CHEMICAL ENGINEERING TRANSACTIONS

VOL. 81, 2020

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Petar S. Varbanov, Qiuwang Wang, Min Zeng, Panos Seferlis, Ting Ma, Jiří J. Klemeš

**ISBN** 978-88-95608-79-2; **ISSN** 2283-9216

DOI: 10.3303/CET2081121

# Data-Driven Robust Optimization for Greenhouse Temperature Control Using Model Predictive Control

Wei-Han Chen, Fengqi You*

Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14853 USA
wc593@cornell.edu

This work proposes a novel data-driven robust model predictive control (DDRMPC) framework for automatic control of greenhouse temperature and $CO_2$ concentration level. The essential concept is to combine dynamic models of greenhouse temperature and $CO_2$ concentration level with data-driven models that identify uncertainty in weather forecast error. By leveraging a machine learning approach, support vector clustering with weighted generalized intersection kernel, data-driven uncertainty sets for ambient temperature and solar radiation are constructed from historical weather data. A training-calibration procedure that tunes the size of uncertainty sets is implemented to ensure that data-driven uncertainty sets attain appropriate performance guarantee. In order to solve the optimization problem in DDRMPC, an affine disturbance feedback policy that provides tractable approximations of optimal control is utilized. A case study of controlling temperature and $CO_2$ concentration level in a greenhouse is carried out. The results show that the proposed DDRMPC framework can prevent the greenhouse climate from becoming harmful to plant and fruit. DDRMPC approach ends up with 20 % less total economic cost than rule-based control strategy. The proposed DDRMPC approach also gives better control performance comparing to certainty equivalent MPC and robust MPC.

## 1. Introduction

Controlling greenhouse climate within a suitable range is an important task for ensuring plant and fruit growth (Cheng et al., 2018). Among several different environment conditions that should be considered such as temperature, lighting, $CO_2$ concentration, and humidity, temperature is the most important factor that should be carefully controlled. Regulating temperature in an appropriate range can not only increase fruit production but also can prevent plants from heat stress or cold damage. However, to prevent temperature in controlled environment agriculture from dropping during winter season in northern regions requires heavy energy consumption which leads to expensive total production cost (Ahamed et al., 2019).

Among various approaches, model predictive control (MPC) is an effective strategy that utilizes prediction of disturbances to optimize future system behavior under certain constraints (Garcia et al., 1989). At each time step, the controller solves an optimization problem based on a model that shows the relationship between system states, control inputs, and disturbances. Only the first control input is implemented while the rest is discarded. This process repeats for all time steps to derive control trajectories. MPC is an ideal framework for building control because building dynamics are slow and the system model incorporates disturbances and constraints can be derived from first principles models (Serale et al., 2018). Another advantage of MPC on greenhouse climate control is that a greenhouse can usually be considered as a large room (Chu et al., 2015). In this way, the model is easier to be obtained in contrast to a building with multiple rooms which involve number of system states is large. The benefits of MPC on building control have been demonstrated in comparison to conventional methods such as rule-based control (RBC). Numerous studies on greenhouse climate control that adopt the MPC have been investigated. There are different approaches such as one based on Volterra series model (Gruber et al., 2011) or one solved by particle swarm optimization (Zou et al., 2010).

Despite the various advantages of MPC on greenhouse control, it is impossible to perfectly predict weather, which is the major disturbance in greenhouse control problems. For example, weather prediction of ambient temperature may deviate from the true measurement. The uncertainty of disturbances might cause system

Paper Received: 25/03/2020; Revised: 16/06/2020; Accepted: 17/06/2020
Please cite this article as: Chen W.-H., You F., 2020, Data-Driven Robust Optimization for Greenhouse Temperature Control Using Model Predictive Control, Chemical Engineering Transactions, 81, 721-726 DOI:10.3303/CET2081121

