Computers and Electronics in Agriculture 148 (2018) 280–290

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

A knowledge-and-data-driven modeling approach for simulating plant growth and the dynamics of CO_2/O_2 concentrations in a closed system of plants and humans by integrating mechanistic and empirical models

Xing-Rong Fan[a,b], Xiujuan Wang[b,f], Mengzhen Kang[b,c,*], Jing Hua[b,c], Shuangsheng Guo[d,*], Philippe de Reffye[e], Bao-Gang Hu[f]

[a] Electronic Information and Networking Research Institute, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[b] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[c] Qingdao Academy of Intelligent Industries, Qingdao 266109, China
[d] National Key Laboratory of Human Factors Engineering, China Astronaut Research and Training Center, Beijing 100094, China
[e] Cirad-Amis, UMR AMAP, TA 4001 Avenue Agropolis, F-34398 Montpellier, Cedex 5, France
[f] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[f] Beijing Engineering Research Center of Intelligent Systems and Technology, Beijing, 100190, China

ARTICLE INFO

Keywords:
Closed ecological life support system
Functional-structural plant model
Knowledge-and-data-driven model
Mass balance model
Model evaluation
And model verification

ABSTRACT

Modeling and the prediction of material flows (plant production, CO_2/O_2 concentrations, H_2O) is an important but challenging task in the design and control of closed ecological life support systems (CELSS). The aim of this study was to develop a novel knowledge-and-data-driven modeling (KDDM) approach for simultaneously simulating plant production and CO_2/O_2 concentrations in a closed system of plants and humans by integrating mechanistic and empirical models.

The KDDM approach consists of a 'knowledge-driven (KD)' sub-model and a 'data-driven (DD)' sub-model. The KD sub-model describes hourly and up to daily plant photosynthesis, respiration and assimilation partitioning using the components of GreenLab and TomSim models. The DD sub-model describes the dynamics of CO_2 production and O_2 consumption by the crew member using a piecewise linear model. The two sub-models were integrated with a mass balance model for CO_2/O_2 concentrations in a closed system.

The KDDM was applied with a two-person, 30-day integrated CELSS test. This model provides accurate computation of both the dry weights of different plant compartments and CO_2/O_2 concentrations. The model also quantifies the underlying material flows among the crew members, plants and environment.

This approach provides a computational basis for lifetime optimization of cabin design and experimental setup of CELSS (*e.g.*, environmental control, planting schedule). With extension, this methodology can be applied to a half-closed system such as a glasshouse.

1. Introduction

Closed Ecological Life Support Systems (CELSS) are self-supporting life support systems for space stations and colonies, typically using controlled closed ecological systems. To date, CELSS have been widely acknowledged as playing a vital role in future regenerative life support systems for long-term human deep space exploration, space technology development, and space colonization (Guo et al., 2014a; Wheeler and Sager, 2006). These systems can provide basic life-support requirements for crew members, such as food, oxygen and drinking water, using plants as the central recycling component. Therefore, research programs on CELSS have been implemented at the national space agencies and universities, such as the University of Arizona (Biosphere 2, USA), the Institute of Biophysics in Krasnoyarsk (BIOS-3, Russia), Beijing University of Aeronautics and Astronautics (Yuegong-1, China), and the European Space Agency (MELiSSA). One of the most important elements of CELSS is the growth of higher plants in a controlled environment for the production of food and oxygen (O_2) from 'waste' carbon dioxide (CO_2) (Finetto et al., 2008; Guo et al., 2008; Hezard et al., 2012;Wheeler, 2015).

Since the experiments of CELSS are high-cost and time-consuming, a mass-balance model for life support systems needs to be developed in at

---

[*] Corresponding authors at: The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (M.Z. Kang), and National Key Laboratory of Human Factors Engineering, China Astronaut Research and Training Center, Beijing 100094, China (S.S. Guo).

E-mail addresses: mengzhen.kang@ia.ac.cn (M. Kang), guoshuangsheng@sina.com (S. Guo).

https://doi.org/10.1016/j.compag.2018.03.006
Received 23 May 2017; Received in revised form 28 February 2018; Accepted 4 March 2018
0168-1699/

least two dimensions: firstly, it must predict important fluxes (*e.g.*, edible biomass, CO_2/O_2 concentrations), and secondly, it must provide environmental control of the plant and human compartments. As plants are complex and dynamic systems, their growth and development involves a large number of interconnected ecophysiological processes. Significant progress has been reported in studies of modeling, simulation and visualization of plant growth in recent decades (Diao et al., 2012; de Reffye and Hu, 2003; Fan et al., 2015; Vos et al., 2009; Yin and Struik, 2016). Early process-based models (PBMs) consider the environment as the main variable driving plant growth and focus on plant functioning in relation to environmental conditions, such as TomSim (Heuvelink, 1995, 1999). Typically, PBMs include modeling of growth mechanisms (*e.g.*, leaf and crop photosynthesis, light interception, maintenance respiration, biomass production) and the interactions between plants and environmental conditions (*e.g.*, temperature, light, CO_2). A relatively weak component of PBMs is the allocation of assimilates among different organs (leaves, internodes and fruits), which limit their potential application in various environmental scenarios.

