Greenhouse Climate Setpoint Optimization: An Online Decision Strategy

Yuanping Su¹, Lihong Xu²

¹ School of Energy and Mechanical Engineering, Jiang Xi University of Science and Technology, Shuanggang RD., 330013 Nanchang, China
² College of Electronics and Information Engineering, Tongji University, Caoan RD., 201804 Shanghai, China

Corresponding author: Yuanping Su (e-mail: suyuanping_2003@163.com).

This work was supported in part by the National Natural Science Foundation of China [grant number 61863015] and [grant number 61973337], in part by the U.S. National Science Foundation’s BEACON Center for the Study of Evolution in Action under cooperative Agreement [grant number DBI-0939454].

ABSTRACT Since the greenhouse climate setpoint greatly impacts the energy consumption and crop yield, optimizing the setpoint can significantly improve the energy efficiency. However, the great uncertainty of the weather makes such optimization problem be difficult to solve. Therefore, how to handle the great uncertainty becomes a challenge for setpoint optimization. To solve this problem, this work proposes an online receding horizon multi-objective optimization method to make a trade-off between the total energy consumption and the final crop yield to obtain a set of good greenhouse climate setpoints. The proposed method uses a surrogate assisted multi-objective optimization algorithm to minimize the energy consumption and maximize the crop yield to obtain the optimal mean temperatures of the crop development stages. Since the proposed method does not directly optimize the setpoint, a serialization method is proposed to transform the mean daily temperature into the setpoint of the inside temperature. In addition, since daily predicted weather is usually rough, this work developed an interpolation method for weather, which was validated using real weather data collected in a Venlo-type greenhouse. The proposed greenhouse climate setpoint optimization method was compared with the Priva system, and the results indicate the advantage of the proposed method.

INDEX TERMS greenhouse climate setpoint; receding horizon optimization; weather interpolation; multi-objective optimization; RBF surrogate model

I. INTRODUCTION

A. ABBREVIATIONS AND ACRONYMS

The advantage of greenhouse production is that the greenhouse climate can be accurately regulated to create a favourable growth environment when some control actions such as heating, fogging, CO₂ injecting are employed [1-2]. Under such situation, the yield and quality of the agricultural product can be improved greatly. However, the energy consumption cost of the greenhouse production is usually high, such that it is difficult for the farmer to get satisfactory financial return. Actually, the energy consumption cost is more than 78% of total greenhouse production cost, and the heating energy demand represents 65-85% of total greenhouse energy demand [3]. Therefore, how to reduce the energy consumption becomes a crucial issue of greenhouse production. To improve the production efficiency, many energy saving methods, such as optimizing the structure and materials of the greenhouse and introducing the biomass energy, have been proposed in recent years [3-5]. But most methods usually require additional construction cost, which cannot be afforded by the farmers. Therefore, developing a low-cost energy conservation method is of importance.

Greenhouse production practices indicate that the greenhouse climate setpoint significantly impact the energy cost and crop yield of greenhouse production. Therefore, dynamically optimizing greenhouse climate setpoints not only can improve crop yield, but also can reduce the total energy consumption of the greenhouse. Therefore, optimizing the greenhouse climate setpoint is an effective way to save energy without additional cost.

Currently, there are two ways that can be used to solve the optimal greenhouse cultivation problem. One is to directly optimize the setpoint of the greenhouse climate by maximizing the economic profit or minimizing the energy consumption, and the other is to obtain the optimal control strategy by solving the optimal control problem of the greenhouse cultivation. However, both ways have difficulty
to deal with the great uncertainty of the long-term weather and the multi-timescale issue of the system.

According to temperature integration theory, the crop growth mainly depends on the accumulated temperature over a certain period rather than the transient temperature. Therefore, if it can ensure the given accumulated temperature over a certain period, the crop growth will not be significantly influenced, which means if the weather is cold, it is better to set a small temperature setpoint, and the loss of the accumulated temperature can be compensated in the warm days. In this case, it only optimizes the accumulated temperature or the mean temperature over a certain period to generate the temperature setpoint. Based on such idea, many setpoint optimization methods have been proposed in the past years, for example, Jones et al reported a method to determine greenhouse daily temperature based on the POLY-2 greenhouse climate model and the TOMGRO tomato growth model [31]. Tchamitchian et al [6] developed a decision support system “SERRISTE” for tomato crop to determine the mean daily temperature setpoint. Sigrimis [7] proved that temperature integration based greenhouse climate setpoint programming is energy saving. Additionally, another advantage of the temperature integration based method is that it can solve the multi-timescale problem of the system, and it also can greatly reduce the number of the design variables of the setpoint optimization problem, which benefit to the convergence of the optimization process.

The second way mentioned early is the so-called two timescale receding horizon optimal control. The basic idea is that the fast sub-system, i.e., the greenhouse climate dynamics, is considered as a quasi-steady system, such that the greenhouse climate variables can be viewed as a function with respect to the control inputs. Therefore, using the Pontryagin’s Maximum Principle and the historical weather, the optimal control problem of the crop growth can be solved, and the optimal control strategy and the optimal setpoint trajectory can be obtained. Since the historical weather may be different from the present weather, the optimal control strategy obtained in the slow sub-system may not be reliable. Therefore, to deal with the uncertainty of the short-term weather, it can use the obtained co-state to solve the optimal control problem of the greenhouse climate on a small timescale, and such process is repeated every hour or half hour. In this respect, van Henten [8] did the groundbreaking work, and Tap [9] and Gonzalez[10] extended van Henten’s method to the on-line case. Since this method greatly depends on the system model, if there is large model error, then it is difficult to ensure the globally optimal performance of the control strategy or setpoint trajectory. Therefore, to improve the model performance using the online sampled data, Xu [11] proposed an adaptive two timescale receding horizon optimal control method.

Although both ways can achieve good results in some cases, their disadvantages still cannot be ignored. The traditional temperature integration based setpoint optimization is a kind of local optimization, i.e., it usually only minimizes the local energy consumption or maximizes local the crop growth of a certain period. Although such methods can effectively deal with the uncertainty of the short-term weather, the economic performance of the greenhouse cultivation may not be globally optimal. In addition, the temperature integration theory only considers the effect of the accumulated temperature on the crop growth, but in fact, the crop growth also significantly depends on the transient-temperature, photosynthetic active radiation and CO₂ concentration, and higher solar radiation usually requires higher temperature. Therefore, setpoint optimization not only considers the accumulated temperature, but also must consider the effect of the transient temperature and the solar radiation on the photosynthesis rate.

Theoretically, if the long-term weather is exactly known, then the two-timescale receding horizon optimal control method can achieve globally optimal control performance. However, its disadvantage is that if the control step-size of the slow sub-system is small such as one hour, even half hour, then the co-states obtained by using the historical weather is not accurate in practice. In addition, the dimension of the design space of the Hamiltonian minimization problem is very high. Therefore, in this case, the polynomial approximation is difficult to approximate the cost function, and the convergence of the nonlinear programming problem of the optimal control is difficult to ensure [12]. Therefore, how to improve reliability of the calculate co-states and reduce the dimension of the design space becomes the crux of such a nonlinear programming problem.

Actually, the historical weather reveals the fact that on a large timescale such as several weeks or one month, the average weathers across different years are usually similar. Therefore, based on long-term average weather, if the temperature integration optimization can be performed on a large timescale such as several days, the optimization result may be more reliable than the result of the receding horizon optimal control. In addition, if it can consider the effect of the solar radiation and the transient temperature on the the photosynthesis, the obtained setpoint of the greenhouse climate would be better.

According to such idea, we proposed a multi-layer hierarchical optimization method to optimize the setpoints of the greenhouse climate [38]. In this method, the setpoint optimization problem can be divided into a multi-objective optimization sub-problem with large timescale and a series of local single optimization sub-problems with small timescale that can be solved in different optimization layers. The multi-objective optimization sub-problem is an offline global optimization issue based on the historical long-term weather. The aim is to obtain the optimal mean temperature of the development stages of the crop. Based on the optimal mean temperature of the development stages and the short-term weather forecast, the optimally daily mean temperature can be obtained by minimizing the energy.
consumption online. This method can achieve good optimization results due to the global and local optimizations if the historical long-term weather is similar to the current long-term weather.

