PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

A digital-twin framework for genomic-based optimization of an agrophotovoltaic greenhouse system

R. S. Isied, E. Mengi and T. I. Zohdi

Department of Mechanical Engineering, University of California, Berkeley, 6102 Etcheverry Hall, Berkeley 94720-1740, CA, USA

RSI, 0000-0003-0095-8211; EM, 0000-0002-5759-3033; TIZ, 0000-0002-0844-3573

Research

Check for updates

Cite this article: Isied RS, Mengi E, Zohdi TI. 2022 A digital-twin framework for genomic-based optimization of an agrophotovoltaic greenhouse system. *Proc. R. Soc. A* **478**: 20220414.
https://doi.org/10.1098/rspa.2022.0414

Received: 14 June 2022
Accepted: 25 October 2022

Subject Areas:
computer modelling and simulation, computational mechanics, power and energy systems

Keywords:
agrophotovoltaics, greenhouses, optimization, simulation

Author for correspondence:
R. S. Isied
e-mail: ruisied@berkeley.edu

Agrophotovoltaic systems combine solar energy generation with agricultural production. In this work, a computational framework is developed to trace light rays through agrophotovoltaic greenhouses, in order to calculate the power generated by greenhouse solar cells, as well as power absorbed by crops within the greenhouse. A geometric ray-tracing algorithm is developed to track the propagation, reflection and refraction of light interacting with a translucent greenhouse. Genomic-based optimization techniques are used to meet a target greenhouse power generation level, as well as a targeted photosynthetic power absorption by optimizing the geometry, translucency and material characteristics of the greenhouse. Representative numerical examples are provided. The framework can be used to generate tailored, temporal and location-specific greenhouse designs.

1. Introduction

Agriculturally viable land has been the target of renewable energy production efforts, such as solar panels and wind turbines. The competition between energy production and agricultural production has led to restrictions on non-agricultural activity on agricultural land in California [1] while pushing for sustainable agriculture and renewable energy production to reach the state's carbon-neutrality goals [2]. Alternative solutions are needed to alleviate such problems. One possible way is by employing systems which combine solar energy generation with agricultural production, so-called agrophotovoltaic (APV) systems

THE ROYAL SOCIETY
PUBLISHING

Figure 1. Example agrophotovoltaic greenhouse. Photo above from public domain: https://pixabay.com. (Online version in colour.)

[3–5]. While there are many examples of APVs installed in open outdoor fields, a similar set-up could be used for greenhouses, which offer more precise climate control of plant growth. Carbon-neutral APV greenhouses, combined with vertical farming, can reduce agricultural land shortage problems and increase crop yields, regardless of the weather/season. These greenhouses also have the potential to be used in prospective space-exploration applications, where climate-control and efficient use of habitable space is crucial.

Several experimental studies have been conducted on the feasibility of APV greenhouse systems [6–10]. While experimental set-ups can provide insight into the best configuration for an APV greenhouse where plant growth and energy production are maximized, physically testing all possible configurations is time-consuming and requires substantial financial costs. In order to alleviate these problems, this work considers the digital-twin (a digital replica of the system) approach whereby various greenhouse designs are computationally generated, and light rays are tracked through a representative domain. As an example, the absorption and reflection of solar rays hitting three-dimensional thin film panels wrapped around the greenhouse is simulated (figure 1). This simulation technology is combined with genomic-based algorithms in order to ascertain system parameters to optimize the greenhouse response. While the APV studies previously mentioned have either experimentally or computationally tested APV designs to find the best design, the proposed framework in figure 2 will enable the user to quickly evaluate solar greenhouse designs with the reduced order model and find the 'optimal' configuration within the design space using the genomic optimization framework. This would reduce the experimental testing of greenhouse designs to the top designs found by the genomic-based optimizer.

In this study, the system parameters are greenhouse shape, solar panel translucency and solar panel refractive index. It is important to note that the complexity in determining the optimal translucency of the solar panels poses a significant challenge. La Notte *et al.* [11] present novel translucent PV-cell technologies that use organic solar cells, dye-sensitized solar cells and perovskite solar cells. These approaches use the wavelength sensitive nature of photosynthesis, which favours radiation in the 400–700 nm range, called photosynthetically active radiation (PAR). These special solar cells absorb light outside of the PAR and transmit PAR to plants, where it is absorbed by the chlorophyll. Wavelength-specific transmittance of the solar panels can be part of the optimization algorithm to determine the optimal solar-cell technology to be used, as well as to aid the design of new solar-cells with specific transmittance parameters on a per-plant basis.

The system presents many challenges with regard to optimizing solar farm characteristics to ensure energy and agricultural needs are simultaneously met. The goal would be to balance

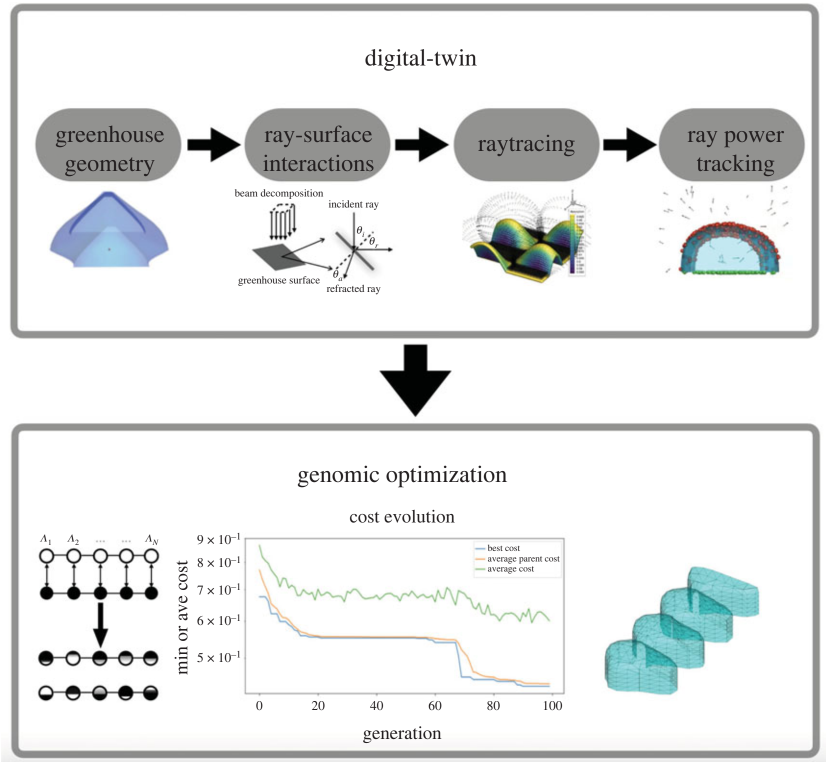

Figure 2. Digital-twin and genomic optimization framework. (Online version in colour.)

plant growth and energy generation in a way that the greenhouse is self-sufficient and is carbon neutral. Accordingly, in this work, a computational framework is developed to trace light rays through APV greenhouses, in order to calculate the power generated by greenhouse solar cells, as well as power absorbed by crops within the greenhouse. A geometric ray-tracing algorithm is developed to track the propagation, reflection and refraction of light interacting with a translucent greenhouse. Genomic-based optimization techniques are used to meet a target greenhouse power generation level, as well as a targeted photosynthetic power absorption by optimizing the geometry, translucency and material characteristics of the greenhouse. Representative numerical examples are provided. The framework can be used to generate tailored, temporal and location-specific and greenhouse designs.