states to violate specified constraints and damage crop production. In order to cope with uncertainties, robust MPC (RMPC) can be adopted (Ning et al., 2019). When uncertainty is bounded, RMPC could ensure that system states would not violate the constraints even when the worst-case scenario occurs. The control inputs may be more expensive to compensate for robustness. RMPC are implemented on greenhouse in some studies (Chen et al., 2018). RMPC may lead to over-conservative results, which is not favorable. Although it is guaranteed that constraints in RMPC would not be violated in the worst-case scenario, the probability of such scenario to happen could be excessively low. To prevent extreme cases, more expensive control inputs are required, which would lead to a waste. In this work, a data-driven RMPC (DDRMPC) framework for greenhouse temperature model is proposed to reduce the conservatism. Firstly, the state-space model of the greenhouse is generated based on building elements construction. Secondly, historical weather forecast data and historical weather measurement data are gathered. These two sets of data could generate uncertain forecast errors. High-density region of the uncertain prediction errors is captured by a machine learning technique, support vector clustering (SVC) (Shang et al., 2019). The historical data information can then be incorporated into RMPC, and conservatism is reduced. To solve the optimization problem in DDRMPC, affine disturbance feedback (ADF) policy is utilized for tractable approximations. The optimization problem is solved by off-the-shelf solvers after the robust counterpart is derived. The contributions of this paper are summarized below:

- A novel data-driven robust model predictive control framework to control greenhouse temperature;
- A simulation of greenhouse temperature control based on real weather data demonstrates better control performance of DDRMPC comparing to other conventional methods.

## 2. Model formulation

### 2.1 Greenhouse model formulation

In greenhouse MPC, a model is required for predicting greenhouse climate (e.g. temperature,) as a function of control inputs (e.g. heating power) and disturbances (e.g. ambient temperature), so that the climate can be constrained in a satisfied range. The cost function and constraints of control inputs are also required to model.

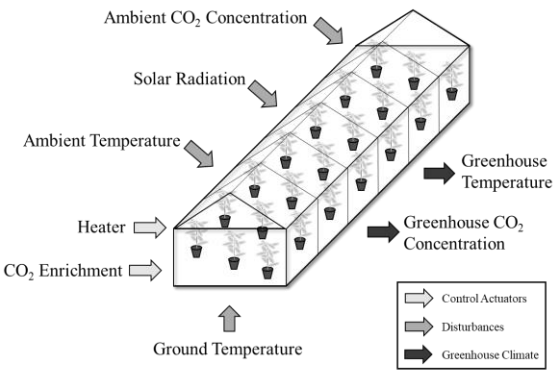


*Figure 1: Greenhouse structure model*

The BRCM MATLAB toolbox uses first-principles models to derive building models which are specially designed for MPC. The main approach of BRCM toolbox is thermal resistance-capacitance modeling. Building elements are served as resistances and capacitances so the model is an analogy to electrical circuit modeling where temperature corresponds to voltage, heat flux to current, thermal capacitance to electrical capacitance, and thermal resistance to electrical resistance. Hence, the model is described by linear ordinary differential equations. To control greenhouse $CO_2$ concentration, a model to calculate net uptake rate of $CO_2$ by crops per unit greenhouse area is required, and it can be estimated by an empirical model of net photosynthesis (Kläring et al., 2007). The dynamic greenhouse climate model in this work combines the temperature model generated by BRCM model. It can be expressed as the following discrete-time linear time-invariant system

$$x_{k+1} = Ax_k + B_u u_k + B_v v_k + B_w w_k \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state, $u_k \in \mathbb{R}^m$ is the control input, and $v_k \in \mathbb{R}^p$ is the weather disturbance at time step *k*, respectively. The matrices *A*, $B_u$, $B_v$, and $B_w$ are of appropriate sizes. As prediction error is unavoidable in weather forecast, $w_k$ is implemented to represent prediction error.

In this work, the states we consider are greenhouse temperature, floor temperature, ceiling temperature, wall temperature, and greenhouse $CO_2$ concentration. The control input is the heating power and $CO_2$ enrichment. The disturbances are solar radiation, ambient temperature, and ground temperature. We assume ground temperature is perfectly known so prediction errors are only for solar radiation and ambient temperature. The structure of greenhouse model including input and disturbances is shown in Figure 1.