However, the current long-term weather may be different from the historical long-term weather at a small timescale, i.e., the long-term weather has great uncertainty. In this circumstance, the optimal solution of the globally offline multi-objective optimization may be far from the true optimal solutions, and the obtained setpoints may not be energy-saving. In addition, the setpoint serialization must solve several optimization problems to obtain the setpoint sequence, which spends much computational time, and the setpoint curve may not be smooth.

If the weather over a long period can be roughly predicted, for example, we may predict the average temperature of a month, then using online multi-objective optimization may achieve better energy-saving result. In fact, the energy-saving performance of the multi-objective optimization and machine learning has been validated in many engineering fields [33-35], so for saving the computational time of the optimization method and deal with the great uncertainty of the long-term weather at a small timescale, this work proposes a receding horizon surrogate assisted multi-objective optimization method.

However, the difficulty is that the crop growth cycle is usually long, for example, the growth cycle of a kind of infinite-growth-typed tomato is more than 300 days. Therefore, the simulations of the greenhouse climate and crop growth are computationally expensive, which greatly impacts the real-time performance of the optimization results. To solve the computationally expensive optimization problem, this work proposes a surrogate assisted multi-objective method. Since the computation of the surrogate is usually much cheaper than the true objective, using the surrogate to explore the globally optimal can save much time of the optimization process. Due to such advantage, surrogate assisted optimization has been excessively studied in the past decade [13]. For multi-objective optimization, there are several typical algorithms including ParEGO [14], MOEA/DEGO [15], and EIMEGO [16]. They uses Gaussian process (GP) or kriging surrogate to approximate the true objectives. However, if the number of the decision variables is large, the Gaussian process and kriging may not be effective. Since the number of the decision variables of the greenhouse climate setpoint optimization problem is more than twenty, GP based optimization algorithms may not ensure the globally optimal performance and the real-time performance. Therefore, this work proposed a new weighted multi-objective optimization method based on radius basis function (RBF) model.

B. CONTRIBUTION

The contribution of this work lies in forth aspects: (1) this work proposes an online receding horizon multi-objective optimization method for greenhouse climate setpoint optimization, which can solve the multi-timescale optimization problem and deal with the uncertainty of the long-term weather; (2) this work proposes a new interpolation method of weather; (3) a serialization method of inside temperature is developed to convert the daily and stage mean temperatures to the daily temperature trajectory for the simulations of the crop yield and greenhouse climate control; (4) this work proposes a RBF surrogate based multi-objective optimization method by weighting the approximated objectives.

The main differences between the proposed method and the one of [38] are as follows:

(1). The method proposed in this work has one optimization layer, i.e., the global multi-objective optimization is performed online every day to obtain the optimal mean temperature of the development stages, while the method presented in [38] has two optimization layers. In the first layer, the global multi-objective optimization is carried out once offline to obtain the optimal the optimal mean temperatures of the development stages. Next, based on such optimal mean temperature, the local single objective optimization is performed to minimize the energy consumption of each stage to obtain the optimal daily mean temperature in the second layer.

(2). In the proposed method, a surrogate assisted multi-
objective optimization algorithm is developed to solve the global multi-objective optimization problem of the energy consumption and the crop yield, which can deal with the great uncertainty of the long-term weather forecast. But the multi-layer optimization method uses NSGA-II to offline maximize the crop yield and minimize the energy consumption based on the historical weather data. The former is a data-driven optimization method, while the latter is a deterministic optimization.

(3). In the proposed method, the daily mean temperature inside the greenhouse is generated according to the proportions of the daily mean temperature outside of the greenhouse, while in the multi-layer optimization framework, the daily mean temperature inside the greenhouse is obtained by minimizing the energy consumption.

(4). The proposed method uses sine function and exponential function to construct the setpoint curves, while the method presented in [38] use parabolic function and straight line to construct the setpoint curve according to the allocation of the accumulated temperature. Generally, the former is smoother than the latter, but the latter may have better energy-saving performance due to the optimization of the accumulated temperature.

II. ONLINE RECEeding HORIZON OPTIMIZATION FRAMEWORK

For greenhouse cultivation, the timescale of the control step-size is much smaller than the overall crop growth cycle. Therefore, there are many decision variables in the nonlinear programming problem of the optimal control. For example, if the crop growth cycle is 260 days, and the step-size is set as 15 minute, then there are 24960 control steps. If it directly optimizes the temperature setpoint, then it must set 24960 decision variables for the setpoint optimization problem. Obviously, the limited function evaluations are usually difficult to ensure the convergence of optimization algorithm within such high-dimensional design space, and the weather has greater uncertainty at such a small timescale. Therefore, it must divide the crop growth cycle into N crop development stages, and each stage includes M days. Thus, the mean temperature of the stages can be considered as the decision variables. If M is set as 10 days, then there are 26 decision variables, such that the number of the decision variables is greatly reduced. Since long-term weather forecast is unavailable, it must combine the historical weather with the current weather forecast to construct the weather dataset. Using the weather dataset,
the surrogate assisted multi-objective optimization of the energy consumption and crop yield can be carried out at 00:00 every day to determine the mean daily temperature of the next day. When the mean daily temperature is determined, the serialization method can be used to convert the mean daily temperature to the setpoint. The principle of the receding horizon optimization is illustrated in Fig. 1, and the flowchart is shown in Fig. 2.

The initial states of the greenhouse climate and crop growth is updated every day using the measured values of the greenhouse climate and crop growth states, which is equivalent to introducing the state feedback to reduce the model error of the system. To a certain extent, such feedback can deal with the uncertainty of the long-term weather forecast. It should be noted that the timescale of the first decision variable is not fixed, and the number of the decision variables reduces with the crop growth.

As illustrated in Fig. 1, after the plant completes a development stage, the number of the decision variables will be reduced by one. From dynamic multi-objective optimization point of view, the proposed online multi-objective receding horizon optimization is a special dynamic multi-objective optimization problem.

In summary, the receding horizon optimization of the greenhouse climate setpoint must solve the following problems:

1. Transforming the roughly predicted weather into the time series of the weather.
2. Transforming the mean daily temperature into the time series of the temperature setpoint.
3. Using a surrogate assisted multi-objective optimization algorithm to find a best trade-off solution of the energy consumption and crop yield.

In the following sections, the above problems will be presented in detail.

### III. INTERPOLATION OF THE PREDICTED WEATHER

Generally, the weather forecast only predicts the maximum and minimum values of the outside temperature, humidity, wind speed and radiation, but such prediction is rough for the simulations of the greenhouse climate and crop growth. Therefore, it must use such weather forecast information to generate the minutely weather.

#### A. INTERPOLATION OF OUTSIDE TEMPERATURE

Generally, the daily temperature changes with different laws within different time intervals, e.g., it usually changes slower during the night than the day. Therefore, when simulating the daily temperature outside the greenhouse, it is better to use different curve to approximate the outside temperature of the different time intervals. Actually, to generate hourly temperature using the observed maximum and minimum temperature of a day, many approaches have been proposed in the past years [17-18]. These methods use a piecewise cosine or sine function to approximate the temperature. However, their simulation error may be large in some cases due to the insufficient information, and the junction between the consecutive days may not be continuous and smooth. Therefore, to improve the interpolation performance, additional parameters for the interpolation of the temperature are introduced in the proposed method. They are the outside temperature values at the sunrise and sunset instants, the time instants at which the maximum and minimum temperatures appear, and the starting temperature value and the last temperature of a day.

As presented in [19], according to the change law of the daily temperature, it is reasonable to divide a day into 4 time intervals, and to use different curves to approximate the temperatures within the time intervals, as shown in Fig. 3.