There are high-fidelity light and plant modelling tools, such as HELIOS and Raytrace3D, which would most accurately simulate the energy distribution in the agro-solar environments and provide useful insight about the power generation and crop production while needing extensive libraries and relatively more computational power. In this study, a reduced-order digital-twin is proposed that allows the user to simulate a day of irradiation in seconds, which allows for optimization of system parameters on the order of minutes. Thus, the simulated APV greenhouse can be regarded as a digital twin that can be simulated and optimized in real time. It is noted that the optimized designs outputted by the genomic optimization algorithm can then be fed into higher fidelity models to evaluate the best designs further. This virtual set-up allows the user to reduce the time and capital spent on the experiments to develop a carbon-neutral greenhouse APV system. We aim to create a digital replica of a solar greenhouse to optimize land use and energy generation by calculating the ground and solar panel power absorption due to solar light-scattering within the system. This approach has the potential to deliver a tailored, situation-specific and self-sufficient APV greenhouse system.

2. Physical model and system optimization

(a) Creating solar panel geometries

Surface functions are assumed to be known for the roof and side walls of the greenhouse $F(x_1, x_2, x_3)$. The ground is assumed to be flat at a constant height $x_3 = 0$. We check for light interactions between the greenhouse or ground by checking if $F(x_{1,j}, x_{2,j}, x_{3,j}) \leq 1$ for the greenhouse surface or if $x_{3,j} \leq 0$ for the ground. The ray position at the time of surface impact is used to compute the absorptivity and reflectivity of that beam.

We can use a generalized equation for an ellipsoid appended with a generalized sinusoid equation to represent a broad range of greenhouse geometries centred at the origin

$$F(x_1, x_2, x_3) = \left|\frac{x_1}{R_1}\right|^{p_1} + \left|\frac{x_2}{R_2}\right|^{p_2} + \left|\frac{x_3}{R_3}\right|^{p_3} + a_1 \sin(\omega_1 x_1) + a_2 \sin(\omega_2 x_2) \leq 1 \tag{2.1}$$

where (R_1, R_2, R_3) are the generalized radii, (p_1, p_2, p_3) are the exponents of the generalized ellipsoid, (a_1, a_2) are the amplitudes of the sinusoid, and (ω_1, ω_2) are the associated frequencies of the sinusoid component. These scaling factors will be the design parameters used to optimize the shape of our greenhouse. It is assumed that the solar panels will be form-fitted to the shape of the greenhouse. The resulting topology of the greenhouse can range from simple geometrical shapes to complex sinusoidal shapes, as visualized in figure 3. The generalized radii and exponents control the ellipsoidal features of the greenhouse while the amplitudes and frequencies control the sinusoidal features. The contour nature of the greenhouse equation enforces an infinitely long greenhouse along its length. For visual purposes, the domain of the greenhouse has been constrained by the domain of simulated light rays. In practice, one can extend the cyclical greenhouse design according to land available.

(b) Reflection and absorption of rays

We assume the rays travel through a vacuum and thus we can use the nominal speed of light ($3 \times 10^8\,\mathrm{m\,s^{-1}}$). The design string parameters that are refined by the genomic-based optimizer can be selected by the user. The total power per surface area is given by P_{tot} which is evenly distributed among the rays based on the total area of light cover being considered A_b. The ray positions are generated randomly over a square region with a side length of $2s_{\mathrm{Reg}}$. The centre of the region is defined to follow the sun's trajectory during the day, meaning that each run of the simulation will base the square beam at a point where the beam is calculated to hit the greenhouse at that specific beam angle.

With these parameters, we can define the power per ray in a light pulse as follows:

$$P_r = \frac{P_{\mathrm{tot}} A_b}{N_r}. \tag{2.2}$$

We follow a standard euclidean basis $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ indicating horizontal, vertical and in-depth directions. To obtain the angle of incidence of the ray of light, θ_i, we first compute the inward unit surface normal vector $\boldsymbol{n}$ of the greenhouse surface F given by

$$\boldsymbol{n} = \frac{-\nabla F}{||\nabla F||}, \tag{2.3}$$

where

$$\nabla F = \frac{\partial F}{\partial x_1}\boldsymbol{e_1} + \frac{\partial F}{\partial x_2}\boldsymbol{e_2} + \frac{\partial F}{\partial x_3}\boldsymbol{e_3} \tag{2.4}$$

is the gradient of the greenhouse surface equation. Note that for the flat ground, the inward normal vector is constant and defined by $\boldsymbol{n}_g = [0, 0, -1]$. Next, we can then compute the angle

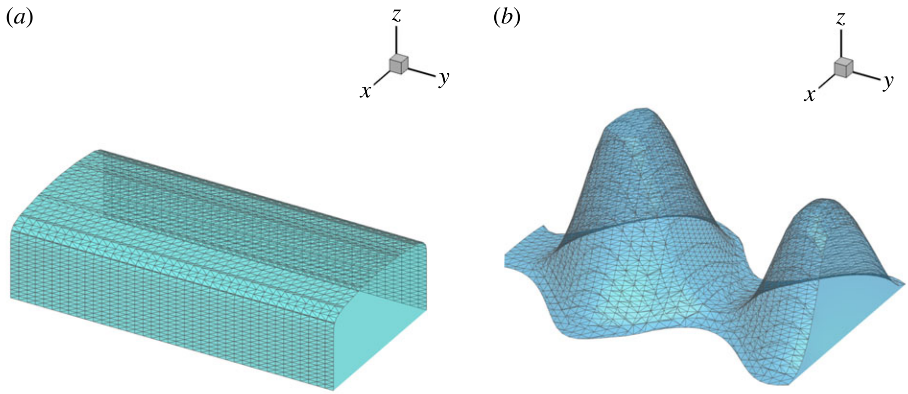

Figure 3. Example greenhouse shapes for (*a*) $(R_1, R_2, R_3) = (0.5, 5, 0.5)$, $(p_1, p_2, p_3) = (20, 20, 20)$, $(a_1, a_2) = (0, 0)$, $(\omega_1, \omega_2) = (0.1, 0.1)$ and (*b*) $(R_1, R_2, R_3) = (0.5, 5, 0.5)$, $(p_1, p_2, p_3) = (0.5, 20, 0.5)$, $(a_1, a_2) = (0.5, 0.5)$, $(\omega_1, \omega_2) = (4, 7)$. (Online version in colour.)

of incidence (θ_i) via the cosine formula between the ray velocity vector ($\boldsymbol{v}$) and the inward unit normal vector of the solar panel surface ($\boldsymbol{n}$)

$$\theta_i = \cos^{-1}\left(\frac{\boldsymbol{v}_j \cdot \boldsymbol{n}_j}{||\boldsymbol{v}_j||||\boldsymbol{n}_j||}\right). \tag{2.5}$$

The component of the ray velocity normal to the surface of the solar panel is given by

$$\boldsymbol{v}_{j,\perp} = ||\boldsymbol{v}_j|| \cos\theta_i \boldsymbol{n}_j. \tag{2.6}$$

We can calculate the outgoing reflected velocity ($\boldsymbol{v}_j^{\mathrm{ref}}$) by turning the inbound normal velocity outward by subtracting it twice

$$\boldsymbol{v}_j^{\mathrm{ref}} = \boldsymbol{v}_j - 2\boldsymbol{v}_{j,\perp}. \tag{2.7}$$

Next, we consider the material properties of the solar panel. We define $\hat{n}$ as the ratio of the refractive indices of the ambient (incident) medium (n_i) and absorbing medium (n_a) such that

$$\hat{n} = \frac{n_a}{n_i}. \tag{2.8}$$

The absorbing medium refractive index, n_a (solar panel), is to be user-designed based on the optimization model. We assume the incident refractive index to be that of a vacuum as $n_i = 1$. Figure 4 outlines the decomposition of an individual ray on the greenhouse surface.