### 2.2 Uncertainty Set Formulation

Before constructing an uncertainty set for ambient temperature prediction error $w_{temp}$, pairs of historical forecast data and historical measurement are required (Ning et al., 2019). Samples of temperature prediction errors can then be calculated from $w_{temp} = \tilde{v}_{temp} - \hat{v}_{temp}$ where $\tilde{v}_{temp}$ is the historical measurement for ambient temperature, and $\hat{v}_{temp}$ is the historical forecast. To obtain the uncertainty set from these *N* samples, machine learning methods can be utilized to develop data-driven uncertainty sets (Shang et al., 2019). Notable data-driven uncertainty sets include principal component analysis and kernel smoothing based sets (Ning et al., 2019), dirichlet process mixture model based disjunctive sets (Ning et al., 2016), among others. In this work, we adopt SVC (Ben-Hur et al., 2002), which tries to find the radius of the minimal sphere that can capture data without considering outliers. Weighted generalized intersection kernel (WGIK) is implemented when solving the dual form of SVC optimization problem (Shang et al., 2017). Unlike some common kernels (e.g. radial basis function, polynomial) which would cause a burden when solving robust optimization, WGIK is especially suited for robust optimization due to its linearity. The data-driven uncertainty set is shown as

$$w_{\text{temp}} \in D_{\text{temp}} = \left\{ w_{\text{temp}} \,\middle|\, \sum_{i \in SV} \alpha_i \left\| \mathbf{Q}(w_{\text{temp}} - w_{\text{temp}}{}^{(i)}) \right\|_1 \le \theta \right\} \tag{2}$$

where **Q** is a weighting matrix that can be obtained from the covariance matrix of $w_{temp}$. Model parameters $\{\alpha_i\}$ and uncertainty set parameters $\theta$ are determined after solving the dual form of SVC using WGIK. It is basically the same procedure for constructing solar radiation uncertainty set as temperature uncertainty set. Solar radiation prediction errors are calculated from $w_{sol} = \tilde{v}_{sol} - \hat{v}_{sol}$ where $\tilde{v}_{sol}$ is the historical measurement for solar radiation, and $\hat{v}_{sol}$ is the historical forecast for solar radiation. The SVC-based uncertainty set is then given as,

$$w_{\text{sol}} \in D_{\text{sol}} = \left\{ w_{\text{sol}} \,\middle|\, \sum_{i \in SV} \alpha_i \left\| \mathbf{Q}(w_{\text{sol}} - w_{\text{sol}}{}^{(i)}) \right\|_1 \le \theta \right\} \tag{3}$$

which is in the same form as the SVC-based ambient temperature uncertainty set.

## 3. Control strategies

In this work, different control strategies are simulated to compare their control performances. These are RBC, certainty equivalence MPC (CEMPC), RMPC, and the proposed DDRMPC.

### 3.1 Robust model predictive control

RMPC guarantees constraint satisfaction for the worst case of the bounded disturbances (Bemporad and Morari, 1999). In order to ensure the tractability of the RMPC problem, ADF policy is adopted, and control input $u_t$ is parameterized according to the past disturbances as follows (Goulart et al., 2006)

$$u_t = h_t + \sum_{j=0}^{t-1} M_{t,j} w_j \tag{4}$$

It also can be written in a compact form as

$$\mathbf{u} = \mathbf{h} + \mathbf{M}\mathbf{w} \tag{5}$$

where

$$\mathbf{M} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ M_{1,0} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ M_{H,0} & M_{H,1} & \cdots & 0 \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_H \end{bmatrix} \tag{6}$$

become decision variables that should be solved to determine the control inputs.

We assume the uncertainty set for RMPC to be $L_1$-norm-based as follows

$$\mathbf{w} \in W = \left\{ \mathbf{w} \middle| \|\mathbf{w}\|_1 \le \Omega \right\} \tag{7}$$

where Ω is the budget parameter that can adjust the conservatism. When Ω is larger, the uncertainty set becomes bigger. The RMPC problem can now be solved easily after the ADF policy is adopted.

