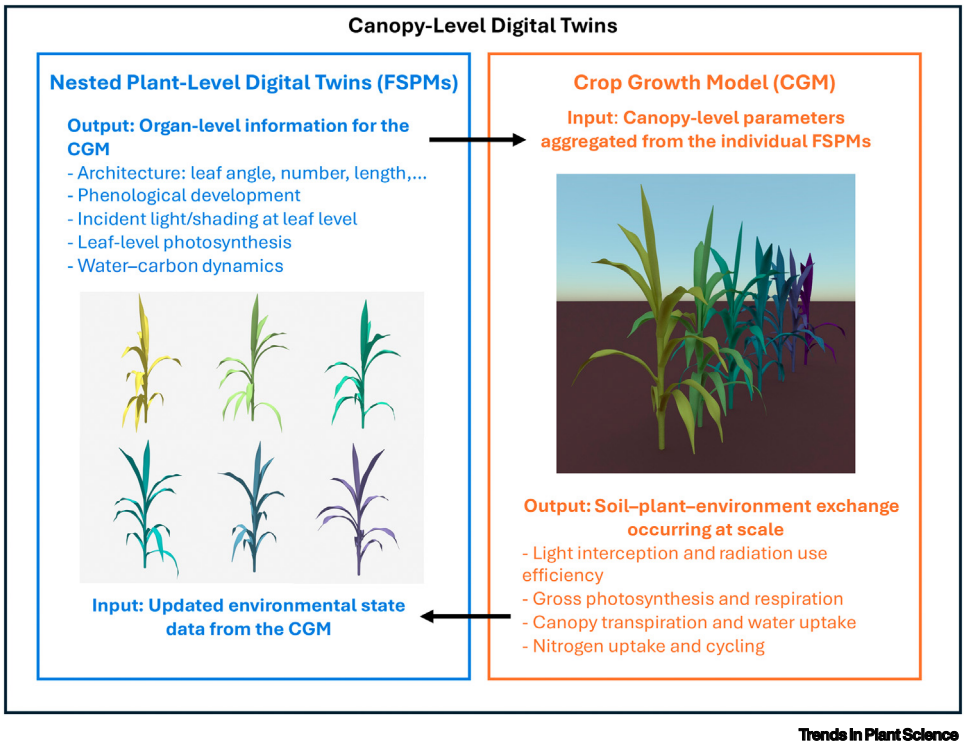

Figure 3. Canopy-level digital twins (DTs) utilizing a nested FSPM–CGM. Canopy level DTs can be constructed using collections of multiple plant-level DTs, each represented by a functional–structural plant model (FSPM), which together form an *in silico* crop canopy [70]. To instantiate the canopy, the individual plant instances are parameterized using phenotypic measurements (e.g., leaf angle, length, and number) to represent natural variation in plant architecture. The FSPMs simulate organ-level processes such as leaf photosynthesis, water–carbon dynamics, phenological development, and light interception. Outputs from the plant-level DTs are aggregated and passed to a crop growth model (CGM), which simulates canopy-scale processes including radiation use efficiency, canopy photosynthesis, transpiration, and nutrient dynamics. The CGM is run and outputs updated environmental and resource conditions that are fed back to the plant-level FSPMs to guide subsequent growth simulations. Data exchange between models occurs through an AI-enabled interface within a large multimodal computing environment, enabling continuous updates from new phenotypic observations from the physical system. This nested FSPM–CGM framework provides a mechanistic pathway for linking plant-level traits and physiology with emergent canopy-level processes.

and integration between FSPMs, CGMs, and phenotyping pipelines [80,81]. Through this FSPM–CGM coupled framework, capturing how plant-level traits impact and drive canopy-level processes could be achieved.

DTs as a framework for closing the G2P gap

Plant DTs can help drive improved crop performance by bridging the G2P gap through connecting genetic variation to dynamic plant functions. This is important because genetic effects on complex agronomic traits are rarely direct. Instead, genetic variation is expressed through developmental, architectural, and physiological processes that interact with environmental and management conditions to produce emergent traits such as water use efficiency, stress resilience, and yield [82,83]. Plant DTs can be used to help connect G2P by extending traditional quantitative genetic approaches, exploring model-derived parameters as functional phenotypes, and using **genomic prediction (GP)** as a means to simulate crop growth and performance in user-defined scenarios.