More recently, a new generation of plant models, often known as functional-structural plant models (FSPMs), has emerged, which aim to explicitly describe the topology and spatial geometry of plant structure, the interactions among plant structural elements (*e.g.*, shape and orientation of organs), the function of organs (*e.g.*, leaf photosynthesis), the allocation of assimilates among organs, and the feedback between plant growth and development (Vos et al., 2007). To date, FSPMs have been regarded as potential tools for predicting and simulating plant growth and structural development (Renton, 2013), such as with the GreenLab model (de Reffye and Hu, 2003). GreenLab is a generic and mechanistic functional-structural plant model that was developed to simulate plant growth at an organ scale during the organogenesis process. To date, GreenLab has been successfully applied to various species of agricultural crops (Guo et al., 2006; Kang et al., 2012; Qi et al., 2010; Vavitsara et al., 2017); its key advantage over other plant models, which are commonly limited to simulation, is its parametric identification (Christophe et al., 2008). Because of the mathematical formalism of GreenLab, hidden model parameters can be identified using inverse methods from measurement data (Guo et al., 2006; Zhan et al., 2003). Although FSPMs aim to simulate plant-level production in a mechanistic way, the sub-models that simulate a certain process, such as photosynthesis, sometimes take a simplified, empirical approach.

In predicting mass fluxes in the CELSS in previous work, photosynthesis and respiration reactions were modeled based on plant physiology and biochemical reaction knowledge, and the mass balance model for predicting total biomass and CO_2/O_2 concentrations was developed based on stoichiometric equations. However, no humans were involved in the closed system (Hezard et al., 2012; Maclean et al., 2010). Moreover, the developmental stages of plant were absent from the model, and consequently, it is difficult to demonstrate the long-term effects of plant behavior, extending from seedling to mature plant stages, on CO_2/O_2 concentrations.

In this study, we proposed a novel knowledge-and-data-driven modeling (KDDM) approach for simulating plant growth and the dynamics of CO_2/O_2 concentrations in a CELSS that includes plants and humans. This model consists of a 'knowledge-driven (KD)' sub-model and a 'data-driven (DD)' sub-model. The KD sub-model is a combined model of GreenLab and TomSim (GreenLab+). The DD sub-model is a piecewise linear model (PLM) of the CO_2 production and O_2 consumption by the crew member. The two sub-models were integrated through a mass balance model with metabolic stoichiometries, which were derived for CO_2/O_2 concentrations in a closed system. A three-step parameter estimation method was developed to identify the proposed model parameters. Finally, the KDDM approach was evaluated using real data from plant cultivation experiments in a closed system of plants and humans.

2. Materials and methods

2.1. Plant materials and measurements

The data were collected from a two-person, 30-day CELSS integrated test from Nov. 1st to Dec. 1st, 2012 in Beijing, China (Guo et al., 2014b). Lettuce (*Lactuca sativa* L. var. Dasusheng) was planted in the CELSS Integration Test Platform (CITP) of the China Astronaut Research and Training Center, in Beijing, China. The platform was tightly sealed and consisted of such elements as a plant cabin, crew cabin, temperature and humidity control system, plant illumination system, nutrient solution control system, effluent collection and disposal equipment; the volume and area of the CITP was 308 m^3 and 88 m^2, respectively. During the experiment, the cultivation area of the plant was 36 m^2, and the planting density was 56 plants m^{-2}. All of the plants were started from seeds and grew inside the plant cabin for their entire production cycle using a recirculating nutrient hydroponic technique. The Hoagland nutrient solution used nitrate as the sole source of nitrogen. The solution pH was automatically controlled between 6.15 and 6.45 with additions of 1 M nitric acid, and the electrical conductivity (EC) was maintained between 0.195 and 0.205 S m^{-1} with automatic additions of a concentrated stock solution. Light emitting diodes (LED) were used as light sources, which consisted of 90% red light (wavelength 637 nm) and 10% blue light (wavelength 465 nm). The photoperiod was 24 h with photosynthetically active radiation (PAR) of 500 μmol m^{-2} s^{-1} at a distance of 30 cm below the light source. The relative humidity was maintained between 64% and 76%. Water consumption and displacement were monitored and controlled, including water intake, urine, sanitary water, disposed and recycled effluent, and water condensate used for the nutrient solution; the effluent was disposed of and then partly recycled into the nutrient solution, and the condensate water was completely transformed into nutrient solution. The closure of air, water and food in the CITP were at 100%, 90% and 13.9% respectively, with the total material closure at 95.1%. On November 1st, when there were approximately 17 visible leaves, two crew members (male, 32 years, 170 cm, 72.0 kg; male, 38 years, 173 cm, 62.5 kg) entered the crew cabin, which was connected to the plant cabin through ventilation. Beginning on November 24th, a gas balance regulation test was performed (Table 1). The illumination area on the plants was adjusted by turning off a portion of the overhead LED lights to test the gas exchange with less plant photosynthesis.

The collected (hourly average) temporal data included air temperature, atmospheric pressure and CO_2/O_2 concentrations in the atmosphere of the cabin of CITP. During the 30-day experiment, the dry weights of the blades, petioles and stem were measured destructively during five stages along the growing period (Table 2). Furthermore, detailed topological observations were made on six plants twice a week, including the numbers of leaves and phytomer ranks (internode number counting from the base) on the main stem. For a more detailed explanation of the experimental setup of the environmental conditions and the crew members, please refer to Guo et al. (2014b).

Table 1
Setup of the gas balance regulation test.[a]

Duration	Illumination area on the plants
Before test	36 m^2
Day 24–27	24 m^2
Day 27–29	30 m^2
Day 29–30	27 m^2

[a] Each time, the illumination area on the plants was adjusted at 09:00 h by turning off a portion of the overhead LEDs.