Although weather forecast can predict the maximum and minimum values of the daily outside temperature $T^{\text{max}}_{\text{end}}$ and $T^{\text{min}}_{\text{end}}$, the times that both temperatures appear are not predicted. Generally, the times that the maximum and minimum temperatures appear usually depend on the geographic position and cloudy condition. By analyzing the historical weather in Shanghai, China, we found that the maximum and minimum values of the outside temperature, humidity, wind speed and radiation, but such prediction is rough for the simulations of the greenhouse climate and maximum temperature usually appears between 13:00 and 14:00, and the minimum temperature often appears at the time near to the sunrise time if there is no cloud, i.e., it can...
set $T_0 = T_{\text{min}}^{\text{out}}$. Generally, the length of daylight (number of hours) varies with the seasons, so the sunrise time and sunset time are usually not fixed. According to William's study [20], the length of daylight can be calculated by:

$$DL = 24 - \frac{24}{\pi} \cos^{-1} \left[ \sin \left( \frac{p \cdot \pi}{180} \right) + \sin \left( \frac{L \cdot \pi}{180} \right) \cdot \sin \phi \right]$$ (1)

$$\phi = \sin^{-1} \left[ 0.39795 \cos \theta \right]$$ (2)

$$\theta = 0.2163108 + 2 \tan^{-1} \left( 0.9671396 \tan \left[ 0.0086(k-186) \right] \right)$$ (3)

where $L$ is the latitude, $p$ is the day length coefficient, which can be set as 0.8, and $k$ is the day index of year. From the historical weather data, it can be found that the sunrise time and sunset time are usually near to 6:00 and 18:00, respectively, so the sunrise time $t_{\text{sunrise}}$ and the sunset time $t_{\text{sunset}}$ can be roughly estimated respectively by:

$$t_{\text{sunrise}} = 6 \cdot 00 + \lceil 0.5(12 - DL) \rceil$$ (4)

$$t_{\text{sunset}} = 18 \cdot 00 + \lceil 0.5(DL - 12) \rceil$$ (5)

where $\lceil \rceil$ is the time operator, i.e., if $0.5(DL - 12)$ is decimal, the time operator transforms it into minute. The time that the maximum value of the outside temperature appears $t_{\text{max}}$ can be determined by

$$t_{\text{max}} = t_{\text{sunrise}} + 0.5DL$$ (6)

The historical weather data indicates that the outside temperature at 8:00 is usually close to the mean daily temperature of a day, as shown in Fig. 4.

![Figure 4](image-url)

Fig. 4 correlation between the mean daily temperature and the temperature at 20:00

Therefore, it is reasonable to set the start time $t_s$ as 20:00 of the former day, and the terminal time $t_{\text{end}}$ is set as 20:00 of the current day, so the temperature value of point S in Fig. 3 can be set as the mean daily temperature of the former day $T_{\text{out}}^{\text{daily}}$, and the temperature value of point F can be set as the mean daily temperature of the current day $T_{\text{out}}^{\text{daily}}$, i.e.,

$$T_S = T_{\text{out}}^{\text{daily}} - T_{\text{out}}^{\text{daily}}$$

In addition, the two middle points A and E are introduced, and their temperature values can be calculated by $T_A = p_1(T_S + T_D)$ and $T_E = p_2(T_D + T_T)$ with $0 < p_1 < 0.5$ and $0 < p_2 < 0.5$, respectively. The parameters $p_1$ and $p_2$ can be obtained by using parameter identification.

As described in [19], the temperature at sunset time can be estimated by $T_D = T_{\text{out}}^{\text{max}} - c \cdot (T_{\text{out}}^{\text{max}} - T_{\text{out}}^{\text{min}})$, where $c$ is a fitting coefficient, and depends on the month of year. In this work, $c$ is set as 0.39. According to [21], the mean daily temperature can be usually expressed as the linear combination of the maximum and minimum temperature, i.e.,

$$T_{\text{out}}^{\text{daily}} = \alpha \cdot T_{\text{out}}^{\text{max}} + \beta \cdot T_{\text{out}}^{\text{min}}$$

The parameter $\alpha$ and $\beta$ can be set as the fixed values 0.522 and 0.449, respectively. Thus, the temperatures of the segment 1 and segment 4 can be fitted using parametric cubic spline curve, and the temperatures of segment 2 and segment 3 can be estimated by sinc functions. They are described respectively by:

Segment 1:

$$T_{\text{out}}(t) = t^3 + a_1 t^2 + b_1 t + c_1$$ (7)

Segment 2:

$$T_{\text{out}}(t) = T_{\text{out}}^{\text{min}} + 0.5(T_{\text{out}}^{\text{max}} - T_{\text{out}}^{\text{min}}) \left[ 1 + \sin \left( \frac{\pi}{2} \right) \right]$$ (8)

Segment 3:

$$T_{\text{out}}(t) = T_D + (T_{\text{out}}^{\text{max}} - T_D) \cdot \sin \left( \frac{1 + \frac{t - t_{\text{max}}}{t_{\text{max}}} \pi}{2} \right)$$ (9)

Segment 4:

$$T_{\text{out}}(t) = t^3 + a_2 t^2 + b_2 t + c_2$$ (10)

The coefficients of the spline curves (7) and (10) can be identified using the known point S, A, B, D, E, F, but the disadvantage of the parametric cubic spline is that it must solve an equation set to obtain the coefficients. Therefore, for simplification, the method reported in [19] can be used to estimate both the temperature curves. Then, the outside temperature can be described by

$$T_{\text{out}}(t) = \begin{cases} T_{D - 1} + \Delta_i(t + t_{\text{end}} - t_{\text{min}}) & t_{\text{min}} \leq t \leq t_{\text{max}} \\ T_D + \Delta_H(t - t_{\text{min}}) & t_{\text{max}} \leq t \leq t_{\text{end}} \end{cases}$$ (11)

$$\Delta_i = \frac{T_{\text{out}}^{\text{max}} - T_{\text{out}}^{\text{end}}}{(t_{\text{max}} + t_{\text{end}} - t_{\text{min}})} \cdot \Delta_H = \frac{T_{\text{out}}^{\text{end}} - T_{\text{out}}^{\text{min}}}{(t_{\text{min}} + t_{\text{end}} - t_{\text{max}})}$$

where $z$ is the curvature parameter of the parabolic function, and is usually set as 0.5, superscript “-1” denotes day before, and superscript “+1” denotes day after.
Meanwhile, the temperature value of point D can be calculated by \[ T_d = T_{\text{out}}^{\text{max}} - c \cdot (T_{\text{out}}^{\text{max}} - T_{\text{out}}^{\text{min}}) \], and the parameter \( c \) is set as 0.39. Using these piecewise functions, it can obtain the temperature curve of a day.

**B. INTERPOLATION OF SOLAR RADIATION**

This work assumes that solar radiation sum can be predicted. If the solar radiation sum is given, then it should be transformed into the solar radiation series. Several solar radiation models [22-24] indicate that the solar radiation during the day usually changes along sine rule. Therefore, for a given solar radiation sum, it can use sine function to approximate the solar radiation curve, which can be described by:

\[
I_{\text{glob}}(t) = I_{\text{glob}}^{\text{max}} \sin \omega t' \quad (12)
\]

and

\[
t' = t - t_{\text{sunrise}} \\
\omega = \frac{\pi}{t_{\text{sunset}} - t_{\text{sunrise}}}
\]

where \( I_{\text{glob}}^{\text{max}} \) is the predicted maximum solar radiation.