With the above parameters defined, the refractive angle of incoming light can be obtained using Snell's Law, namely

$$\theta_r = \sin^{-1}\left(\frac{1}{\hat{n}}\sin(\theta_i)\right). \tag{2.9}$$

This angle is used for the light rays travelling into and out of the greenhouse. It is assumed that once refracting through the solar panel, the inner medium of the greenhouse is the same as that outside the greenhouse. It is also assumed that the solar panel is sufficiently thin such that the change in curvature between the outer and inner surface is negligible. For computational efficiency, the refracted light rays were not tracked through time within the surface of the solar panel. Rather, refracted rays were manually translated through the medium according to figure 5.

In this figure example, the light travels from the ambient environment, through the solar panel, and into the greenhouse. We define t as the thickness of the solar panel, θ_r as the refractive angle and d_{light} as the translational distance the light ray travels before exiting the solar panel. Since it is assumed that the medium inside and outside of the greenhouse is the same, the light retains its

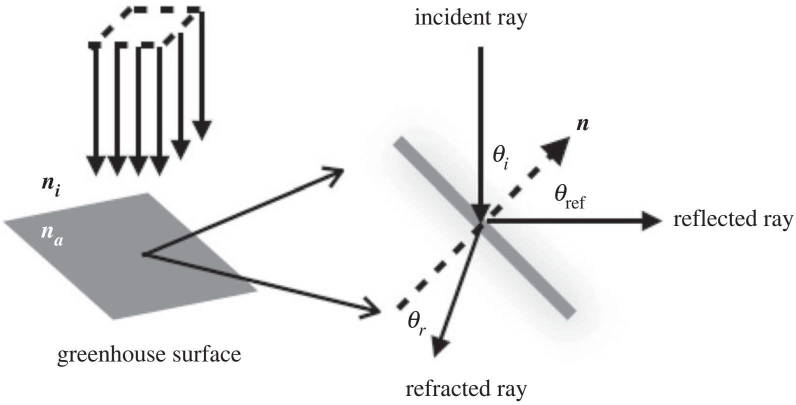

Figure 4. Beam decomposition for a geometric ray-tracing model.

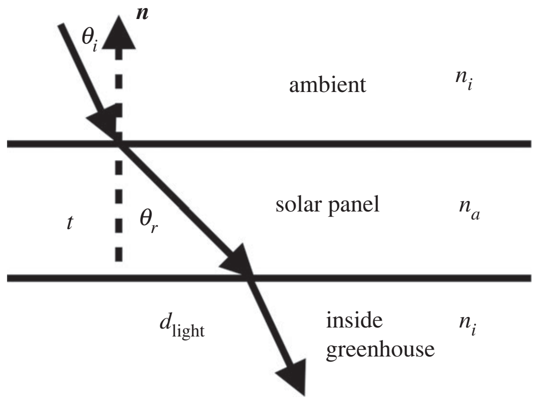

Figure 5. Evolution of light refraction through wall of greenhouse.

initial incidence angle θ_i once it enters the greenhouse. Accordingly, the translational distance can be calculated as follows:

$$d_{\text{light}} = t\tan(\theta_r). \tag{2.10}$$

In practice, each light ray is manually translated according to equation (2.10) when it comes in contact with the solar panel surface for both incoming and outgoing rays.

(c) Power tracking and splitting

We consider a ray of light incident upon a material interface which produces a reflected ray and a transmitted/absorbed (refracted) ray. The ratio of reflected electromagnetic power (I_r) to the total incident electromagnetic power (I_i) defines the total reflectance $\mathcal{IR} \equiv (I_r/I_i)$, where $0 \le \mathcal{IR} \le 1$ for unpolarized electromagnetic radiation. We refer the reader to Zohdi [12] for a detailed derivation of $\mathcal{IR}$. The reflectance is a function of the angle of incidence of the incoming rays, the medium which the rays travel through, and the material which the rays intersect with. For this model, we will consider applications with non-magnetic media and frequencies where the magnetic

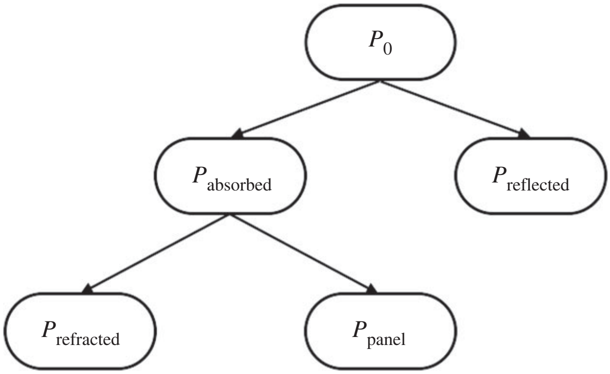

Figure 6. Power distribution on solar panels.

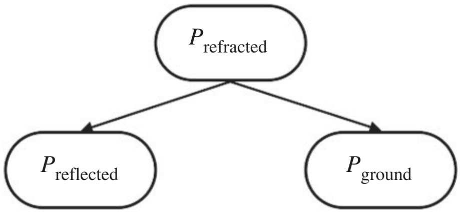

Figure 7. Power distribution on the ground.

permeability is virtually the same for both the incident and absorbing medium. Following Zohdi [12], we define the reflectivity $\mathcal{IR}$ as follows:

$$\mathcal{IR}(\hat{n},\theta_i) = \frac{I_r}{I_i} = \frac{1}{2}\left(\frac{\hat{n}^2\cos\theta_i - (\hat{n}^2 - \sin^2\theta_i)^{1/2}}{\hat{n}^2\cos\theta_i + (\hat{n}^2 - \sin^2\theta_i)^{1/2}}\right)^2 + \frac{1}{2}\left(\frac{\cos\theta_i - (\hat{n}^2 - \sin^2\theta_i)^{1/2}}{\cos\theta_i + (\hat{n}^2 - \sin^2\theta_i)^{1/2}}\right)^2. \tag{2.11}$$

The reflectance is used to obtain the total amount of absorbed power by a material as follows:

$$P_{\text{abs}} = (1 - \mathcal{IR})P_r. \tag{2.12}$$

We track the total power and position of a ray and stop tracking it if either the ray has moved outside of the user-defined domain, or the ray's power is reduced below a user-defined threshold.

The power absorbed by the ground and solar panel surfaces are obtained using the following flowcharts, shown in figures 6 and 7 during reflectance/absorptivity calculations that occur within the ray-tracing algorithm.

The incident light rays which come in contact with the solar panel are split into reflected, refracted and power-converted light rays. The light that refracts into the greenhouse is based on the transmissibility of the solar panel defined by $\gamma \in [0, 1]$. The refracted power $P_{\text{refracted}}$ is set to be proportional with the transmissibility, namely

$$P_{\text{refracted}} = \gamma P_{\text{absorbed}}, \tag{2.13}$$

whereas the used light contains the remaining power, namely

$$P_{\text{panel}} = (1 - \gamma)P_{\text{absorbed}}. \tag{2.14}$$

Similarly, the power distribution of the refracted ray when it hits the ground surface is determined by the reflectance parameter.