### 3.2 Data-Driven robust model predictive control

DDRMPC adopts SVC to construct uncertainty sets that could tackle outliers. Furthermore, the performance guarantee is ensured after tuning uncertainty sets by the calibration data set. The approach to solving the optimization problem in DDRMPC also uses ADF policy shown as

$$\begin{aligned} & \min_{\mathbf{M},\mathbf{h}} \mathbf{c}^T \mathbf{h} \\ & \text{s.t. } \mathbf{F}_x \left[ \mathbf{A} x_0 + \mathbf{B}_u \mathbf{h} + \mathbf{B}_v \mathbf{v} + (\mathbf{B}_u \mathbf{M} + \mathbf{B}_w) \mathbf{w} \right] \le \mathbf{f}_x, \forall \mathbf{w} \in D \\ & \quad \mathbf{F}_u \left[ \mathbf{M}\mathbf{w} + \mathbf{h} \right] \le \mathbf{f}_u, \forall \mathbf{w} \in D \end{aligned} \tag{8}$$

where $x_0$ is the initial state, and Eq(8) is a convex optimization problem that can be solved effectively.

## 4. Case study

### 4.1 Problem description

In this work, a greenhouse located at Brooklyn, NY, USA is simulated for closed-loop temperature and $CO_2$ concentration control under different control strategies. The dimension of the greenhouse is 40 m × 13 m × 4 m. The material for roof and walls is 10 mm twin-wall polycarbonate which provides good insulation against heat. The floor is made of concrete. The disturbances considered are solar radiation, ambient temperature, and ground temperature. To control the greenhouse air temperature and $CO_2$ concentration, heater and $CO_2$ enrichment are implemented. Historical weather forecast data and historical weather measurement data from January 2018 to June 2018 are collected from Meteogram Generator (Iowa State University, 2018).
The climate control is for tomatoes to grow in the greenhouse. Therefore, the control goal for daytime and nighttime is treated differently (Mesquita et al., 2019). However, the required temperature setpoints would also change because of different periods of tomato growth. The setpoints are obtained by maximizing the profit of tomato crop. From 6 am to 10 pm, the greenhouse temperature should be between 22-24 ℃ according to different growing period, and in the rest of time, the greenhouse temperature should be above 18 ℃. Ground temperature is assumed to be a constant at 18 ℃. The maximum heating power for heating system is 300,000 W. For RBC, constant heating power is set as 60,000 W, and the threshold to turn on heater is 15 ℃. For CEMPC, RMPC, and DDRMPC, the prediction horizon $H$ is 5 intervals, and the sampling interval is 1 hour. For RMPC, budget parameter Ω ranges from 0 to 8 to reveal different levels of conservatism. For DDRMPC, another set of historical weather data from January to June 2017 is obtained for constructing data-driven uncertainty set. The maximal violation probability and confidence level are set as $\epsilon = 0.05$ and $\beta = 0.10$, leading to $N_{\text{calib}} = 45$ samples.

### 4.2 Results and discussions

The results of different control strategies throughout 60 days are simulated, and Figure 2 is an example that presents the temperature profile of January. The profile has a diurnal cycle, with higher temperature during daytime. Most of the time, DDRMPC leaves some margin from the temperature constraints to ensure weather forecast error to happen. Therefore, due to the margin saved for prediction error, temperature constraint would not be violated, or at most only very minor violation when extreme cases strike. Similar to DDRMPC, RMPC also leaves some margin for the temperature constraints. However, the margin left by RMPC is usually larger than DDRMPC's case, which clearly demonstrates the ability of DDRMPC to reduce conservatism. On the other hand, CEMPC violates the constraints frequently and severely. The reason of the violation is that CEMPC does not consider prediction error when optimization problem is being solved at each step. CEMPC shows the least conservative control profile. Although CEMPC consumes the least heating power shown in Table 2, it ends up with the most constraint violations in Table 1. Assuming prediction as true value makes CEMPC violate constraints. Whenever the weather is colder than predicted, CEMPC would violate the constraint. Some violations are so severe that the greenhouse air temperature drops to 15℃. This is undesirable in a greenhouse because crops are usually sensitive to temperature. Even a 2-3 ℃ difference may cause serious damages to crop. RMPC is tuned according to the budget parameter Ω. When Ω = 0, RMPC is the same as CEMPC which does not consider uncertainty. When Ω = 8, RMPC results in nearly no more constraint violation. To achieve this result, RMPC requires more heating power than CEMPC and DDRMPC. However, DDRMPC also results in nearly zero constraint violation.