**C. INSTANTANEOUS WIND SPEED SIMULATION**

According to Ephrath’s study [25], if the maximum and minimum values of the wind speed of a day are known, then the change curve of the daily wind speed can be approximated by a sine function, which is given by

\[
V_{\text{wind}}(t) = \begin{cases} 
V_{\text{min}} + V_{\text{max}} \times \sin \left( \pi \cdot \frac{t - t_{\text{min1}}}{t_{\text{min2}} - t_{\text{min1}}} \right) & t_{\text{min1}} \leq t \leq t_{\text{min2}} \\
V_{\text{min}} & \text{else}
\end{cases} 
\]

(13)

where \( V_{\text{min}} \) and \( V_{\text{max}} \) are the minimum and maximum values of the wind speed of a day, respectively. \( t_{\text{min1}} \) is the time that the minimum wind speed appears before the noon, and \( t_{\text{min2}} \) is the time that the minimum wind speed appears in the afternoon.

**D. INTERPOLATION OF RELATIVE HUMIDITY**

Relative humidity significantly depends on the atmosphere temperature \( T_{\text{out}} \) and the dew point temperature \( T_{\text{dew}} \), and can be calculated by:

\[
R \H_{\text{out}}(t) = \frac{E_m}{E_s} \times 100\% \quad (14)
\]

where \( E_m \) (Pa) is the measured vapour pressure, and \( E_s \) (Pa) is the saturation vapor pressure, which can be described by

\[
E_m = 2.229 \times 10^{11} \exp \left( -\frac{5383}{273.15 + T_{\text{out}}(t)} \right) \quad (15)
\]

The measured vapor pressure \( V_PA \) can be calculated by

\[
E_s = 2.229 \times 10^{11} \exp \left( -\frac{5383}{273.15 + T_{\text{dew}}(t)} \right) \quad (16)
\]

The dew point temperature can be determined by \( T_{\text{dew}} = \min(T_{\text{out}}, T_{\text{dew}}^{\text{max}}) \), where \( T_{\text{dew}}^{\text{max}} \) is the seasonal maximum dew point temperature of a certain geographical region.

Since the atmospheric CO\(_2\) concentration is usually stable, it usually changes around a constant, for example, in Shanghai, the CO\(_2\) concentration usually varies in 761±22 mg/m\(^3\). Therefore, the CO\(_2\) concentration outside the greenhouse can be set as a constant.

**IV. SERIALIZATION OF INSIDE MEAN DAILY TEMPERATURE AND ESTIMATIONS OF THE OBJECTIVES**

Since the proposed method does not directly optimize the setpoints of the greenhouse climate but the mean temperature, the optimization results cannot directly be used to simulate the crop growth and the greenhouse climate. Therefore, the mean temperature must be transformed into the minutely temperature to evaluate the crop yield and energy consumption. In the online receding horizon optimization framework, the training dataset of the objectives must be updated every day using the latest weather forecast, such that the final crop yield and total energy consumption should be re-calculated every day. However, the timescale of the decision variables of the optimization problem is too large to be directly used to calculate the crop yield and energy consumption. Therefore, the mean stage temperatures, i.e., the decision variables, must be transformed into the minutely temperature, which is so-called the temperature setpoint. Therefore, it must introduce a setpoint estimation method to generate the setpoint candidate.

In contrast to the traditional temperature integration based setpoint optimization and the receding horizon optimal control, the proposed method considers the long-term accumulated temperature and the short-term effect of the transient temperature inside the greenhouse on the photosynthesis. Essentially, the crop growth mainly depends on the photosynthesis. According to the photosynthesis rate model [26], for a given CO\(_2\) concentration, raising the temperature and the solar radiation can improve the photosynthesis, as shown in Fig. 5.

From Fig. 5, it can be seen that for a given CO\(_2\) concentration, if the photosynthetic active radiation (PAR) is low, then raising the temperature cannot significantly improve the photosynthesis rate, but it will consume much more heating energy. Therefore, from the energy saving point of view, if the solar radiation is small, then the temperature setpoint should be set as a small value. The basic idea of the proposed serialization approach of the mean daily temperature is to ensure the temperature...
setpoint to change with the solar radiation. Thus, it can save much heating energy without loss of the crop yield.

![Image](https://creativeworks.org/licenses/by/s.png)

Fig. 5 effects of canopy temperature and solar radiation on the photosynthesis when the CO2 concentration is 600 mg m⁻³

Denotes $\bar{T}_{\text{stage}}^i$ as the mean stage temperature of the $i$-th crop development stage, and let $\Delta t_{\text{sampling}}(s)$ be the sampling period of the greenhouse microclimate. Then, the aim of the setpoint serialization is to transform the mean stage temperature $\bar{T}_{\text{stage}}^i$ into $(86400 \cdot M) / \Delta t_{\text{sampling}}$ temperature values. To protect the crop growth from the extreme weather, the inside temperature must be bound by the minimum and maximum temperatures $T_{\text{min}}^\text{in}$ and $T_{\text{max}}^\text{in}$. If the weather is cold, then the inside temperature may be lower than the lower bound. In this case, it must provide the minimum heat energy to ensure the temperature to be within the bounds, and the corresponding temperature trajectory can be regarded as the basic temperature curve in the warm days. Therefore, the simulated temperature $T_{\text{in}}^0(t)$ must be corrected using the lower and upper bounds to obtain the basic temperature curve $T_{\text{basic}}^\text{in}(t)$. If the simulated temperature $T_{\text{in}}^0(t)$ is lower than $T_{\text{min}}^\text{in}$, then $T_{\text{in}}^0(t)$ can be set as $T_{\text{in}}^\text{min}$, i.e., $T_{\text{in}}^\text{basic}(t) = T_{\text{in}}^\text{min}$. If the maximum value of the simulated temperature $T_{\text{in}}^0(t)$ is higher than the upper bound, then the times $t_1$ and $t_2$, at which $T_{\text{in}}^0(t)$ first reaches the optimal temperature $T_{\text{in}}^\text{opt}$ in the morning and afternoon, can be determined, as shown in Fig. 6. Thus, the temperature between $t_1$ and $t_2$ can be approximated by a sine function, which is described by:

$$T_{\text{in}}^\text{basic}(t) = T_{\text{in}}^\text{opt} + \left( T_{\text{in}}^\text{max} - T_{\text{in}}^\text{opt} \right) \sin \left( \frac{\pi (t - t_1)}{t_2 - t_1} \right)$$  \hspace{1cm} (17)

Obviously, (17) means that the setpoint candidate of the inside temperature is not higher than the upper bound.

Therefore, , the basic temperature curve $T_{\text{in}}^\text{basic}(t)$ can be evaluated by

$$T_{\text{in}}^\text{basic}(t) = \begin{cases} T_{\text{in}}^\text{min} & \text{if } T_{\text{in}}^0(t) < 10 \\ T_{\text{in}}^\text{opt} + \left( T_{\text{in}}^\text{max} - T_{\text{in}}^\text{opt} \right) \sin \left( \frac{\pi (t - t_1)}{t_2 - t_1} \right) & \text{otherwise} \end{cases} \hspace{1cm} (18)$$

Then the mean daily temperature candidate of a day can be estimated by

$$\bar{T}_{\text{daily}} = \bar{T}_{\text{daily}}^\text{basic} + \Delta_j, \text{ with } j = 1, 2, \cdots, M \hspace{1cm} (19)$$

where $\Delta_j$ is the temperature increment of the $j$-th day, and $\bar{T}_{\text{daily}}^\text{basic}$ is the daily mean value of the basic temperature curve $T_{\text{in}}^\text{basic}(t)$. They can be calculated respectively by:
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3119295, IEEE Access

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

\[ T_{\text{basic}}^{\text{daily}} = \frac{1}{t_{\text{end}} - t_{\text{start}}} \int_{t_{\text{start}}}^{t_{\text{end}}} [T_{\text{in}}^{\text{basic}}(t)] \, dt \]  
\[ \Delta_j = \frac{M \cdot T_{\text{in}}^{\text{basic}, j}}{M \cdot T_{\text{in}}^{\text{basic}, \text{daily}}} \sum_{j=1}^{N} T_{\text{in}}^{\text{basic}, j} \left( T_{\text{max}}^{\text{daily}} - T_{\text{in}}^{\text{basic}, j} \right) \text{ with } i, j = 1, 2, \ldots, N \]  

where \( T_{\text{max}}^{\text{daily}} \) is the maximum value of the mean daily temperature inside the greenhouse. Then, for a given mean stage temperature, the accumulated temperature increment of the \( j \)-th day of the \( i \)-th development stage can be evaluated by \( \Delta ST_j = (t_{\text{end}} - t_{\text{start}})(T_{\text{in}}^{\text{daily}} - T_{\text{in}}^{\text{basic}, j}) \). To generate the temperature setpoint candidate, the following two cases are considered:

**Case 1:** the maximum value on the basic temperature curve is higher than the optimal temperature, i.e., \( T_{\text{in}}^{\text{basic}, \text{max}} > T_{\text{opt}}^{\text{in}} \)

In this case, it must first determine the times \( t_1 \) and \( t_2 \).