While the power conversion efficiencies of the solar panels and the plants are dependent on ambient parameters, this study only takes the surface power absorption into account for simplicity.

(d) Time-stepping algorithm

We use the power propagation and ray-tracing algorithm described above in conjunction with an explicit time-stepping scheme (forward Euler method) to track the rays from time $t = 0$ to $t = t_{\text{final}}$ or until all rays are deactivated, whichever comes first. The time-stepping algorithm is as follows for all light rays $j = 1, \ldots, N_r$:

1. Initialize ray positions $\boldsymbol{r}_j(t = 0)$ and velocities $\boldsymbol{v}_j(t = 0)$
2. Iterate ray positions in time using

$$\boldsymbol{r}_j(t + \Delta t) = \boldsymbol{r}_j(t) + \Delta t \boldsymbol{v}_j(t). \tag{2.15}$$

3. Check for surface-ray collisions. If ray has collided with a surface
   - (a) Update power absorbed by the surfaces ($P_{\text{abs}} = (1 - \mathcal{IR})P_r$) and remaining power for all rays ($P_{\text{ref}_{,j}} = \mathcal{IR}_j P_{r,j}$),
   - (b) Calculate new ray velocity values after reflection $\boldsymbol{v}_j^{\text{ref}}$,
   - (c) Calculate ray refractions if ray hits the greenhouse surface,
4. Check for active rays:
   - (a) **If no active rays are remaining:** End simulation,
   - (b) **Otherwise:** Move to the next step,
5. Increment the time step to ($t = t + \Delta t$) and go back to Step 2.

The time-step size Δt is chosen accordingly to capture all ray surface interactions using the formula $\Delta t = \xi(h_0^{\text{ray}}/c)$, where h_0^{ray} is the initial height of the generated rays, c is the speed of light, and ξ is a tunable parameter such that $\xi \in (0, 1]$. The high velocity of the light rays requires a time step size scaled to accurately observe the motion of the rays with a sufficient number of time steps. For this study, the parameter was chosen to be $\xi = 0.01$. This parameter was chosen to be sufficiently small to capture all ray interactions, while not causing a significant bottleneck for simulation time. Further refinement of this parameter can be obtained by conducting a convergence study in which an average F is calculated for all surface ray interactions as a function of ξ.

(e) Process optimization

The 'design string' for this process contains all eight controllable constants

$$\boldsymbol{\Lambda}^i \equiv \{\Lambda_1^i, \ldots, \Lambda_N^i\} \equiv \{\gamma, \hat{n}_s, p_1, p_3, a_1, a_2, \omega_1, \omega_2\}, \tag{2.16}$$

where γ is the transmissibility of the solar panels and $\hat{n}_s$ is the refractive index of the solar panels. These parameters can be controlled within user-specified bounds. All system parameters used in this simulation are displayed within table 2.

A good set of characteristic parameters of our greenhouse design will allow for user-defined absorptivity of the solar panels (for energy production) as well as absorbed light by the ground (for plant growth). These two opposing objectives force the optimizer to design a solution which balances between the two goals. It should be noted that this model does not penalize for over-irradiation/heating of crops in a time period throughout the day given that this framework is developed independently of specific crop characteristics. An application of this framework to a specific set of crops should consider this in choosing $P_{\text{plant,des}}$ and add a penalty term to the cost function which drastically increases the crops if the photosynthetic power exceeds a power threshold detrimental to the crops' growth.

With these objectives in mind, we construct the following cost function,

$$\Pi = w_1 \left| \frac{P_{\text{solar,des}} - P_{\text{elec}}}{P_{\text{solar,des}}} \right| + w_2 \left| \frac{P_{\text{plant,des}} - P_{\text{photosynth}}}{P_{\text{plant,des}}} \right|, \tag{2.17}$$

where the weights are chosen to be $w_1 = 2$ and $w_2 = 1$. The weights can be assigned arbitrarily by the user as long as they are positive and reflect the relative importance of each objective term. The weights in this work's numerical example are chosen to prioritize designs in which the electrical power is as close to the desired solar power as possible. Each term in equation (2.17) attempts to set power absorbed by the solar panels and power absorbed for plant growth as close as possible to their desired counterparts, $P_{\text{solar,des}}$ and $P_{\text{plant,des}}$. Note that all terms in the cost function are non-dimensional.

(f) Process optimization scheme: genetic algorithm

The greenhouse digital twin is optimized using a genetic algorithm. All genetic algorithm parameters and search bounds chosen are displayed in table 2. For applying a genetic algorithm, the algorithm is as follows:

1. Generate S random genetic strings, where $\boldsymbol{\Lambda}_i \in [\Lambda_i^-, \Lambda_i^+]$

$$\boldsymbol{\Lambda} = (\boldsymbol{\Lambda}^{(1)}, \boldsymbol{\Lambda}^{(2)}, \ldots, \boldsymbol{\Lambda}^{(i)}, \ldots, \boldsymbol{\Lambda}^{(S)}), \tag{2.18}$$

where

$$\boldsymbol{\Lambda}^{(i)} = \begin{pmatrix} \gamma^- \leq \gamma^{(i)} \leq \gamma^+ \\ \hat{n}_s^- \leq \hat{n}_s^{(i)} \leq \hat{n}_s^+ \\ p_1^- \leq p_1^{(i)} \leq p_1^+ \\ p_3^- \leq p_3^{(i)} \leq p_3^+ \\ a_1^- \leq a_1^{(i)} \leq a_1^+ \\ a_2^- \leq a_2^{(i)} \leq a_2^+ \\ \omega_1^- \leq \omega_1^{(i)} \leq \omega_1^+ \\ \omega_2^- \leq \omega_2^{(i)} \leq \omega_2^+ \end{pmatrix} \tag{2.19}$$

2. Compute fitness of each string by evaluating $\Pi(\boldsymbol{\Lambda}^{(i)})\ \forall i$
3. Rank the genetic strings where the top rank has the minimum cost function $\Pi(\boldsymbol{\Lambda}^{(i)})$
4. Mate the top pairs of genetic strings to obtain two children, such that

$$\boldsymbol{\Lambda}^{(ci)} = \begin{pmatrix} \gamma^{pi}\phi_1 + \gamma^{p(i+1)}(1-\phi_1) \\ \cdots \\ \cdots \\ \omega_2^{pi}\phi_8 + \omega_2^{p(i+1)}(1-\phi_8) \end{pmatrix},$$

where $\phi_j \in \text{rand}[0,1]$.

$$\boldsymbol{\Lambda}^{(c(i+1))} = \begin{pmatrix} \gamma^{p(i+1)}\hat{\phi}_1 + \gamma^{pi}(1-\hat{\phi}_1) \\ \cdots \\ \cdots \\ \omega_2^{p(i+1)}\hat{\phi}_8 + \omega_2^{pi}(1-\hat{\phi}_8) \end{pmatrix},$$

where $\hat{\phi}_j \in \text{rand}[0,1]$.

5. Remove bottom $S-P$ original strings from population. Generate $S-P-P$ new random genetic strings.
6. Repeat steps 2–5 with a new population until either one of these conditions is met
   - G generations has been reached.
   - $\min(\Pi) \leq \mathrm{TOL}$.

3. Results

(a) Convergence study

A convergence study was conducted to observe the dependence of the physical model on sources of randomness of the simulation. The variance of the cost was observed as a function of the number of rays in the simulation. To do so, a single set of design parameters was chosen. For each number of rays tested, the simulation was run 100 times, and the cost parameters were saved.