Plant DTs can transform discrete measurements into a continuous representation of plant development and response to environmental conditions, which has important implications for genetic analyses. A persistent constraint of phenotyping, including high-throughput phenotyping, is that measurements are collected at discrete timepoints. A plant may be imaged daily, weekly, or at key developmental stages, but those data still represent snapshots of dynamic biological processes playing out over time [32,84]. Traits such as flowering time, leaf appearance, organ expansion rate, and stomatal regulation inherently have a temporal component that changes as a function of development and environmental conditions [82], so traditional phenotyping approaches are limited in their ability to fully capture these traits. The DT framework overcomes this limitation by using phenotypic and environmental data collected on the physical twin to initialize and update a DT, enabling it to be studied or trait values extracted at any point in time. DTs can, therefore, provide estimates of unobserved states between measurements and capture transient responses to environmental or stress events. By providing a continuous representation of dynamic traits, DTs could add depth to approaches such as genome-wide association studies and transcriptomic analyses by better linking genetic variation to when, where, and how traits emerge during the season or in response to environmental stimuli such as abiotic stress. Rather than associating loci with single trait values captured by current phenotyping approaches, DTs could enable genetic analyses of trait formation, expression, and plasticity across time.

While these temporally based approaches are useful for studying traditional traits such as leaf area, canopy temperature, and biomass, they often describe outcomes rather than the underlying biological processes. We argue that DT-derived model parameters could serve as dynamic, mechanistic phenotypes that capture genetic variation in plant processes and their responses to environmental or management conditions [85]. This is because model parameters represent biological processes governing growth, development, and response to the environment [86,87]. These model parameters are useful as phenotypes because they can mechanistically describe how plants function and provide a clear link between gene action and plant performance [88]. Model parameters would also enable the genetic dissection of more complex traits because they are biologically interpretable and more genetically stable [89]. An example of the utility of using model parameters as phenotypes relates to the genetic mapping of canopy temperature. Current approaches that use canopy temperature directly as a phenotype may reveal loci associated with the trait, but they provide limited insight into how those loci act mechanistically. By contrast, model-derived parameters from DTs that have integrated canopy temperature data, such as stomatal sensitivity to vapor pressure deficit, could make these associations easier to interpret by linking genetic variation to the physiological processes that regulate canopy

temperature. This would provide a clearer understanding of the functional basis of the trait and the mode of action through which genetic variation influences plant water balance. Through these types of analyses, model parameter phenotypes could provide a missing link between genetic information, functional processes, and phenotypic output, bridging the G2P gap.

A final way that DTs could help address the G2P gap is through integration with GP [90]. In a DT framework, GP could be used to predict the physiological, developmental, and architectural model parameters needed to initialize a genotype-specific DT, rather than directly predicting endpoint traits such as yield [91,92]. The most useful targets for this approach are likely to be component traits with stable genetic control and strong influence on downstream plant performance. Examples would include traits such as flowering time, phenological development, photosynthetic capacity, and resource acquisition traits. This approach would require a training population of DTs, developed on corresponding physical twins, to allow GP models to estimate the relationships between marker data and DT-derived parameters. Once trained, these models could be applied to newly genotyped breeding lines to predict their DT parameters and initialize synthetic genotype-specific functional models. Unlike calibrated DTs developed from a corresponding physical twin, these synthetic functional models would lack a real-world counterpart. Instead, these virtual plants would be simulated and 'phenotyped' across time and in contrasting target environments with various management scenarios. This simulation capability is important because GP of complex agronomic traits, including yield and nitrogen/water use efficiency, often loses accuracy across environments since these traits emerge from interactions among genetic, environmental, and management factors [93]. Rather than asking GP to directly predict these highly context-dependent traits, DT-enabled GP would focus on genetically stable component processes that interact with the environment and management to produce complex agronomic outcomes.