Since the basic temperature between \( t_1 \) and \( t_2 \) is higher than the optimal temperature, the heating action is usually not carried out. Therefore, if \( \Delta ST_j > 0 \), the greenhouse heating action should be carried out in the time intervals \([t_1, t_2] \) and \([t_2, t_{\text{end}}] \). For simplification, the morning and afternoon are allocated the same accumulated temperature, i.e., \( \Delta ST_j^{\text{morning}} = \Delta ST_j^{\text{afternoon}} \), and it must satisfy the condition \( \Delta ST_j^{\text{morning}} + \Delta ST_j^{\text{afternoon}} = \Delta ST_j \). Next, it must determine the starting time \( t_{\text{start}} \) and the end time \( t_{\text{end}} \) of the heating, as shown Fig. 7(a). Thus, the temperature setpoint curve between \( t_{\text{start}} \) and \( t_{\text{end}} \) can be approximated by two parabolic splines:

\[ T_{\text{opt}}^{\text{point}}(t) = \begin{cases} 
T_{\text{in}}^{\text{point}} & 
0.5(t_2 + t_{\text{start}}) \leq t \leq t_{\text{end}} \\
2[T_{\text{in}}^{\text{point}} - T_{\text{in}}^{\text{basic}, (t_{\text{start}})}](t - t_{\text{start}})^2 & 
t_{\text{start}} \leq t \leq 0.5(t_2 + t_{\text{start}}) \\
2[T_{\text{in}}^{\text{point}} - T_{\text{in}}^{\text{basic}, (t_{\text{end}})}](t - t_{\text{end}})^2 & 
t_{\text{end}} \leq t \leq 0.5(t_2 + t_{\text{end}}) \\
\end{cases} \]  

and it has the following constraint:

\[ \int_{t_{\text{start}}}^{t_{\text{end}}} T_{\text{in}}^{\text{point}}(t) \, dt = \Delta ST_j^{\text{morning}} + \Delta ST_j^{\text{afternoon}} \]  

**Case 2:** the maximum value of the basic temperature is lower than the optimal temperature, i.e., \( T_{\text{in}}^{\text{basic}, \text{max}} \leq T_{\text{opt}}^{\text{in}} \)

In this case, it must find the maximum basic temperature \( T_{\text{in}}^{\text{basic}, \text{max}} \) of the day. Similar to Case 1, it requires to determine the starting time \( t_{\text{start}} \) and the stop time \( t_{\text{end}} \) of the heating, and it can use parabolic spline to approximate the temperature setpoint curves in the time intervals \([t_{\text{start}}, t_{\text{end}}]\).
and \([t_{max}, t_{h2}]\), as shown in Fig. 7(b). Then, the temperature setpoint curves in both time intervals can be described respectively by:

\[
T_{out}^{point}(t) = \begin{cases} 
T_{\text{basic}}^{out}(t_{h1}) + 2(T_{in}^{opt} - T_{\text{basic}}^{out}(t_{h1}))(t - t_{h1})^2 & t_{h1} \leq t \leq 0.5(t_{max} + t_{h1}) \\
T_{\text{in}}^{opt} - 2(T_{in}^{opt} - T_{\text{basic}}^{out}(t_{h1}))(t - t_{max})^2 & 0.5(t_{max} + t_{h1}) \leq t \leq t_{max} 
\end{cases}
\]  

(22)

with the constraints

\[
\int_{t_{h1}}^{t_{max}} T_{in}^{point}(t)dt = \Delta ST_{\text{jorning}} + \int_{t_{h1}}^{t_{max}} T_{\text{basic}}^{out}(t)dt
\]

(23)

\[
\int_{t_{h1}+0.5}^{t_{max}} T_{out}^{point}(t)dt = \Delta ST_{\text{afternoon}} + \int_{t_{h1}+0.5}^{t_{max}} T_{\text{basic}}^{in}(t)dt
\]

(24)

\[
\int_{t_{h1}+0.5}^{t_{max}} T_{out}^{point}(t)dt = \Delta ST_{\text{afternoon}} + \int_{t_{h1}+0.5}^{t_{max}} T_{\text{basic}}^{in}(t)dt
\]

(25)

Once obtaining the time response trajectories of the inside climate and weather, we can input them into the greenhouse climate model and crop growth model to evaluate the final energy consumption and crop yield. In this work, the reduced state variable TOMGRO tomato growth model [27] is used to simulate the crop growth, and the energy consumption model developed in our previous work is used to estimate the energy consumption of the greenhouse production [28]. Then, both objectives can be described respectively as follows:

\[
\text{Yield} = \text{TOMGRO}(I_{\text{glob}}(t), T_{\text{in}}^{out}(t), CO_{\text{2in}}(t))
\]

\[
= \text{obj}_1(x)
\]

\[
\text{Energy} = \text{Energy Consumption}(w(t), \text{climate}(t))
\]

\[
= \text{obj}_2(x)
\]

(26)

(27)

where \(w(t) = [I_{\text{glob}}(t), T_{\text{out}}(t), H_{\text{out}}(t), CO_{\text{2out}}(t), V_{\text{out}}(t)]\) is the weather variable vector, and \(\text{climate}(t) = [T_{\text{in}}(t), H_{\text{in}}(t), CO_{\text{2in}}(t)]\) is the microclimate variable inside greenhouse, \(x\) denotes the design variable vector, which represents the mean stage temperatures, and \(\text{obj}_1(\cdot)\) and \(\text{obj}_2(\cdot)\) are the objective functions. Since both objectives are usually computationally expensive, a number of evaluations of both objectives are difficult to ensure the real-time performance of the optimization result. Therefore, they must be approximated by the computationally cheap surrogates in the optimization process. To minimize the energy consumption and maximize the crop yield, this work proposed a RBF surrogate based multi-objective optimization algorithm, which is presented in the supplementary material. Since there are many solutions in the obtained optimal Pareto Set, it must select a best solution to generate the setpoint of the greenhouse climate. Denote \(p_{\text{energy}}\) and \(p_{\text{crop}}\) as the prices of the fossil fuel and crop fruit, respectively, then the best solution can be selected from the optimal Pareto set by maximizing the financial return, i.e.,

\[
x_{\text{best}} = \arg \max_{x} p_{\text{crop}} \cdot \text{Yield}(x) + p_{\text{energy}} \cdot \text{Energy}(x)
\]

(28)

where \(\Omega\) is the design space.

V. EXPERIMENTAL SETTING

Although the proposed greenhouse climate setpoint optimization algorithm has not been applied in practical greenhouses, the weather and greenhouse microclimate data used in this simulation are the real data. These data were collected in a Venlo-type commercial greenhouse located at Chongming, Shanghai. This commercial greenhouse is equipped with many environmental regulation devices such as pipe heating system, direct air heating system and fogging system. All the actuators in the greenhouse are controlled by Priva system, which is an integrated horticultural management and control system developed by the Priva company in Netherlands. Since the Priva system is a closed source system, the proposed method cannot be validated by this system. But to some extent, the simulation experiment still can validate the effectiveness of the proposed approach by using the real weather data.