Figure 8 outlines the results of the sensitivity study. Figure 8*a–c* depicts the variance of the overall cost, the solar cost and the plant cost parameter over 100 test runs for the same design parameters. Figure 8*d* summarizes the standard deviation of the performance parameters as the number of rays increases. For the chosen number of rays for the numerical example outlined in the next section (500 rays), the standard deviation for all of the parameters falls below 3×10^{-3}.

In choosing a representative numerical example discussed in the following section, it was important to consider a number of rays that would lead to reproducible and accurate results while not sacrificing the efficiency of the simulation.

(b) Numerical example

A numerical example is generated based on the model previously described (source code available at https://github.com/ucb-msol/AgroPV.git). A single 'flash' of light was used to represent an hour of irradiation. The Pysolar Python package was used to determine the azimuth and elevation angle of the incoming rays of light for a specified date, time and location [13]. Based on these angles, the clear sky solar irradiance was determined from the package. The greenhouse was simulated for 1 July 2021 in Berkeley, California. Figure 9 outlines the solar irradiance distribution over the simulated day and location. It was assumed that the greenhouse had no obstruction from the sun in its surroundings, and there were no terrain obstructions on the sun throughout the day. The model can be trivially extended to include terrain obstructions but was not considered for this study.

The simulation was run for every hour throughout the day at which the altitude of the sun was above the horizon. For this particular time of year, the sun was located over the horizon from the hours of 5.00 and 19.00. As such, 15 iterations of the ray-tracing simulation were conducted for each design to determine its associated cost. The physical parameters used in the system are outlined in table 1. The ground refractive index was chosen based on that of a leaf [14]. The fixed geometric exponent in the $\mathbf{e}_2$ axis was chosen to ensure the projected shape of the greenhouse would be rectangular from a aerial view. Lastly, the fixed generalized radii were fixed to allow the topology optimizer to be agnostic to the size of the incoming beam of light. This allowed for topological updates focused on the shape and waveform of the greenhouse rather than its size. A volume penalty term could be added to the cost if space constraints are of concern for a particular system.

Table 2 outlines the parameters used to set up the genetic algorithm in the numerical example shown. The number of design strings, parents and generations were chosen to be relatively low to highlight the efficacy of the computational framework without looking for a true optimized solution. The search bound for the solar panel transmissibility was chosen from opaque to transparent. The solar refractive index was designed for optimal light reflection/refraction angles. It does *not* account for the particular wavelengths of the incoming rays. This is discussed further in §4. Ultimately, the search bound parameters chosen for this example are arbitrary as further

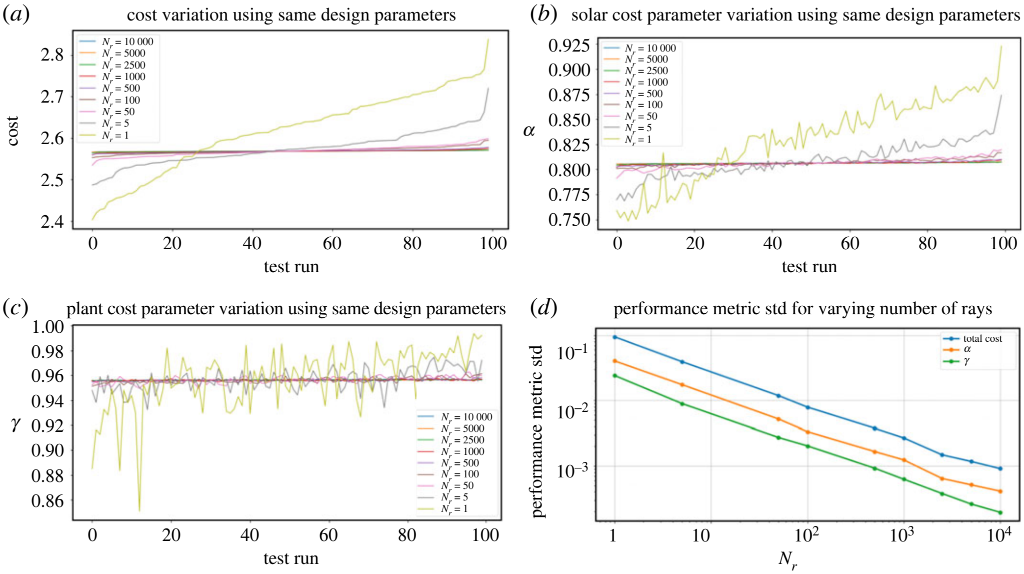

Figure 8. Summarized results of sensitivity study. (*a*) Sensitivity of total cost over 100 simulation runs, (*b*) sensitivity of solar cost parameter over 100 simulation runs and (*c*) sensitivity of plant cost parameter over 100 simulation runs. (*d*) Standard deviation of cost parameters for varying number of rays. (Online version in colour.)

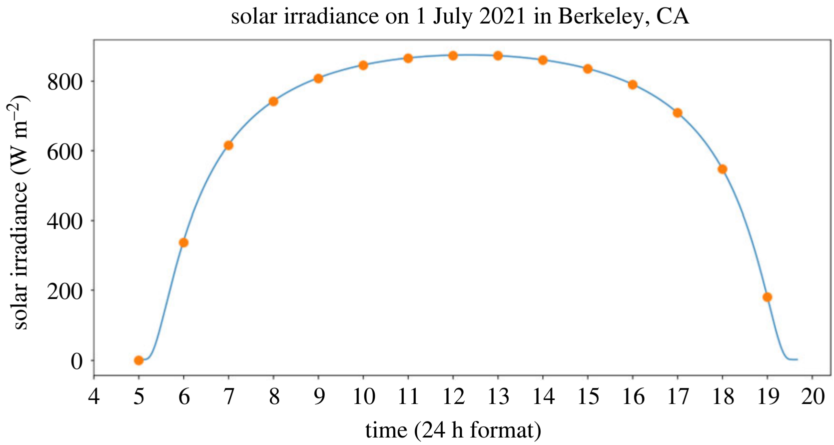

Figure 9. Evolution of clear sky solar irradiance on 1 July 2021 in Berkeley, California. (Online version in colour.)

verification and application of this framework requires an advanced degree of expertise about manufacturing, material and ambient constraints to drive realistic search bounds for the design parameters of this system.