Remaining challenges

While plant DTs offer potential to help accelerate research and breeding efforts, they face challenges to widespread adoption and use. The first challenge is the compatibility and integration of the subcomponent pieces, such as the FSPM, CGM, ray-tracing simulator, and phenomics data pipeline. Enabling these heterogeneous components to communicate seamlessly and operate as a coherent, coupled system remains technically complex. Indeed, integration challenges persist in more mature DT domains such as engineering and medicine, but these obstacles are largely technical rather than conceptual and are therefore not insurmountable. An additional obstacle is that simulating many individual plant DTs in tandem will require substantial computational resources. Moreover, additional dimensions, such as water and nutrient transfer in soils, should be considered. Their simulation will require physics-based engines and discretization of the domain to allow for computational fluid dynamics algorithms to be incorporated.

Beyond technical challenges, logistical barriers also constrain implementation. Limited access to a computationally trained workforce remains pervasive in plant science, alongside constraints in access to the necessary computational hardware and software infrastructure. Moreover, high-quality FSPMs and CGMs may not yet exist for all studied crops, and their development would be a nontrivial task given their complexity and need for validation. Indeed, the overall effort required to deploy a DT remains substantial, encompassing model design, experimental planning, data acquisition, parameterization, validation, and iterative refinement. Developing strategies that reduce this complexity and improve accessibility for plant biologists will be critical for future efforts and is essential for realizing the full potential of plant DTs in research and breeding.

Finally, the broader plant science community would benefit from a collaborative effort to establish a set of standards, open frameworks, and datasets to reduce friction around developing and implementing DTs in novel research contexts. Such standards should include a set of data formats, communication protocols, and interfaces that enable interoperability between models and software frameworks, facilitating the integration of heterogeneous components into a single DT [59,70]. Additionally, standards for model specification and verification would enable comparisons between DTs [59]. With the addition of high-quality, freely accessible data sets, and proactive efforts to ensure that the software powering DTs adheres to FAIR standards (Findable, Accessible, Interoperable, Reusable) [94], both the development of novel DTs and their deployment would be made simpler and more accessible. However, there are many steps, both in research and in deployment, that must be completed first.

Concluding remarks

The development of DTs represents a promising tool in plant science, offering a powerful framework to simulate, analyze, and explore the complexities of plant and crop biology. Critically, the successful development and deployment of DTs hold the promise of bridging the G2P gap by enhancing our understanding of how plant-level traits scale to and impact crop-level performance. As the evolution of DTs continues, through improvements in data integration, modeling, and system development, more applications will become possible as the accuracy of DTs improves. Ideally, success would see the DT become the object of study, with the physical twin largely serving to validate the DT.

However, realizing this potential requires overcoming key challenges, including the integration of the various subcomponents into the DT, validation against empirical data, and making DTs accessible to researchers without extensive computational training (see Outstanding questions). Advancements in AI, sensor technologies, and FSPMs/CGMs are helping to address these gaps, but widespread use and adoption will also depend on building modular, flexible platforms that can support a variety of plant species. Investing in the development and deployment of plant DTs will enable a deeper understanding of the mechanics of plant and crop biology, helping us connect gene function to crop performance under varying environmental conditions (see also Outstanding questions).

Outstanding questions

What is the measure of success for the particular purpose for which the digital twin was designed? How should structural, physiological, and environmental processes be represented to balance mechanistic realism with computational tractability across scales, and where do diminishing returns emerge as model complexity increases to achieve this level of fidelity?

Can plant digital twins capture the full range of phenotypic plasticity and rare genotype-by-environment interactions, or do they represent the 'average plant' and risk missing critical edge-case behaviors? Although phenotypic data on the physical twin can be provided to update the digital twin, are current plant process models capable of representing extreme physiological states or responses to novel environmental conditions within a digital twin framework?

What standards or frameworks will ensure interoperability across different digital twin systems, species, and modeling platforms to enable widespread use? Developing and implementing common vernaculars and data standards will be critical to the adoption of digital twins, as the data and model resources for digital twins are extensive. What is the role of AI in addressing this challenge?