Table 3 shows the trade-off between violation and control effort. Control cost comparison of different control strategies are on the same basis of perfect bound (PB) result. It reveals that although CEMPC induces the least control cost, it violates constraints frequently. RBC results in more control cost and constraint violation percentage than RMPC and DDRMPC, which shows that it performs worse than those two approaches. While DDRMPC violates constraints slightly, it uses the least control cost and has the lower violation percentage. The violation percentage only tells how frequent a control strategy violates constraints but does not show how serious the violation is. In the third row, another indicator is added to compare how serious the violation is for different control strategies. Violation amount calculates the area below the violation, and it has similar results to violation percentage, which CEMPC violates a great amount, and RBC performs worse in this case as well. Although DDRMPC violates 0.72 ℃·h in total, it does not affect much given the time horizon of 1,440 h.

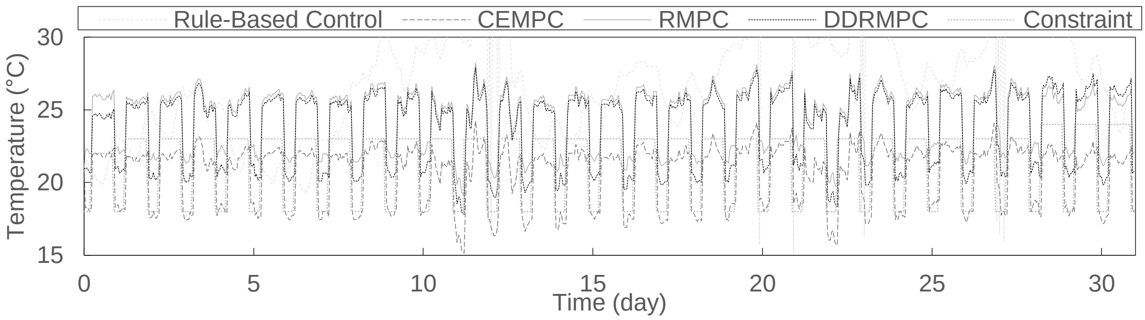


*Figure 2: Greenhouse temperature profile in January 2018*

*Table 1: Constraint violation percentage in each month*

| | Jan. | Feb. | Average. |
|---|---|---|---|
| RBC (%) | 13.84 | 4.76 | 9.30 |
| CEMPC (%) | 82.80 | 81.25 | 82.02 |
| RMPC (%) | 0.00 | 0.89 | 0.45 |
| DDRMPC (%) | 0.67 | 0.00 | 0.34 |

*Table 2: Cost of controlling temperature and $CO_2$ concentration in Each Month*

| | Jan. | Feb. | Total |
|---|---|---|---|
| PB (USD) | 5,939 | 4,177 | 10,115 |
| RBC (USD) | 8,808 | 7,968 | 16,776 |
| CEMPC (USD) | 5,887 | 4,197 | 10,084 |
| RMPC (USD) | 7,705 | 5,892 | 13,597 |
| DDRMPC (USD) | 7,423 | 5,886 | 13,310 |

*Table 3: Trade-off between control effort and violation indicators*

| | RBC | CEMPC | RMPC | DDRMPC |
|---|---|---|---|---|
| Control cost* (%) | 66 | 0 | 34 | 32 |
| Violation percentage (%) | 9.3 | 82 | 0.45 | 0.34 |
| Violation amount (℃·h) | 193 | 1,631 | 2.94 | 0.72 |

* Additional control cost in % of PB

## 5. Conclusions

In this paper, a data-driven robust model predictive control framework for greenhouse climate control was developed. To prevent greenhouse climate from becoming harmful to plant and fruit due to inherent error lying in weather forecast, uncertainty sets for temperature and solar radiation were constructed by adopting the SVC approach on historical weather data. A case study of controlling simultaneously the temperature and $CO_2$

concentration of a greenhouse is demonstrated. DDRMPC reduces 21 % and 2.1 % controlling cost comparing to RBC and RMPC, while only 0.34 % of time violates temperature constraint throughout 1,440 h. The results showed that DDRMPC had better control performance compared to RBC, CEMPC, and RMPC.

## References


Ahamed M. S., Guo H., Taylor L., Tanino K., 2019, Heating demand and economic feasibility analysis for year-round production in Canadian Prairies greenhouses, Information Processing in Agriculture, 6, 81-90.