In practical greenhouse climate control process, the crop growth states such as the LAI and plant height can be real-time measured, but in the simulation, they must be generated by the crop growth model. Therefore, generally, an accurate system model, such as Vanthoor’s model [26] and TOMGRO [29], is useful to improve the optimization performance. However, an accurate models are usually complex and computational expensive, a simple greenhouse climate model with 3 state variables, which was developed in our previous work [30], was used to simulate the greenhouse climate in this work, and a reduced state TOMGRO developed by Jones [36] was used to simulate the crop growth. It should be pointed out that although the greenhouse system model is assumed to be accurate, the true energy consumption and crop yield may be still different from their predicted values, because the control system of the greenhouse climate cannot always ensure the controlled greenhouse climate to accurately track the setpoints. Therefore, online updating the crop states in the dynamic optimization process can be regarded as a kind of feedback, which can ensure the global optimization performance of the obtained setpoints.

In this simulation, the weather data of two growth cycles were used for the online multi-objective optimization. The
weather collected from Step. 1, 2014 to May 18, 2015 is considered as the historical weather, and the weather collected from Step. 1, 2014 to May 17, 2016 is regarded as the current weather forecast. The online surrogate assisted multi-objective optimization is carried out every day. Assume that the crop is transplanted on Step. 1, 2015, and the harvest time is set as May 17, 2016. The growth cycle has 260 days, and the overall growth cycle can be divided into 26 development stages, each of duration ten days. Some model parameters of TOMGRO and the initial crop states are listed in Table 1.

Table 1 model parameters of TOMGRO and initial crop states

| Symbol | Description                  | value | unit |
|--------|------------------------------|-------|------|
| LAI    | leaf area index              | 0.1   | -    |
| DMCrop | total crop dry weight        | 0.28  | g/m² |
| DMFruit| fruit dry weight             | 0     | g/m² |
| Nnode  | number of nodes              | 6     |      |
| ρ      | Plant density                |       |      |
| NFF    | Nodes per plant when first fruit appears | | |
| LAImax | Maximum leaf area index      |       |      |

VI. RESULTS

A. VALIDATION OF WEATHER INTERPOLATION

Two validation criteria including the average absolute error (AE) and the root mean square error (RMSE) are used to validate the weather interpolation method. They are given respectively by

\[
AE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| 
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} 
\]

Since the times that the minimum wind speed appears \( t_{\text{min}1} \) and \( t_{\text{min}2} \) are difficult to predict in practice, their empirical values can be used for the wind speed interpolation. Both parameters can usually be fixed as \( t_{\text{min}1} = 4:00 \) and \( t_{\text{min}2} = 20:00 \), respectively. The interpolation results are illustrated in Fig. 8 and Table 2.

Table 2 validation indices of weather interpolation

| variables | AE    | RMSE  |
|-----------|-------|-------|
| Tout      | 0.846591748754285 | 1.25056742372483 |
| Hout      | 1.30531362634531  | 1.67710330279087 |
| Iglob     | 27.1294226473084  | 58.3481738654183 |
| Vwind     | 1.07558382210378  | 1.35709798466276 |

From Fig. 8 and Table 2, we can observe that the
predicted outside temperature, solar radiation and wind speed can approximate their measured values well, but the predictive error of humidity is large, because the used dew point temperature is a fixed value. In fact, the dew point temperature greatly depends on the weather condition, and is usually difficult to measure. Therefore, the dew point temperature is a time-varying parameter rather than a fixed value, and has great uncertainty. Due to these reasons, the predicted humidity usually has large simulation error. As shown in Fig. 8(d), the predicted wind speed can reflect the basic changing trend, but in many cases, the simulation error is large, because in fact, the wind speed prefers to be a random variable rather than a deterministic variable. Therefore, Eq. (13) can be viewed as a kind of trend function of the wind speed.

Beside the microclimate inside the greenhouse, the total energy consumption and crop yield also depend on the outside temperature, CO2 concentration, solar radiation and wind speed. Therefore, although the predicted humidity outside the greenhouse has large simulation error, it does not significantly impact both objectives. But the predicted humidity greatly impacts the inside humidity due to the ventilation. The interpolation results of the predicted weather indicate that the proposed interpolation method is effective, and can achieve good interpolation performance.

**B. MULTI-OBJECTIVE OPTIMIZATION OF ENERGY CONSUMPTION AND CROP YIELD**

Assume that the fuel oil price and the tomato selling price are $1/L and $2.2/kg, respectively. The surrogate assisted multi-objective optimization is carried out at 12:00 p.m. every day to find a best solution to maximize the financial return. The parameters of the proposed surrogate assisted multi-objective optimization algorithm are set as follows:

(a) Latin hypercubic sampling method is used to sample the initial training data set.
(b) The size of the initial training data set was set as \( n = 11d - 1 \).
(c) The number of the function evaluations after initial sampling is set as 300.

To validate the optimization of the proposed receding horizon multi-objective optimization, it is compared with the multi-layer optimization method proposed in our previous work [37]. In addition, to validate the surrogate assisted multi-objective optimization algorithm on the greenhouse climate setpoint optimization problem, it is compared with EIMEGO and K-RVEA [38].

The initial states of the plant, temperature, humidity and CO2 concentration inside the greenhouse and structural parameter of the greenhouse are set as follows:

(a) Leaf area index (LAI): 0.1
(b) Dry matter of the plant: 0.03 (g/m²)
(c) Number of nodes: 6
(d) Dry matter of the fruit: 0 (g/m²)
(e) The temperature, humidity and CO2 concentration inside the greenhouse are 20.3 °C, 24 g/m³, 376 mg/m³.
(f) Greenhouse floor area: 100 m²
(g) Greenhouse height: 5 m
(h) The setpoint of the CO2 concentration inside the greenhouse is fixed as 414 mg/m³.

For better investigation, we used NSGA-II [32] to offline optimize both objectives based on the historical weather (i.e., from Step 1, 2014 to May 18, 2015) and the current weather (i.e., from Step 1, 2015 to May 17, 2016). The result is shown in Fig. 9. From Fig. 9, it can be seen that the historical weather is better for the crop growth than the current weather. Therefore, the best solution must move from the offline optimal pareto front obtained with the historical weather towards the off-line optimal pareto front obtained with the current weather, which means that both fronts can be regarded as the upper and lower boundaries of the fronts obtained by the online surrogate assisted multi-objective optimizations, as illustrated in Fig. 9. However, some solutions obtained by the online optimization may be outside of the boundaries, because the combination of the historical weather and the current weather may be worse for the crop growth than the historical weather and the current weather.
because for the optimization problem with 26 design variables, the 585 function evaluations are not sufficient to ensure the surrogates to approximate the true objectives. In fact, the surrogate assisted optimization with high dimensional design space is a great challenge in the optimization community. If there are only a few design variables, the convergence performance of the proposed algorithm is good, for example, the Pareto optimal solutions obtained on the last day converge to the upper boundary, as shown in Fig. 9. Due to the insufficient convergence of the surrogate assisted optimization, some Pareto points may be outside the boundaries. Actually, some Pareto points, which are outside of the boundaries, are non-optimal solutions.

For clear investigation, we used NSGA-II to off-line optimize both objectives based on the combination of the historical weather and the current weather, and the results are illustrated in Fig. 10. From Fig. 10, we can observe that the Pareto front online obtained becomes narrower and narrower with the reduction of the number of the design variables, and on the last day, the Pareto front almost becomes a single point. The final best solution is very close to the front off-line obtained under the current weather, which means online optimization can deal with the great uncertainty of the long-term weather, and can ensure the global optimization performance.