How do we make digital twin technologies accessible to breeders and researchers without computational backgrounds, especially in resource-limited settings? The technical training and experience needed to develop and implement digital twins are extensive, so are there approaches that can be used to reduce this barrier? Is it conceivable, with advancements in generative AI, specifically in code development, that this will no longer be a barrier?

How can recent advances in remote sensing, data infrastructure, and AI be aligned with the increasing computational, storage, and networking demands of digital twins as their scale and precision grow? Can algorithm development match these advancements to extract information from the ever-increasing amount of data and the different modalities that arise?

Acknowledgments

We would like to thank Alyssa Kearly for providing her editorial services to improve the manuscript quality. The authors would like to acknowledge funding support through the National Science Foundation (award numbers 2102120, 2417510, 2417511, 2412928, and 2023310), the U.S. Department of Agriculture (2022-67013-36205), the U.S. Department of Agriculture - National Institute of Food and Agriculture (1032382 and 1032672), and Cotton Incorporated (23-890 and 23-926).

Declaration of interests

The authors declare no competing interests.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT in order to generate illustrative figures to convey the ideas contained in the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

1. Ariesen-Verschuur, N. *et al.* (2022) Digital Twins in greenhouse horticulture: a review. *Comput. Electron. Agric.* 199, 107183
2. Willcox, K. *et al.* (2023) *Foundational Research Gaps and Future Directions for Digital Twins*, National Academies Press
3. Allen, D.B. (2021) Digital twins and living models at NASA. In *Keynote Presentation at the ASME Digital Twin Summit*
4. Douthwaite, J.A. *et al.* (2021) A modular digital twinning framework for safety assurance of collaborative robotics. *Front. Robot. AI* 8, 758099
5. Committee on Foundational Research Gaps and Future Directions for Digital Twins, Board on Mathematical Sciences and Analytics, Committee on Applied and Theoretical Statistics,