Bemporad A., Morari M., 1999, Robust MPC, Robustness in Identification and Control, 245, 207.

Ben-Hur A., Horn D., Siegelmann H. T., Vapnik V., 2002, Support vector clustering, Journal of Machine Learning Research, 2(2), 125-137.

Chen L., Du S., He Y., Liang M., Xu D., 2018, Robust model predictive control for greenhouse temperature based on particle swarm optimization, Information Processing in Agriculture, 5(3), 329-338.

Cheng F., Jin H., Shen H., 2018, Design of real-time monitoring system for $CO_2$, $CH_4$ and various environmental factors in intelligent greenhouse, Chemical Engineering Transactions, 71, 163-168.

Chu Y., You F., 2015, Model-based integration of control and operations: Overview, challenges, advances, and opportunities, Computers & Chemical Engineering, 83, 2-20.

Garcia C. E., Prett D., Morari M., 1989, Model predictive control: theory and practice, Automatica, 25, 335-348.

Goulart P. J., Kerrigan E. C., Maciejowski J. A., 2006, Optimization over state feedback policies for robust control with constraints, Automatica, 42(4), 523-533.

Gruber J. K., Guzman J. L., Rodriguez F., Bordons C., Berenguel M., Sanchez J. A., 2011, Nonlinear MPC based on a Volterra series model for greenhouse temperature control using natural ventilation, Control Engineering Practice, 19(4), 354-366.

Iowa State University, 2018, Meteogram generator, Iowa State University <www.meteor.iastate.edu/~ckarsten/bufkit/image_loader.phtml> accessed 01.01.2020.

Kläring H. P., Hauschild C., Heißner A., Bar-Yosef B., 2007, Model-based control of $CO_2$ concentration in greenhouses at ambient levels increases cucumber yield, Agricultural and Forest Meteorology, 143, 208.

Mesquita M., Dos Santos De Paula A., Pureza Machado A., Fonseca De Oliveira H., Casaroli D., Alves Junior J., 2019, Qualitative characteristics of processing tomato cultivated under water deficit induced in the vegetative growth stage, Chemical Engineering Transactions, 75, 175-180.

Ning C., You F., 2017, Data-driven adaptive nested robust optimization: general modeling framework and efficient computational algorithm for decision making under uncertainty, AIChE Journal, 63, 3790-3817.

Ning C., You F., 2017, A data-driven multistage adaptive robust optimization framework for planning and scheduling under uncertainty, AIChE Journal, 63, 4343-4369.

Ning C., You F., 2018, Data-driven stochastic robust optimization: General computational framework and algorithm leveraging machine learning for optimization under uncertainty in the big data era, Computers & Chemical Engineering, 111, 115-133.

Ning C., You F., 2018, Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods, Computers & Chemical Engineering, 112, 190-210.

Ning C., You F., 2019, Optimization under uncertainty in the era of big data and deep learning: When machine learning meets mathematical programming, Computers & Chemical Engineering, 125, 434-448.

Ning C., You F., 2019, Data-driven adaptive robust unit commitment under wind power uncertainty: a bayesian nonparametric approach, IEEE Transactions on Power Systems, 34, 2409-2418.

Serale G., Fiorentini M., Capozzoli A., Bernardini D., Bemporad A., 2018, Model predictive control (MPC) for enhancing building and HVAC system energy efficiency, Energies, 11(3), 631.

Shang C., Chen W.-H., Stroock A.D., You F., 2020, Robust model predictive control of irrigation systems with active uncertainty learning and data analytics, IEEE Transactions on Control Systems Technology, 28, 1493-1504.

Shang C., Huang X., You F., 2017, Data-driven robust optimization based on kernel learning, Computers & Chemical Engineering, 106, 464-479.

Shang C., You F., 2019, A data-driven robust optimization approach to scenario-based stochastic model predictive control, Journal of Process Control, 75, 24-39.

Shang C., You F., 2019, Data analytics and machine learning for smart process manufacturing: recent advances and perspectives in the big data era, Engineering, 5, 1010-1016.

Zou Q., Ji J., Zhang S., Shi M., Luo Y., 2010, Model predictive control based on particle swarm optimization of greenhouse climate for saving energy consumption, 2010 World Automation Congress, 123-128.