Figure 10 depicts the convergence of the cost function across 200 generations. The figure on the left indicates the cost of the best-performing design after each generation while the figure on the right highlights the average cost of the parent strings at the end of each generation. The optimal design string parameters at the end of the final generation are displayed in table 3. The uncertainty in the cost was determined using the results of the convergence study, and is reported with a 95% confidence interval over 100 test runs. Given the limited number of generations used,

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 478: 20220414

Table 1. Numerical example—greenhouse system parameters.

symbol	type	units	value	description
N_r	scalar	none	500	number of light rays
n_g	scalar	none	1.4	ground refractive index
c	scalar	m s^{-1}	3×10^8	speed of light
t	scalar	m	0.1	solar panel thickness
$[R_1, R_2, R_3]$	scalar	none	[0.5, 5, 0.5]	generalized radii
p_2	scalar	none	20	geometric exponent

Table 2. Numerical example—genomic optimization parameters.

symbol	type	units	value	description
parents	scalar	none	6	surviving strings for breeding
S	scalar	none	20	designs per generation
G	scalar	none	200	total generations
$[\gamma^-, \gamma^+]$	scalar	none	[0.25,1]	solar panel transmissibility
$[\hat{n}_s^-, \hat{n}_s^+]$	scalar	none	[2,5]	solar panel rrefractive index
$[p^-_{1\,\mathrm{or}\,3}, p^+_{1\,\mathrm{or}\,3}]$	scalar	none	[1, 20]	geometric exponents
$[a^-_{1\,\mathrm{or}\,2}, a^+_{1\,\mathrm{or}\,2}]$	scalar	none	[0, 1.75]	sinusoid amplitudes
$[\omega^-_{1\,\mathrm{or}\,2}, \omega^+_{1\,\mathrm{or}\,2}]$	scalar	none	[0, 10]	sinusoid frequencies
w_1	scalar	none	2	weight of solar panel power in net cost
w_2	scalar	none	1	weight of photosynthetic power in net cost
$P_{\mathrm{solar,des}}$	scalar	W	$1/3P_0$	desired power absorbed by solar panel
$P_{\mathrm{plant,des}}$	scalar	W	$1/6P_0$	desired power absorbed by plants

uncertainty in the optimal design parameters is not trivially quantifiable given the random nature of the genomic-based optimizer and the non-convexity of the cost function.

These 'optimal' design parameters translate into the 'optimal' greenhouse illustrated in figure 11. The ellipsoidal components of the generalized greenhouse contour equation are highlighted by the average value between the peaks, and follow a unique version of the shape depicted in figure 3*a* while the alternating patterns denote the sinusoidal component of the generalized greenhouse contour equation as depicted in figure 3*b*. The intended nature of this optimized shape and physical intuition is further discussed in §4.

Figure 12 outlines key snapshots of the simulation at different times of the day. The times are broken down by every 2 h to illustrate significant changes in the angle of incoming light to the greenhouse. The time was chosen to depict collision of the light rays with the greenhouse and ground as well as reflection of the rays simultaneously. In the earlier part of the day, the light rays are observed to become entrapped in reflections within the 'sawtooth' convex section of the greenhouse, allowing for increased reflections. The convex portions of the greenhouse on the opposite end of the y-axis (figure 12) are not as large in magnitude, but still provide an entrapment effect for the greenhouse.

A closer look at a full simulation is outlined in figure 13 which outlines an incoming flash of light at 12.00 on 1 July 2021. As expected, the light rays enter the domain nearly perpendicular to the ground. The initial impact of the rays leads to pure reflection from the concave portion of the greenhouse. Refraction is not observed until rays come in contact with the convex portion

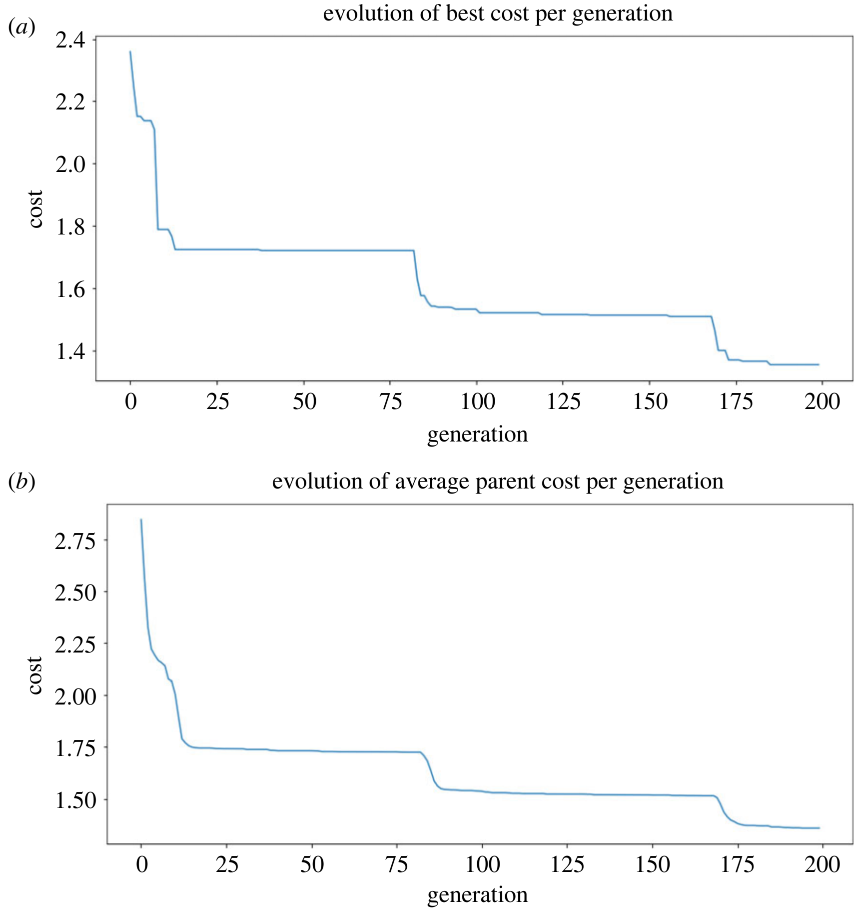

Figure 10. Evolution of best cost (*a*) and average parent cost (*b*). (Online version in colour.)

Table 3. Numerical example—optimal greenhouse design parameters.

Π	γ	$\hat{n}$	p_1	p_2	a_1	a_2	ω_1	ω_2
1.35 $\pm$ 0.0028	0.26	2.00	12.45	6.64	1.00	1.55	4.76	7.51

of the greenhouse situated along the x-axis. At this point, there is a period of reflection and refraction within the 'sawtooth' until the rays are reflected out or move below the designated power threshold. The simulation is stopped when there are no active rays remaining in the domain.

4. Discussion

The physical intuition of the optimized results can be understood through the design of the cost function. Based on the cost, the optimization scheme worked to maximize the power absorbed by both the solar panels as well as the crops (defined by the ground within the greenhouse). The physical nature of an APV system deems the majority of the light will come in contact with the greenhouse exterior before refracting into its interior. As such, the algorithm is able to find an optimal path by first designing the greenhouse to generate as much power as possible since it will

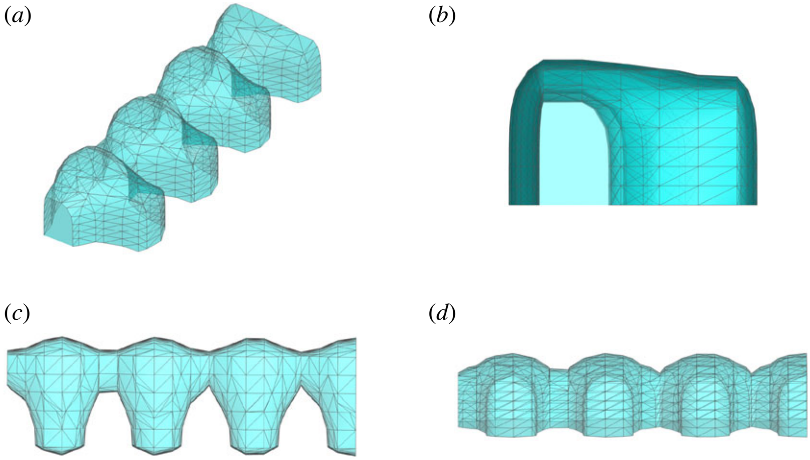

Figure 11. Detailed views of optimized greenhouse design. (*a*) Isometric view of optimized greenhouse design, (*b*) side view of optimized greenhouse design, (*c*) top view of optimized greenhouse design and (*d*) front view of optimized greenhouse design. (Online version in colour.)

originally capture the clear sky radiation. It is observed that the optimal design was chosen such that the greenhouse was 74% opaque and thus greenhouse power generation is still prioritized.