In Shanghai, the weather is usually warm before November, and the weather conditions during this period in different years are similar. Generally, it does not need to heat the greenhouse during this period, so the Pareto fronts obtained before November are similar, as shown in Fig. 9. After November, the weather becomes cold, and the weather is very different from the historical weather. Then, the Pareto front obtained by the surrogate assisted optimization is gradually far from the front obtained on the first day. Due to the great uncertainty of the weather, the total energy consumption and final crop yields optimized on every day are different, but in ideal case, the final best solution should converge to the Pareto front obtained by the off-line optimization, as illustrated in Fig. 9. Since the proposed method uses the historical weather and current predicted weather to estimate the crop yield and the total energy consumption on every day, the different combinations of the historical weather and the current predicted weather may result in different Pareto front. Therefore, to investigate the effect of the uncertainty of the current weather on the solution with maximum income, we recorded the predicted crop yield, total energy consumption and maximum income of each day, as shown in Fig. 11. Fig. 11 reflects the fact that in the early stages, the predicted crop yield, total energy consumption and maximum income are similar. When the predicted weather becomes worse, the predicted values of the three indices also become worse. From Fig. 11 (b) it can be seen that the predicted crop yield begin to increase since February, because during this period the solar radiation begin to get better. On the other hand, the predicted total energy consumption almost maintains the same, which means that it can improve the predicted maximum income. Since the different combinations of the historical weather and current predicted weather result in different yield and energy consumption, the predicted values of both objectives always perturbs within a certain range, as shown in Fig. 11.

To illustrate the advantage of the obtained optimal setpoint of the greenhouse climate, we made a comparison with the real greenhouse climate controlled by the PRIVA system, and recorded the accumulation of the fruit dry matter and the energy consumption of each day. The results are illustrated in Table 1. From Table 1 it can be seen that the predicted values always perform worse than the real values. After the first week, the predicted values of solar radiation almost maintain the same, which means that the predicted values can maintain the same if the solar radiation is sufficient. If the solar radiation is limited, the predicted values will perform worse than the real values.
are plotted in Fig. 12. From Fig. 12, it can be seen that the optimal setpoint obtained by the online optimization result in higher energy consumption and crop yield, and can achieve the financial return $21.87 per square meter, while the real greenhouse climate can only achieve $15.98 per square meter. It is clear that comparing with the real
greenhouse climate, the setpoint optimization can improve the financial return by 36.82%. However, from Fig. 12, we also see that the multi-layer optimization can achieve $22.5 per square meter, which is better than that of the receding horizon multi-objective optimization method. There are two reasons. One is that the multi-layer optimization method uses the accurate short-term weather forecast to drive the greenhouse climate, while the receding horizon multi-objective optimization method uses the simulation weather curve; the second is that due to the high-dimensional design space, the surrogate assisted multi-objective algorithm may not completely converge to the optimal pareto front, so the global optimization result is worse than the globally offline optimization result obtained by NSGA-II. But it does not mean that the receding horizon surrogate assisted multi-objective optimization method is worse than the multi-layer optimization, because multi-layer optimization requires that short-term weather can be accurately predicted, and the historical long-term weather is close to the current long-term weather. However, the short-term weather forecast is usually rough, we often can predict the mean values, upper and lower values of the weather instead of the hourly or minutely weather. Therefore, when the weather has great uncertainty, the online global optimization may achieve better result.

To observe the optimization performance of different surrogate assisted multi-objective optimization algorithms, a comparison study was made. The proposed surrogate assisted multi-objective optimization method is compared with K-RVEA and EIMEGO on the greenhouse climate setpoint optimization problem. Fig. 13 illustrates the results obtained by the three algorithms. From Fig. 13, it can be seen that the energy consumption and crop yield obtained by the proposed algorithm are higher than that of K-RVEA and EIMEGO, and the optimization results of K-RVEA and EIMEGO are very close. The financial returns obtained by the later two algorithms are $21.72 and $21.77 per square meter. The economic performances of the three algorithms are similar. The main reason is that K-RVEA and EIMEGO still run well during the later development stages due to the reduction of the number of the design variables, and the crop growth is robust to the environment. This comparison result also reflect the fact that multi-objective optimization can find more control strategies of the greenhouse climate control to achieve the best production efficiency.

VII. CONCLUSION

Good greenhouse climate setpoint not only can improve the crop yield, but also can greatly reduce energy consumption. Therefore, the setpoint optimization is an effective way for the energy saving of the greenhouse production without additional investment cost. However, the difficulty is that the long-term weather forecast has great uncertainty, and the long-term weather is very difficult to accurately predict at a small timescale. Therefore, it is difficult to predict the total energy consumption and final crop yield. In this case, we have to combine the historical weather data with the current short-term weather forecast to drive the setpoint optimization. Since the weather forecast can only predicts a little of weather information such as the minimum and maximum values of the daily temperature, the maximum value of the daily solar radiation, such weather information is not sufficient for the calculations of the energy

Fig. 13. Energy consumption and fruit dry weight accumulated every day under the three surrogate assisted multi-objective algorithms
consumption and crop yield. Therefore, it must use a weather interpolation method to transform the weather information into the time response curve of the weather, and this work proposes an effective weather interpolation approach to do this. The interpolation results indicate that the predicted weather can approximate the measured weather. To provide more selections between the energy consumption and crop yield for the farmer, this work proposes a surrogate assisted multi-objective optimization method to online optimize the objectives. The advantage of the online optimization is that it can effectively handle the uncertainty of the short-term weather forecast. In the online optimization process, it can use the latest weather forecast to update the weather data, and the uncertainty of the objectives can be reduced. With the reduction of the uncertainty of the weather, the Pareto front becomes narrower and narrower, and at the end of the growth cycle, the Pareto front almost becomes a single point which is close to the front obtained by the off-line multi-objective optimization, which means the online optimization can find the global optimal solution in spite of the great uncertainty of the weather. Nevertheless, the practical performance of the proposed method needs to be validated and calibrated in greenhouse climate management.

The proposed method is difficult to ensure the global optimal solution may not be reliable. Therefore, improving the optimization performance, which means that the obtained best solution may not be reliable. Therefore, improving the greenhouse climate model and crop growth model is important work.

REFERENCES

[1] van Straten, G., Challa, H., Buwalda, F. “Towards user accepted optimal control of greenhouse climate.” Comput Electron Agric., vol. 26, pp. 221-238, 2000.

[2] Su, Y., Xu, L., & Goodman, ED. “Nearly dynamic programming NN-approximation-based optimal control for greenhouse climate: A simulation study.” Optim Control Appl Meth., vol.39, pp. 638-662, 2018.

[3] Runkle, E., Both, A. “Greenhouse energy conservation strategies. Michigan State University – Extension Bulletin, 2011, 3160, 2011.

[4] Bibbiani, C., Fantozzi, F., & Gargari, C., et al. “Wood Biomass as Sustainable Energy for Greenhouses Heating in Italy.” Agriculture and Agricultural Science Procedia, vol.8, pp.637-645, 2016.

[5] Chua, KJ., Chou, SK, & Yang, WM. “Advances in heat pump systems: a review.” Applied Energy, vol. 87, no.12, pp.3611–24, 2010.

[6] Tchamitchian, M., Montbroussaud, B., Jeannequin, B., & Martin-Cloaure, R. “SERRISTE: Daily Greenhouse Climate Set-Point Determination for Tomatoes.” IFAC Proceedings Volumes, vol. 30, no. 26, pp. 25-30, 1997.

[7] Sigurin, N., Ferentinos, K.P., Arvanitis, K.G., & Anastasiou, A. “A Comparison of Optimal Greenhouse Heating Setpoint Generation Algorithms for Energy Conservation.” Preprints of the 2nd IFAC-CIGR Workshop on Intelligent Control for Agricultural Applications, IFAC Proceedings, vol. 3, no.11, pp.61-66, 2001.

[8] Van Henten, E.J., & Bontsema, J. “Time-scale decomposition of an optimal control problem in greenhouse climate management.” Control Engineering Practice, vol. 17, no.1, pp.88-96, 2009.