Computer Science and Telecommunications Board, Board on Life Sciences, Board on Atmospheric Sciences and Climate *et al.* (2024) *Foundational Research Gaps and Future Directions for Digital Twins*, National Academies Press
6. Grieves, M.W. (2003) *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*, White paper
7. Aivaliotis, P. *et al.* (2019) The use of Digital Twin for predictive maintenance in manufacturing. *Int. J. Comput. Integr. Manuf.* 32, 1067–1080
8. Cimino, C. *et al.* (2019) Review of digital twin applications in manufacturing. *Comput. Ind.* 113, 103130
9. Huang, Z. *et al.* (2021) A survey on AI-driven digital twins in Industry 4.0: smart manufacturing and advanced robotics. *Sensors (Basel)* 21, 6340
10. Katsoulakis, E. *et al.* (2024) Digital twins for health: a scoping review. *NPJ Digit. Med.* 7, 77
11. Chen, J. *et al.* (2024) Generative-AI-driven human digital twin in IoT healthcare: a comprehensive survey. *IEEE Internet Things J.* 11, 34749–34773
12. Tang, C. *et al.* (2024) A roadmap for the development of human body digital twins. *Nat. Rev. Electr. Eng.* 1, 199–207
13. Laubenbacher, R. *et al.* (2024) Digital twins in medicine. *Nat. Comput. Sci.* 4, 184–191
14. Shahat, E. *et al.* (2021) City digital twin potentials: a review and research agenda. *Sustainability* 13, 3386
15. Peldon, D. *et al.* (2024) Navigating urban complexity: the transformative role of digital twins in smart city development. *Sustain. Cities Soc.* 111, 105583
16. Liu, Z. *et al.* (2024) Plant digital twins based on model phenotyping and functional-structural plant modeling. In *Proceedings of the 2024 IEEE 4th International Conference on Digital Twins and Parallel Intelligence (DTPI)*, pp. 211–214, IEEE
17. Escribà-Gelonch, M. *et al.* (2024) Digital Twins in agriculture: orchestration and applications. *J. Agric. Food Chem.* 72, 10737–10752
18. Grieves, M. and Vickers, J. (2017) Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary Perspectives on Complex Systems* (Kahlen, J. K. *et al.*, eds), pp. 85–113, Springer International Publishing
19. Caldarelli, G. *et al.* (2023) The role of complexity for digital twins of cities. *Nat. Comput. Sci.* 3, 374–381
20. Anderson, P.W. (1972) More is different. *Science* 177, 393–396
21. Sharma, A. *et al.* (2022) Digital Twins: state of the art theory and practice, challenges, and open research questions. *J. Ind. Inf. Integr.* 30, 100383
22. Yao, J.-F. *et al.* (2023) Systematic review of digital twin technology and applications. *Vis. Comput. Ind. Biomed. Art* 6, 10
23. Gonsard, A. *et al.* (2024) Digital twins for chronic lung diseases. *Eur. Respir. Rev.* 33, 240159
24. Baillargeon, B. *et al.* (2014) The Living Heart Project: a robust and integrative simulator for human heart function. *Eur. J. Mech. A Solids* 48, 38–47
25. Qian, S. *et al.* (2025) Developing cardiac digital twin populations powered by machine learning provides electrophysiological insights in conduction and repolarization. *Nat. Cardiovasc. Res.* 4, 624–636
26. Tardini, E. *et al.* (2022) Optimal treatment selection in sequential systemic and locoregional therapy of oropharyngeal squamous carcinomas: deep Q-learning with a patient-physician digital twin dyad. *J. Med. Internet Res.* 24, e29455
27. Surian, N.U. *et al.* (2024) A digital twin model incorporating generalized metabolic fluxes to identify and predict chronic kidney disease in type 2 diabetes mellitus. *NPJ Digit. Med.* 7, 140
28. Kovatchev, B.P. *et al.* (2025) Human-machine co-adaptation to automated insulin delivery: a randomised clinical trial using digital twin technology. *NPJ Digit. Med.* 8, 253
29. Xiong, H. *et al.* (2023) The Digital Twin Brain: a bridge between biological and artificial intelligence. *Intell. Comput.* 2, 0055
30. Zhang, Y. *et al.* (2023) Predicting ventricular tachycardia circuits in patients with arrhythmogenic right ventricular cardiomyopathy using genotype-specific heart digital twins. *Elife* 12, RP88865