This work provides a framework for producing fast, optimized APV greenhouse designs given a particular location and time of year. Tuning parameters such as opacity to enforce the algorithm to allow a baseline power of light within the greenhouse could allow for less prioritization of greenhouse power generation. An alternative and potentially more effective strategy for equal prioritization of both power generation and plant growth is to consider wavelength sensitive greenhouses. One such example of these systems is explored by Loik *et al.* [7]. This APV system uses thin-film solar cells surrounding a greenhouse which work to filter designated wavelengths of light which are solar cell sensitive while refracting wavelengths of light that are crop sensitive for photosynthesis. This not only refines the power sharing strategy outlines in figures 6 and 7, but also allows for more efficient processing of light for power generation and photosynthesis.

A wavelength-dependent computational framework also provides merit in addressing the cost reduction limitations of this physical system. These limitations are apparent in the cost evolution shown in figure 10. The cost reduction after 200 generations is limited to roughly a 42.5% improvement from the best design in the initial generation. Ultimately, a wavelength-dependent simulation would allow for a more refined cost reduction if the PAR band of plants is used when designing the greenhouse system.

The optimal shape of the greenhouse follows a sawtooth cyclical pattern. The top view shown in figure 11*c* most clearly illustrates this pattern. The physical intuition of this optimized geometry follows providing convex caves as a means of entrapping the incoming light throughout a single day. These entrapment zones are aligned along the x-axis, consistent with the incoming direction of the radiation. It is also observed that these entrapment zones are biased towards the rays in the first half of the day. This can be attributed to the particular nature of the solar irradiance for this particular day and location. As shown in figure 9, the sun is above the horizon just before 5:00 in the morning, and the sun is below the horizon just after 19:00. As such, the first half of the day simulation includes one more flash of light in comparison to the second half of the day leading to a skewed entrapment of light for the first half of the day. This design, nonetheless, can be used as inspiration for potential transparent greenhouse designs in which the greenhouse is symmetric about the x-axis to account for multiple different seasons within the year.

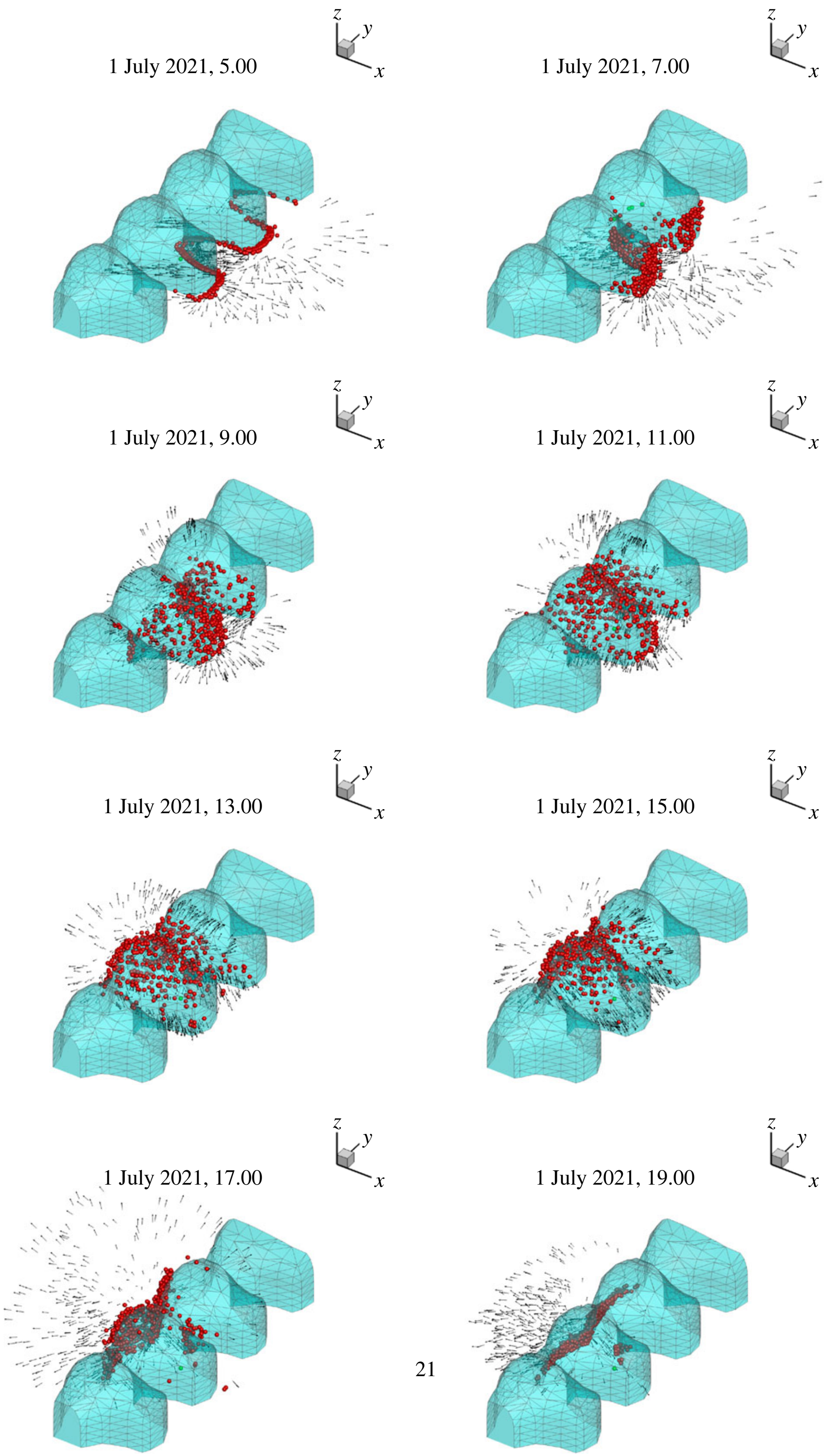

Figure 12. Full day simulation of best design. (Online version in colour.)