[9] Tap, R.F. “Economics-based optimal control of greenhouse tomato crop production.” PhD Thesis. Wageningen Agricultural University, Wageningen, The Netherlands, 2000.

[10] Gonzalez, R., Rodriguez, F., Guzman, J.L., Berenguel, M. “Robust constrained economic receding horizon control applied to the two-time scale dynamics problem of a greenhouse.” Optim. Contr. Appl. Meth., vol. 35, no. 4, pp. 435-453, 2014.

[11] Xu, D., Da, S., & van Willigenburg, L.G. “Adaptive two-time-scale receding horizon optimal control for greenhouse lettuce cultivation.” Computers and Electronics in Agriculture, vol.146, pp.93–103, 2018.

[12] van Straten, G., van Willigenburg, L.G., & Tap, R.F. “The significance of the crop co-states for receding horizon optimal control of greenhouse climate.” Control engineering practice, vol.10, pp. 625-632, 2002.

[13] Regis, RG. “Constrained optimization by radial basis function interpolation for high-dimensional expensive black-box problems with infeasible initial points.” Engineering Optimization, vol. 46, no.2, pp. 218–243, 2014.

[14] Knowles, J. “ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems.” IEEE Trans. Evol. Comput., vol. 10, no.1, pp. 50–66, 2006.

[15] Zhang, Q., Liu, W., Tsang, E., & Virginas, B. “Expensive multiobjective optimization by MOEA/D with Gaussian process model.” IEEE Trans. Evol. Comput., vol. 14, no.3, pp. 456–474, 2010.

[16] Zhan, D., Cheng, Y. & Liu, J. “Expected Improvement Matrix-Based Infill Criteria for Expensive Multiobjective Optimization.” IEEE Trans. Evol. Comput., vol. 21, no. 6, pp.956-975, 2017.

[17] Reicosky, D.C., Winkelman, L.J., Baker, J.M. & Baker, D.G. “Accuracy of Hourly Air Temperatures Calculated From Daily Minima and Maxima.” Agricultural and Forest Meteorology, vol. 46, pp. 193-209, 1989.

[18] Chow, D.H., & Levermore, G. J. “New algorithm for generating hourly temperature values using daily maximum, minimum and average values from climate models.” Building Serv. Eng. Res. Technol., vol.28, no.3, pp.237-248, 2007.

[19] Eccel, E. “What we can ask to hourly temperature recording. Part II: hourly interpolation of temperatures for climatology and modelling.” Italian Journal of Agrometeorology, vol.15, no.2, pp.45-50, 2010.

[20] William, C., Forsythe, E. J., Rykiel, Jr., & Randal, S., et al. “A model comparison for daylength as a function of latitude and day of year.” Ecological Modelling, vol. 80, pp. 87-95, 1995.

[21] Ma, Y., & Guttrop, P. “Estimating daily mean temperature from synoptic climate observations.” International Journal of Climatology, vol. 33, pp.1264–1269, 2013.

[22] Wong, L.T., Chow, W.K. “Solar radiation model.” Applied Energy, vol. 69, pp.191–224, 2001.

[23] Machler, MA., & Iqbal, M. “A modification of the ASHRAE clear-sky irradiation model.” ASHRAE Transactions, vol. 91, no.1, pp.06–15, 1985.
[24] Kaplanis, S., & Kaplani, E.. “A model to predict expected mean and stochastic hourly global solar radiation values.” Renewable Energy, vol.32, pp.1414–1425, 2007.

[25] Ephrath, J.E., Goudriaan, J., & Marani, A.. “Modelling diurnal patterns of air temperature, radiation wind speed and relative humidity by equations from daily characteristics.” Agricultural Systems, vol.51, no.4, pp.377-393, 1996.

[26] Vanthoor, B.H.E. A model - based greenhouse design method [PhD thesis]. Wageningen University, Wageningen, The Netherlands, 2011.

[27] Jones, J.W., Kenig, A., & Vallejos, C.E.. Reduced state-variable tomato growth model. Trans ASAE, vol.42, no.1, pp.255-265, 1999.

[28] Shen, Y., Wei, R., & Xu, L.. Energy Consumption Prediction of a Greenhouse and Optimization of Daily Average Temperature. energies, vol.11, no.1, pp.65-81, 2018.

[29] Jones, J. W., Dayan, E., Allen, L. H., Van Keulen, H., & Challa, H.. “A Dynamic Tomato Growth and Yield Model (TOMGRO).” Transactions of the ASAE, vol.34, no.2, pp.663-672, 1991.

[30] Su, Y., & Xu, L.. A Greenhouse Climate Model for Control Design, the 15 IEEE International Conference on Environment and Electrical Engineering, 2015, Rome, Italy.

[31] Jones, P., Jones, J.W., & Hwang, Y.. “Simulation For Determining Greenhouse Temperature Setpoints.” Transactions of the ASAE, vol. 33, no.5, pp.1722-1728, 1990.

[32] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T.. “A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans.” Evolutionary Computation., vol. 6, no.2, pp.182–197, 2002.

[33] M. S. Omar et al., "Multiobjective Optimization in 5G Hybrid Networks," IEEE Internet of Things Journal, vol. 5, no. 3, pp. 1588-1597, June 2018

[34] Z. Zhou et al., "Energy-Efficient Resource Allocation for Energy Harvesting-Based Cognitive Machine-to-Machine Communications," IEEE Transactions on Cognitive Communications and Networking, vol. 5, no. 3, pp. 595-607, Sept. 2019.

[35] S. Khan, K. Muhammad, S. Mumtaz, S. W. Baik and V. H. C. de Albuquerque, "Energy-Efficient Deep CNN for Smoke Detection in Foggy IoT Environment," IEEE Internet of Things Journal, vol. 6, no. 6, pp. 9237-9245, Dec. 2019

[36] J. W. Jones, A. Kenig , & C. E. Vallejo. “Reduced state variable tomato growth model,” Transactions of the ASAE, vol.42, no.1 pp.255-265, 1999.

[37] T. Chugh, Y. Jin, K. Miettinen, J. Hakanen, and K. Sindhya. “A Surrogate-Assisted Reference Vector Guided Evolutionary Algorithm for Computationally Expensive Many-Objective Optimization,” IEEE Transactions on Evolutionary Computation, vol. 22, no. 1, pp.129-142, Feb. 2018.

[38] Y. Su, L. Xu, E. D. Goodman. "Multi-layer hierarchical optimisation of greenhouse climate setpoints for energy conservation and improvement of crop yield," biosystems engineering, vol.205, no. 2021, pp. 212-233, May, 2021.

Yuanping Su received the bachelor’s degree in automation in 2002, and the master’s degree in control theory and control engineering in 2009 from the Jiangxi University of Science and Technology, Ganzhou, China, and Ph.D in control theory and control engineering in 2016 from Tongji University, Shanghai, China. He is currently working in Jiangxi University of Science and Technology as an associate professor. His current research interests include nonlinear system control theory, evolutionary optimization, and modeling and control of greenhouse environment.

Lihong Xu received the Ph.D degree in engineering from the Department of Automatic Control, Southeast University, Nanjing, China, in 1991. In 1994, he was appointed as a Professor at Southeast University. He transferred to Tongji University in August, 1997, and has been a Professor with Tongji University since then. His research fields include control theory, computational intelligence, and optimization theory. In 1998 he was funded by the Daylight Project of Shanghai. In 1999 he was funded by the University Backbone Young Tutors Project of Ministry of Education of China. He got a National Science and Technology Progress Award (second prize) of China in 2007. He is now doing joint research work as a Visiting Professor and Advisor of the greenhouse research team of BEACON, USA. Prof. Xu is a member of ACM, a senior member of IEEE and the President of IEEE CIS’s Shanghai Chapter. He was the Co-Chair of the 2009 GEC Summit in Shanghai. He is also a Standing Director of the Chinese Society of Agricultural Engineering and the President of Shanghai Society of Agricultural Engineering.