31. Hansen, J. *et al.* (2024) From transcriptomics to digital twins of organ function. *Front. Cell Dev. Biol.* 12, 1240384
32. Pauli, D. *et al.* (2016) The quest for understanding phenotypic variation via integrated approaches in the field environment. *Plant Physiol.* 172, 622–634
33. Furbank, R.T. and Tester, M. (2011) Phenomics – technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644
34. Lowder, S.K. *et al.* (2016) The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Dev.* 87, 16–29
35. Rada, N.E. and Fuglie, K.O. (2019) New perspectives on farm size and productivity. *Food Policy* 84, 147–152
36. Scharff, L.B. *et al.* (2022) Prospects to improve the nutritional quality of crops. *Food Energy Secur.* 11, e327
37. Burgess, A.J. *et al.* (2023) Improving crop yield potential: underlying biological processes and future prospects. *Food Energy Secur.* 12, e435
38. Ortiz-Bobea, A. *et al.* (2021) Anthropogenic climate change has slowed global agricultural productivity growth. *Nat. Clim. Chang.* 11, 306–312
39. Schlenker, W. and Roberts, M.J. (2009) Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proc. Natl. Acad. Sci. U. S. A.* 106, 15594–15598
40. Yuan, X. *et al.* (2023) A global transition to flash droughts under climate change. *Science* 380, 187–191
41. Snapp, S. *et al.* (2023) Spatially differentiated nitrogen supply is key in a global food–fertilizer price crisis. *Nat. Sustainability* 6, 1268–1278
42. Ray, D.K. *et al.* (2013) Yield trends are insufficient to double global crop production by 2050. *PLoS One* 8, e66428
43. Ray, D.K. *et al.* (2012) Recent patterns of crop yield growth and stagnation. *Nat. Commun.* 3, 1293
44. Li, H. *et al.* (2018) Fast-forwarding genetic gain. *Trends Plant Sci.* 23, 184–186
45. Atlin, G.N. *et al.* (2017) Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Glob. Food Sec.* 12, 31–37
46. Xiong, W. *et al.* (2022) Climate change challenges plant breeding. *Curr. Opin. Plant Biol.* 70, 102308
47. Gill, T. *et al.* (2022) A comprehensive review of high throughput phenotyping and machine learning for plant stress phenotyping. *Phenomics* 2, 156–183
48. Visakh, R.L. *et al.* (2024) Precision phenotyping in crop science: from plant traits to gene discovery for climate-smart agriculture. *Plant Breed.* Published online October 20, 2024. https://doi.org/10.1111/pbr.13228
49. Howard, D.A. *et al.* (2022) A case study of digital twin for greenhouse horticulture production flow. In *Proceedings of the 2022 IEEE 2nd International Conference on Digital Twins and Parallel Intelligence (DTPI)*, pp. 1–6, IEEE
50. Skobelev, P. *et al.* (2024) Multiparametric variety models in digital twin of plant for winter wheat. In *Proceedings of the 2024 X International Conference on Information Technology and Nanotechnology (ITNT)*, pp. 1–9, IEEE
51. Smoleňová, K. *et al.* (2025) Development of a tomato functional–structural plant model for digital twin applications. *In Silico Plants* 7, diaf022
52. Fourcaud, T. *et al.* (2008) Plant growth modelling and applications: the increasing importance of plant architecture in growth models. *Ann. Bot.* 101, 1053–1063
53. Monteith, J.L. (1965) Light distribution and photosynthesis in field crops. *Ann. Bot.* 29, 17–37
54. DeJong, T.M. *et al.* (2011) Using functional–structural plant models to study, understand and integrate plant development and ecophysiology. *Ann. Bot.* 108, 987–989
55. Boote, K.J. *et al.* (1996) Potential uses and limitations of crop models. *Agron. J.* 88, 704–716
56. Holzworth, D.P. *et al.* (2014) APSIM – evolution towards a new generation of agricultural systems simulation. *Environ. Model. Softw.* 62, 327–350
57. Prusinkiewicz, P. and Lindenmayer, A. (1996) *The Algorithmic Beauty of Plants*, Springer, The Virtual Laboratory, p. 228
58. Vos, J. *et al.* (2010) Functional–structural plant modelling: a new versatile tool in crop science. *J. Exp. Bot.* 61, 2101–2115
59. Louarn, G. and Song, Y. (2020) Two decades of functional–structural plant modelling: now addressing fundamental