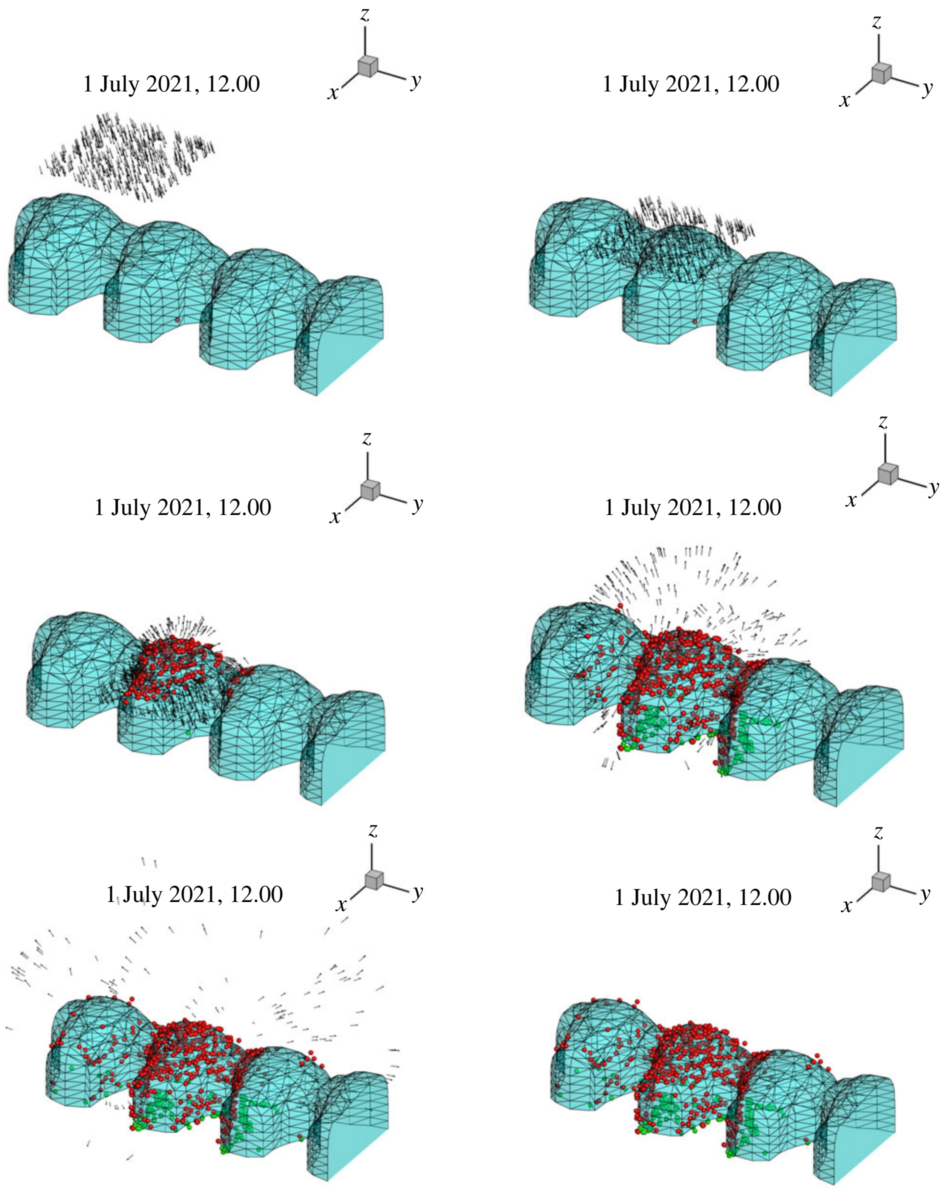

Figure 13. Detailed simulation of light flash at 12.00. (Online version in colour.)

5. Summary and extensions

The presented system captures the essential physics in a computationally efficient manner. However, there are many improvements to the model presented. For example, this work can be extended to include detailed solar panel translucency analysis. As mentioned earlier, solar panel translucency is directly tied to the PAR a plant needs, which will ultimately influence the optimal solar-cell design for a specific plant to be grown in a self-sufficient greenhouse. Another example is the possibility of using alternative optimization techniques, such as Bayesian optimization methods, to extensively probe the given search space for this problem. Current work of the authors includes incorporating solar-panel design into self-sufficient greenhouse modelling and improved optimization frameworks, as well as integrating wavelength sensitivity into the physical model.

Data accessibility. AgroPV Source Code Repository: https://github.com/ucb-msol/AgroPV.git.

Authors' contributions. R.S.I.: investigation, methodology, software, writing—original draft, writing—review and editing; E.M.: conceptualization, data curation, formal analysis, software, visualization, writing—original draft, writing—review and editing; T.I.Z.: conceptualization, formal analysis, funding acquisition, methodology, project administration, resources, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This work has been partially supported by the UC Berkeley College of Engineering, the USDA AI Institute for Next Generation Food Systems (AIFS), USDA award no. 2020-67021-32855.

References

1. California Department of Conservation, Williamson Act Program. See www.conservation.ca.gov/dlrp/wa.
2. California Air Resources Board, CARB releases ambitious draft climate action plan to slash use of fossil fuels and reach carbon neutrality by 2045. See https://ww2.arb.ca.gov/news/carb-releases-ambitious-draft-climate-action-plan-slash-use-fossil-fuels-and-reach-carbon.
3. Beck M, Bopp G, Goetzberger A, Obergfell T, Reise C, Schindele S. 2012 Combining PV and food crops to agrophotovoltaic–optimization of orientation and harvest. In *EUPVSEC Proceedings*, vol. 1. *27th European Photovoltaic Solar Energy Conference and Exhibition, Frankfurt, Germany, 24–28 September 2012*, pp. 4096–4100. München, Germany: EUPVSEC. (doi:10.4229/27thEUPVSEC2012-5AV.2.25)
4. Schindele S *et al.* 2020 Implementation of agrophotovoltaics: techno-economic analysis of the price-performance ratio and its policy implications. *Appl. Energy* **265**, 114737. (doi:10.1016/j.apenergy.2020.114737)
5. Weselek A, Ehmann A, Zikeli S, Lewandowski I, Schindele S, Högy P. 2019 Agrophotovoltaic systems: applications, challenges, and opportunities. A review. *Agron. Sustain. Dev.* **39**, 1–20. (doi:10.1007/s13593-019-0581-3)
6. Hassanien RHE, Li M, Dong Lin W. 2016 Advanced applications of solar energy in agricultural greenhouses. *Renew. Sustain. Energy Rev.* **54**, 989–1001. (doi:10.1016/j.rser.2015.10.095)
7. Loik ME, Carter SA, Alers G, Wade CE, Shugar D, Corrado C, Jokerst D, Kitayama C. 2017 Wavelength-selective solar photovoltaic systems: powering greenhouses for plant growth at the food-energy-water nexus. *Earth's Future* **5**, 1044–1053. (doi:10.1002/2016EF000531)
8. Ezzaeri K *et al.* 2018 The effect of photovoltaic panels on the microclimate and on the tomato production under photovoltaic Canarian greenhouses. *Sol. Energy* **173**, 1126–1134. (doi:10.1016/j.solener.2018.08.043)
9. Cossu M, Ledda L, Urracci G, Sirigu A, Cossu A, Murgia L, Pazzona A, Yano A. 2017 An algorithm for the calculation of the light distribution in photovoltaic greenhouses. *Sol. Energy* **141**, 38–48. (doi:10.1016/j.solener.2016.11.024)
10. Nakoul Z, Bibi-Triki N, Kherrous A, Bessenouci M, Khelladi S. 2014 Optimization of a solar photovoltaic applied to greenhouses. *Phys. Proc.* **55**, 383–389. 8th International Conference on Material Sciences, CSM8-ISM5. See https://doi.org/10.1016/j.phpro.2014.07.055. (doi:10.1016/j.phpro.2014.07.055)
11. La Notte L, Giordano L, Calabrò E, Bedini R, Colla G, Puglisi G, Reale A. 2020 Hybrid and organic photovoltaics for greenhouse applications. *Appl. Energy* **278**, 115582. (doi:10.1016/j.apenergy.2020.115582)
12. Zohdi T. 2021 A digital-twin and machine-learning framework for the design of multiobjective agrophotovoltaic solar farms. *Comput. Mech.* **68**, 357–370. (doi:10.1007/s00466-021-02035-z)
13. Stafford B. 2020 Pysolar (0.9). (doi:10.5281/zenodo.5518074)
14. Jacquemoud S, Baret F. 1990 PROSPECT: a model of leaf optical properties spectra. *Remote Sens. Environ.* **34**, 75–91. (doi:10.1016/0034-4257(90)90100-Z)