questions in systems biology and predictive ecology. *Ann. Bot.* 126, 501–509
60. Zarei, A. *et al.* (2024) PlantSegNet: 3D point cloud instance segmentation of nearby plant organs with identical semantics. *Comput. Electron. Agric.* 221, 108922
61. Leinonen, I. *et al.* (2006) Estimating stomatal conductance with thermal imagery. *Plant Cell Environ.* 29, 1508–1518
62. Serbin, S.P. *et al.* (2012) Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *J. Exp. Bot.* 63, 489–502
63. Meroni, M. *et al.* (2009) Remote sensing of solar-induced chlorophyll fluorescence: review of methods and applications. *Remote Sens. Environ.* 113, 2037–2051
64. Ostermann, I. *et al.* (2024) Sorghum segmentation and leaf counting using in silico trained deep neural model. *Plant Phenome J.* 7, e70002
65. Davis, J.M. *et al.* (2025) 3D reconstruction enables high-throughput phenotyping and quantitative genetic analysis of phyllotaxy. *Plant Phenomics* 7, 100023
66. Xue, J. and Su, B. (2017) Significant remote sensing vegetation indices: a review of developments and applications. *J. Sens.* 2017, 1353691
67. Sel, K. *et al.* (2024) Building digital twins for cardiovascular health: from principles to clinical impact. *J. Am. Heart Assoc.* 13, e031981
68. Gonzalez, E.M. *et al.* (2023) PhytoOracle: scalable, modular phenomics data processing pipelines. *Front. Plant Sci.* 14, 1112973
69. Wang, X. *et al.* (2024) Functional–structural plant model "GreenLab": a state-of-the-art review. *Plant Phenomics* 6, 0118
70. Soualiou, S. *et al.* (2021) Functional–structural plant models mission in advancing crop science: opportunities and prospects. *Front. Plant Sci.* 12, 747142
71. Ghanem, M.E. *et al.* (2015) Physiological phenotyping of plants for crop improvement. *Trends Plant Sci.* 20, 139–144
72. Ort, D.R. *et al.* (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8529–8536
73. Long, S.P. *et al.* (2006) Can improvement in photosynthesis increase crop yields? *Plant Cell Environ.* 29, 315–330
74. Slattery, R.A. and Ort, D.R. (2021) Perspectives on improving light distribution and light use efficiency in crop canopies. *Plant Physiol.* 185, 34–48
75. Croce, R. *et al.* (2024) Perspectives on improving photosynthesis to increase crop yield. *Plant Cell* 36, 3944–3973
76. Vadez, V. *et al.* (2024) Crop traits and production under drought. *Nat. Rev. Earth Environ.* 5, 211–225
77. Boote, K.J. *et al.* (2013) Putting mechanisms into crop production models. *Plant Cell Environ.* 36, 1658–1672
78. Long, Q. *et al.* (2018) An architecture for the integration of different functional and structural plant models. In *Proceedings of the 7th International Conference on Informatics, Environment, Energy and Applications*, pp. 107–113, ACM
79. Evers, J.B. *et al.* (2019) Understanding and optimizing species mixtures using functional–structural plant modelling. *J. Exp. Bot.* 70, 2381–2388
80. Huang, D. *et al.* (2024) From Large Language Models to Large Multimodal Models: a literature review. *Appl. Sci. (Basel)* 14, 5068
81. Hou, X. *et al.* (2026) Model Context Protocol (MCP): landscape, security threats, and future research directions. *ACM Trans. Softw. Eng. Methodol.* Published online February 16, 2026. https://doi.org/10.1145/3796519
82. Hammer, G. *et al.* (2006) Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11, 587–593
83. Hammer, G.L. *et al.* (2010) Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. *J. Exp. Bot.* 61, 2185–2202
84. Pérez-Valencia, D.M. *et al.* (2022) A two-stage approach for the spatio-temporal analysis of high-throughput phenotyping data. *Sci. Rep.* 12, 3177
85. Tardieu, F. (2003) Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends Plant Sci.* 8, 9–14
86. Ben-Haj-Salah, H. and Tardieu, F. (1995) Temperature affects expansion rate of maize leaves without change in spatial distribution of cell length (analysis of the coordination between cell division and cell expansion). *Plant Physiol.* 109, 861–870
87. Granier, C. and Tardieu, F. (1998) Is thermal time adequate for expressing the effects of temperature on sunflower leaf development? *Plant Cell Environ.* 21, 695–703
88. Tardieu, F. *et al.* (2005) Linking physiological and genetic analyses of the control of leaf growth under changing environmental conditions. *Aust. J. Agric. Res.* 56, 937–946
89. Reymond, M. *et al.* (2003) Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol.* 131, 664–675
90. Meuwissen, T.H. *et al.* (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829
91. Technow, F. *et al.* (2015) Integrating crop growth models with whole genome prediction through approximate Bayesian computation. *PLoS One* 10, e0130855
92. Cooper, M. *et al.* (2016) Use of crop growth models with whole-genome prediction: application to a maize multienvironment trial. *Crop Sci.* 56, 2141–2156
93. Crossa, J. *et al.* (2017) Genomic selection in plant breeding: methods, models, and perspectives. *Trends Plant Sci.* 22, 961–975
94. Zeng, Y. *et al.* (2025) Towards an open soil-plant digital twin based on STEMMUS-SCOPE model following open science. *Comput. Geosci.* 205, 106013